**FACULTY
OF MATHEMATICS
AND PHYSICS**
Charles University

DOCTORAL THESIS

Tadeáš Bilka

Time-dependent CP violation in $B^0 \to \eta_c K^0_S$ decays at the Belle experiment, Alignment of the Belle II detector

Supervisor of the doctoral thesis: prof. RNDr. Zdeněk Doležal, Dr.

Study programme: Physics

Study branch: Subnuclear Physics

I declare that I carried out this doctoral thesis independently, and only with the cited sources, literature and other professional sources. It has not been used to obtain another or the same degree.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Sb., the Copyright Act, as amended, in particular the fact that the Charles University has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 subsection 1 of the Copyright Act.

In .... date ....
                                                    Author's signature

Věnováno babičce a dědovi.
Dedicated to my grandparents.

Title: Time-dependent CP violation in $B^0 \to \eta_c K_S^0$ decays at the Belle experiment,
Alignment of the Belle II detector

Author: Tadeáš Bilka

Institute: Institute of Particle and Nuclear Physics

Supervisor: prof. RNDr. Zdeněk Doležal, Dr., Institute of Particle and Nuclear Physics

Abstract:

This thesis deals with two independent yet closely related topics. In the first part, a measurement of branching fraction and time-dependent CP violation in $B^0 \to \eta_c K_S^0$, $\eta_c \to K_S^0 K^{\pm} \pi^{\mp}$ decays is performed. This decay allows access to $\sin 2\phi_1$, where ϕ_1 is an angle of the unitary triangle of the Cabibbo-Kobayashi-Maskawa quark mixing matrix. The measurement is based on the entire dataset of the Belle experiment, which consists of 772×10^6 B meson pairs collected at the KEKB e^+e^- collider. The extracted mixing-induced and direct CP-violation parameters read $\sin 2\phi_1 \simeq \mathcal{S} = 0.59 \pm 0.17$ (stat) ± 0.07 (syst) and $\mathcal{A} = 0.16 \pm 0.12$ (stat) ± 0.06 (syst), respectively. The measured product of branching fractions $\mathcal{B}(B^0 \to \eta_c K_S^0) \times \mathcal{B}(\eta_c \to K_S^0 K^{\pm} \pi^{\mp})$ is $(9.8 \pm 0.6$ (stat) ± 0.4 (syst) ± 2.3 (int)$) \times 10^{-6}$, where the last uncertainty accounts for interference with non-resonant background.

The second part deals with the alignment of the vertex detector and the central drift chamber of the Belle II experiment at the SuperKEKB collider, which is a next-generation Super-B-Factory. With the new pixel detector and the presented alignment method, Belle II achieves approximately twice better impact parameter resolutions than Belle. The presented alignment procedure involves a simultaneous determination of about sixty thousand parameters and accounts for time-dependent detector instabilities. The method is evaluated in simulations, and its performance is validated using data recorded by the Belle II detector and in the first world-leading physics measurements.

Keywords: Belle, CP violation, B-physics, $\sin 2\phi_1$, $\sin 2\beta$, η_c charmonium, Belle II, alignment, calibration, vertex detector, drift chamber, Millepede II

Contents

Preface

Since the discovery of the Higgs boson in 2012 [1], the Standard Model (SM) of particle physics can be considered completed[1]. All its fundamental particles have been observed, and all predictions to date have been confirmed by experiments, with only a handful of exceptions, where the statistical significance is still low to claim a ground-breaking discovery, or there are ongoing discussions about theoretical or experimental uncertainties. However, there is a general consensus that the SM is an incomplete theory, and some new physics beyond the SM is needed. Apart from the most apparent missing part—gravitation—the SM does not account for approximately 95% of the matter-energy content of the universe, composed of two elusive substances: dark energy and dark matter. Even ordinary (baryonic) matter has its secrets. One of the biggest mysteries is the absence of substantial amounts of its mirror counterpart—the antimatter—in the universe. A necessary condition for such a matter-antimatter asymmetry is the violation of charge-parity (CP) invariance at the level of fundamental interactions. While the SM does predict a small amount of CP violation, it is by many orders of magnitude smaller than what is required to match astronomical observations. In the SM, this asymmetry is tightly related to changes in the flavors of quarks, fundamental constituents of matter.

The experimental flavor physics program to observe and determine the amount of matter-antimatter asymmetry (initially, there were free parameters) predicted by the SM using decays of heavy B mesons was started by the construction of two B-Factory experiments in the 1990s. It concluded around 2010 when the Belle experiment [2] at the KEKB (KEK, Tsukuba, Japan) accelerator finished its data-taking. Together with the BaBar detector [3] at the PEP-II accelerator (SLAC, Stanford, USA), these B-Factories collected large datasets, which are still studied to obtain more results completing the extensive physics program, encompassing hundreds of analyses. Part II of this thesis represents one particular piece in this vast program.

The motivation for a new-generation B-Factory experiment, which could accumulate more than an order of magnitude larger dataset, has become even more urgent after several years of LHC operation, without any signs of supersymmetric particles to date [4]. It is very likely that the energy scale of New Physics (NP) beyond the SM is out of reach of the current and possibly even planned high-energy accelerators. Nonetheless, NP might reveal itself in other ways without the need for direct production of new heavy particles. These new particles and the corresponding fields would couple to the SM fields and enter as virtual particle contributions in Feynman diagrams, possibly modifying experimentally observable quantities. However, as mentioned, most measurements are still compatible with the SM predictions. Thus the NP contributions, if any, must be small and require more precise measurements.

---

[1]Neutrino masses and mixing parameters are still to be measured more precisely. One can consider these to be already new physics modifications.

Part III of this thesis describes an essential contribution to this new level of precision, from which many (mainly future) physics analyses will benefit. The larger datasets will reduce statistical uncertainties to the level where systematic errors could become a limiting factor. Therefore the more precise measurements will require better detector performance, better reconstruction and analysis algorithms and tools, and a more complex computing infrastructure to handle the enormous amount of data. An essential task in the quest for the best performance of the detector is its proper calibration. The model of the detector used for physics reconstruction has a tremendous amount of parameters, which are usually assumed to be known precisely in simulations. A very important subset of these parameters are values that describe the exact positions and orientations of the sensitive elements of the detector. The procedure to directly determine the optimal (ideally true) values of these parameters from the recorded data is called track-based alignment (or just alignment). The presented alignment method in the second part is crucial for reaching the ultimate precision for measuring parameters of charged particles.

During his bachelor's studies, the author joined the Belle II Collaboration, developing simulations for beam tests of the vertex detector. The need for an alignment method for such setups and work conducted during the author's master studies resulted in the first version of alignment for the complete vertex detector. Since then, this work has been extended to match the needs of a world-leading high-precision experiment. Belle II started taking the first physics data in 2019, and after further extensions, the author became responsible for all aspects of alignment of both the vertex detector and the central drift chamber. A high-precision alignment will be crucial for future time-dependent CP violation measurements. However, the matureness of the developed method was already proven by several validations and first measurements with the world's leading precision, namely the lifetimes of D^0 and D^+ mesons [5].

As the accumulation of a competitive dataset for time-dependent CP-violation measurements at the Belle II experiment was not guaranteed in a timely manner, the author was given the opportunity to study the existing unique data of the former Belle experiment and perform a physically relevant and competitive measurement in one of the decays, only measured by BaBar with the full integrated luminosity. The studied decay $B^0 \to \eta_c K_S^0$ belongs to the family of $b \to c\bar{c}s$-induced transitions, that allow to access $\sin 2\phi_1$ in a time-dependent analysis. The angle $\phi_1 \equiv \arg[-(V_{cd}V_{cb}^*)/(V_{td}V_{tb}^*)]$ is directly related to the elements of the Cabibbo-Kobayashi-Maskawa quark-mixing matrix. The underlying physics is similar to the $B^0 \to J/\psi K_S^0$ decay, often called the *golden channel* for $\sin 2\phi_1$ extraction. The analysis was started by Z. Drásal, who performed a similar measurement using the $\eta_c \to p\bar{p}$ decay. This work verified the proposed method and tools, which were adjusted for the $\eta_c \to K_S^0 K^\pm \pi^\mp$ decay channel. This decay mode offers smaller statistical uncertainty for $\sin 2\phi_1$ measurement but is more challenging due to a more complex background composition. This study was taken over by the author from Z. Drásal in the state of finished branching fraction measurement on the control channel with all necessary data (including simulations) already in a form suitable for fast analysis. While some of the previous work needed to be reproduced, the initial progress allowed the author to concentrate on work more relevant to physics, having many previously developed

tools at hand.

This thesis is divided into three parts. In part I, the Belle and Belle II experiments will be described in detail and in the context of general physics motivations together, as they share many features and concepts. A bit more emphasis is given to the achieved Belle performance, relevant for the Belle physics analysis part, while for the Belle II detector, more space is devoted to a description of sub-detectors relevant for the alignment part.

Part II is devoted to the physics analysis of Belle data. The phenomenon of *CP* violation and the particular physics measurement of $\sin 2\phi_1$ via time-dependent analysis will be discussed in detail in the theoretical introduction, followed by an exhaustive examination of the analysis method, intermediate results, cross-checks, validations, and finally, a *CP* violation measurement in $B^0 \to \eta_c(\eta_c \to K_S^0 K^+\pi^-)K_S^0$ decays[2].

For part III, devoted to Belle II detector alignment, we have decided to omit the usual pedagogical exposition in search of the optimal alignment configuration in the huge parameter space on simulations. This work was done in the past, and it became practically irrelevant given the constraints and results from real data. Instead, to keep a reasonable length of this thesis, the alignment will be described as a *final product*. In reality, the method's capabilities increased gradually from a couple of parameters in the vertex detector beam tests to about sixty thousand simultaneously determined values also describing every single wire of the drift chamber. Only one chapter will summarize the older history of the alignment, also mentioning past problems and their solutions. Most of the experimental content is devoted to more recent results and the performance of alignment configurations, which are being used for physics analyses and the first world-leading precision measurements.

---

[2]We will generally adopt implicit charge conjugation in this thesis, but occasionally it will be written explicitly, at least for charged particles.

Part I

Experiments and Detectors

CHAPTER 1

Introduction

While the Belle Experiment was explicitly designed to confirm the Kobayashi-Maskawa (KM) picture of the SM (see Sec. 4.3) and measure its fundamental parameters in B meson decays, its upgrade aims to probe physics beyond the SM by pursuing precision measurements in a very similar experimental setup, but with higher precision and using a much larger dataset. We thus attempt to describe both experiments compactly, highlighting the improvements.

The idea of B-Factories (or, more poetically, Beauty Factories) requires an electron-positron collider with asymmetric beam energies at a precisely defined center of mass energy of 10.58 GeV. This energy corresponds into the rest mass of the $\Upsilon(4S)$ resonance which decays to the desired pairs of B mesons. Due to a low cross-section (around 1 nb), the collider must have a substantial luminosity.

Initially, two such machines were constructed: PEP-II at SLAC Laboratory (Stanford, California, USA) and KEKB at the KEK Research Center in Tsukuba, Japan. These two colliders and their detectors BaBar [3] and Belle [2] accumulated two unique datasets, summing up to over 1.5 ab^{-1} (see Fig. 1.1 left), recorded also at various nearby $\Upsilon(nS)$ resonances and off-resonance. Most of the dataset is, however, collected at the $\Upsilon(4S)$ resonance, which for example, corresponds at Belle to 772 million $B\overline{B}$ decays. The KEKB collider held the record for the highest instantaneous luminosity of 2.11×10^{34} cm^{-2}s^{-1} for almost a decade.
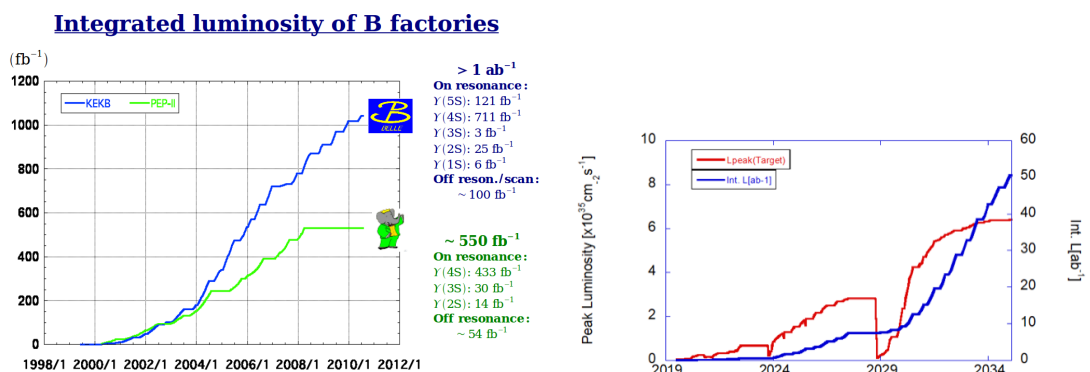
Figure 1.1: Integrated luminosity of the B-Factory experiments Belle (blue) and BaBar (green) [6] (left) and projection of instantaneous (red) and integrated (blue) luminosity for SuperKEKB and Belle II (right) [7].

The B-Factories have a rich legacy. They successfully observed CP violation in B-meson decays and confirmed the KM picture of the SM by measuring parameters of the unitary triangle, see Sec. 4.3. However, the physics reach was

much wider. These were ideal places to gain experimental insight into quarkonium spectroscopy by studying excited states of bottomonium and charmonium, which led to discoveries of a number of un-conventional states of bound quarks. A large physics program was also devoted to studies of charm mesons or τ lepton. The two experimental collaborations also established several analysis tools and techniques, which became standards in the field, from flavor tagging to blind analysis [8].

The need for more precise measurements resulted in proposals to construct a next-generation Super-B-Factory experiment. While the planned SuperB experiment in Frascati (Italy) [9] was not realized, several key ideas were employed in the construction of the high-luminosity KEKB upgrade. The aim is to collect 50 ab^{-1} by the year 2036. Such a large amount of data can only be collected in time if the projected increase in instantaneous luminosity (see Fig. 1.1 right) is achieved, targeting to surpass the KEKB world record by a factor of 30 [10].

The Belle II physics program is ambitious and diverse. The best reference is the Belle II Physics Book [11]. Let us at least briefly sketch a part of the landscape of the possible physics topics. The well-defined initial state of the e^+e^- collisions allows Belle II to handle inclusive modes and final states with neutral particles (π^0, K_S^0, γ, ...) or with missing energy due to undetected neutrinos. Especially its capabilities of neutral-particle reconstruction make Belle II complementary to the LHCb in many indirect searches for NP. In many other analyses, the two experiments will be competitive, which is welcome, as their sources of instrumental and other experimental uncertainties are very different.

The physics of the B decays naturally encompass the flag-ship measurements at a B-Factory. With the 50 ab^{-1}, the angles of the unitary triangle should be known to $\sim 1°$ ($\sim 0.4°$ for ϕ_1). The measurements of the sides of the unitary triangle will also improve. Of particular relevance are magnitudes $|V_{cb}|$ and $|V_{ub}|$, which represent fundamental inputs for the future precision studies. Belle II should help in resolving a long-standing discrepancy between inclusive and exclusive measurements of $|V_{cb}|$ and $|V_{ub}|$ in semileptonic B decays. Semileptonic B decays are also a hot topic in lepton flavor universality tests, and the tensions in ratios $R(D^{(\star)})$ in $B \to D^{(\star)}\tau\nu_\tau$ and $B \to D^{(\star)}\ell\nu_\ell$ decays, where $\ell = e, \mu$, should be definitely confirmed or rejected already with several ab^{-1} of data. Good efficiency for both muons and electrons will also be crucial for confirming similar tensions in lepton-flavor universality in penguin-mediated decays like $B^+ \to K^+\ell^+\ell^-$ [12]. Another important topic is charm-less B decays $B \to K\pi$, which can probe possible NP contribution to loop amplitudes, and there is a long-standing discrepancy known as the $K - \pi$ puzzle in the direct CP asymmetry in $B^0 \to K^+\pi^-$ and $B^+ \to K^+\pi^0$. Belle II will be unique in its ability to access the mode $B^0 \to K^0\pi^0$, needed to disentangle hadronic uncertainties. As one last example, let us mention rare B decays, like $B \to K^{(\star)}\nu\bar{\nu}$. The general idea behind such flavor-changing neutral processes is that their low branching fractions in the SM could be significantly enhanced by NP contributions, providing complementary tests of, e.g., the anomalies in $b \to s\ell^+\ell^-$ decays.

SuperKEKB is also a Charm-Factory and τ-Factory in a sense, as $e^+e^- \to \tau^+\tau^-$ and $e^+e^- \to c\bar{c}$ has a similar cross-section as $\Upsilon(4S)$ production at the center-of-mass energy of SuperKEKB. The program of charm physics involves precise measurements of D meson lifetime, D^0 mixing, and CP-violation param-

eters, or direct *CP* asymmetries. The τ physics program includes precise measurements of τ mass and lifetime, which are fundamental inputs to tests of lepton flavor universality. A wide range of topics are lepton-flavor-violating τ decays, like $\tau \to \mu\gamma$, which are forbidden or highly suppressed in the SM, but NP could enhance their branching fractions significantly.

A broad range of analyses is targetting the dark sector. The hermeticity of Belle II and the clean collision environment allow triggering, for example, events with a single hard photon from initial state radiation (ISR) followed by an NP process producing undetectable particles like a dark photon.

Finally, let us mention charmonium and bottomonium spectroscopy and studies utilizing special SuperKEKB runs at nearby Υ resonances, which shall continue a very successful program started at Belle focused on investigating exotic bound states of quarks [11].

The KEKB and SuperKEKB accelerators are discussed in the next Chapter 2. The detectors installed at the interaction point, where the electron and positron beamlines cross, are compared in Chapter 3.

CHAPTER 2

KEKB and SuperKEKB Colliders

SuperKEKB is constructed in the same tunnel as KEKB. It is located around 10 m underground, with four straight and four curved sections with a total length of about 3 km. The center of mass energy of $\sqrt{s} = 10.58$ GeV is achieved by colliding 8 (7) GeV high energy (HER) electron and 3.5 (4) GeV low energy (LER) positron beams at KEKB (SuperKEKB), which are supplied by a system of a linear accelerator and a damping storage ring for positrons. The beams are driven into the collision point at a large crossing angle of 22 mrad at KEKB [2], further increased to 83 mrad at SuperKEKB [10].
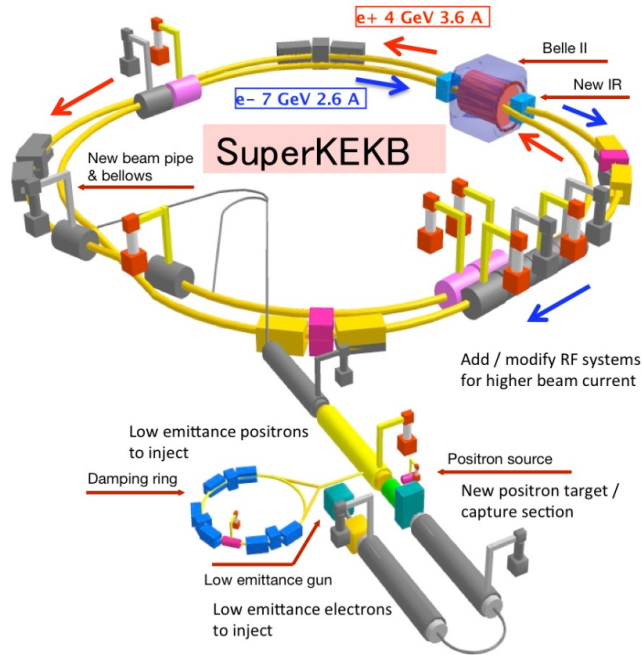
Figure 2.1: Schematic drawing of the SuperKEKB with upgraded or replaced parts highlighted in color [13].

The main achieved KEKB machine parameters are compared to SuperKEKB design values in Table 2.1. In Fig. 2.1, a schematic view shows the SuperKEKB accelerator and highlights systems that were replaced or upgraded from KEKB. The two major factors leading to the increased luminosity are approximately twice higher beam currents and, in particular, twenty times smaller vertical beta function $\beta_y^\star$ at the interaction point (IP). This decrease is possible thanks to adopting the so-called nano-beam scheme. The world's most complex system of

(LER/HER)	KEKB	SuperKEKB	unit
Beam energy (E)	3.5/8.0	4.0/7.0	GeV
Vertical β function at IP ($\beta_y^\star$)	5.9/5.9	0.27/0.30	mm
Half-crossing angle (ϕ)	11	41.5	mrad
Current (I)	1.6/1.2	2.8/2.0	A
Luminosity (L)	2.1×10^{34}	60×10^{34}	$\text{cm}^{-2}\text{s}^{-1}$

Table 2.1: Comparison of the main machine parameters of KEKB and SuperKEKB [10]. Some values are updated to recent un-official estimates according to discussion with experts.

superconducting magnets is designed to squeeze the beams just before the collision. The beam optics system (QCS) includes 55 superconducting coils (dipoles, quadrupoles, sextupoles, and compensating solenoids). A large crossing angle is needed to bring the focusing magnets closer to the IP. In addition, it helps to reduce beam-beam interactions due to faster separation of the particle bunches, allowing for higher beam currents.

The final luminosity of KEKB, surpassing its design requirements by a factor of two, was achieved by installing crab cavities, which rotate the bunches before the collision to increase the effective bunch overlap. The SuperKEKB collider reclaimed the world luminosity record on 15 June 2020, after a brief period when LHC became the most luminous machine in 2018, despite being a hadron collider. The current SuperKEKB luminosity record is 4.71×10^{34} $\text{cm}^{-2}\text{s}^{-1}$ (22 June 2022), still order of magnitude from the design value. Despite machine tuning being more challenging than initially foreseen, the luminosity is being gradually improved. The current integrated luminosity (see Fig. 2.2) is 427.79 fb^{-1}.

The beam asymmetry, which was reduced to limit intra-bunch (Touschek) scattering (limiting beam lifetimes), results in a lower boost of the B mesons. At Belle, their boost factor $\beta\gamma = 0.425$ resulted in a mean separation between the B meson decay vertices $\langle \Delta z \rangle \sim 200$ μm. This factor is reduced to $\beta\gamma = 0.28$ at SuperKEKB. This change is compensated by a smaller beam pipe radius and increased vertex resolution thanks to the installation of an additional high-precision vertex sub-detector closer to the interaction point. This upgrade is discussed among others in Chapter 3, devoted to the detectors.

2.1 Interaction Region and Beam Pipe

The nano-beam scheme results in a tiny vertical size (~ 50 nm at design values) of the interaction region at SuperKEKB, compared to KEKB (~ 90 μm), which is illustrated in Fig. 2.3. Also, the length of the beam overlap during collision is reduced from ~ 1.9 mm to ~ 100 μm[1].

The interaction region is surrounded by a specially designed section of the beam pipe. As the momentum spectrum of particles from B decays is rather soft, vertex resolution is limited by multiple scattering effects, which depend on the material budget of the beam pipe and the distance of the first layer of the vertex

---

[1]Current values for the longitudinal interaction region size are ~ 350 μm.
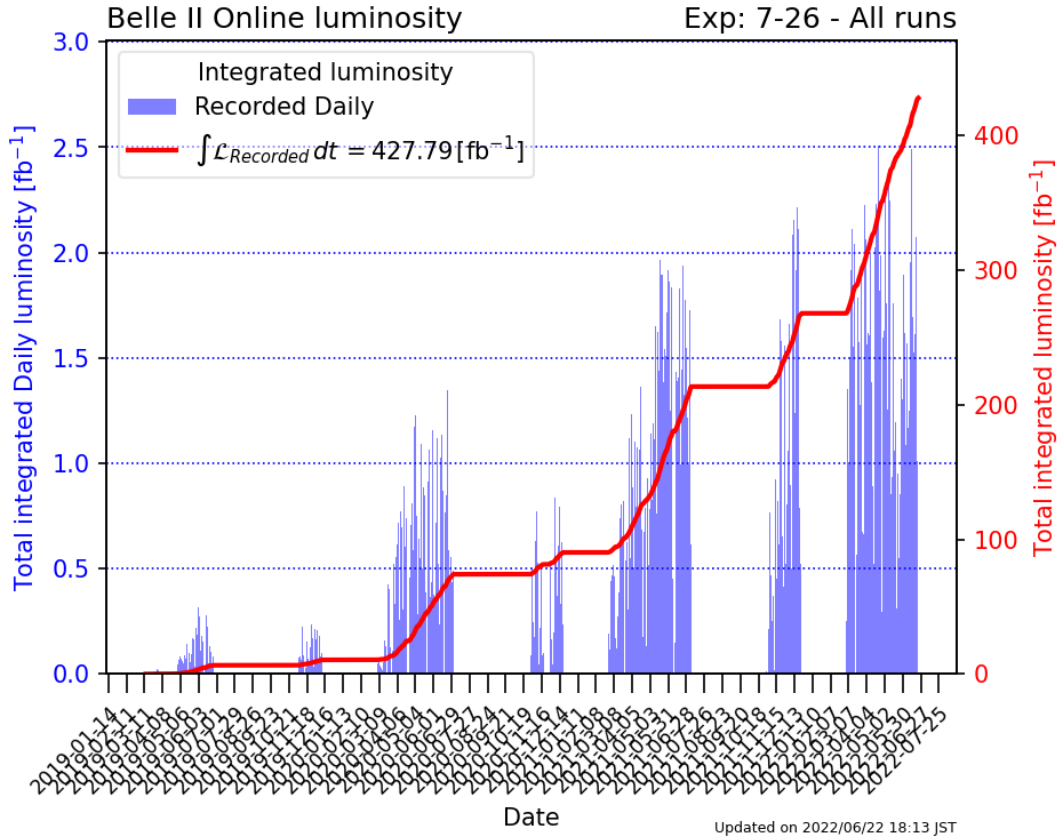
Figure 2.2: Integrated luminosity profile recorded by the Belle II detector at the SuperKEKB collider until 22 June 2022 [7].
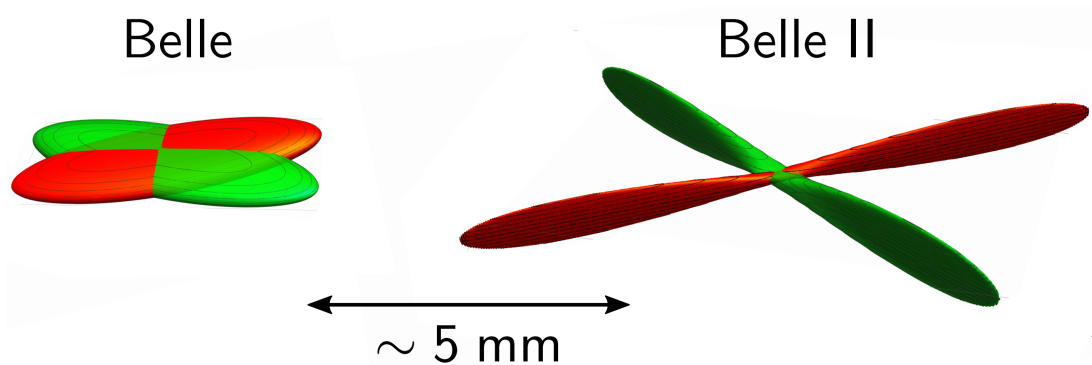
Figure 2.3: Illustration of the interaction region formed by the colliding bunches (without active crab cavities) at KEKB (left) and SuperKEKB (right).

detector from the IP. The beam pipe is thus made from two layers of Beryllium with a cooling gap in between. For additional shielding from X-ray photons, it is covered by thin 20 μm (10 μm) gold plating on the inside at Belle (Belle II).

At Belle, two beam pipe configurations were used for its two vertex detectors named SVD1 and SVD2. The initial beam pipe with an outer radius of 24.25 mm was replaced by a 16.25 mm beam pipe for SVD2. At Belle II, the beam pipe is even smaller, with an outer radius of 10 mm.

CHAPTER 3

Detectors Belle and Belle II

Four experimental halls are located in the straight sections of the main accelerator tunnel, see Fig. 2.1. The experiment is installed at the single point where the electron and positron beamlines cross, in the Tsukuba hall (4th underground floor). The Belle [2] and Belle II [10] experiments share the same concept of a multi-layered cylindrical multi-purpose particle detector with a 1.5 T solenoid magnetic field surrounding the region of colliding beams.

The z-axis of the coordinate system, with the center at the nominal interaction point, points along with the boost in the direction of the electron beam (HER). The vector of magnetic induction from the solenoid points along $+z$. The y-coordinate points upwards and x outside the collider ring, approximately in the direction of Mount Tsukuba (877 m). The cylindrical system is defined by longitudinal z-direction and radial direction $R = \sqrt{x^2 + y^2}$, azimuthal angle ϕ ($\phi = 0$ in $+x$ direction) and polar angle θ (with $\theta = 0$ in forward $+z$ direction). The design is forward-backward asymmetric with 92% solid angle coverage in the range $17° < \theta < 150°$. The two detectors are illustrated in Fig. 3.1.
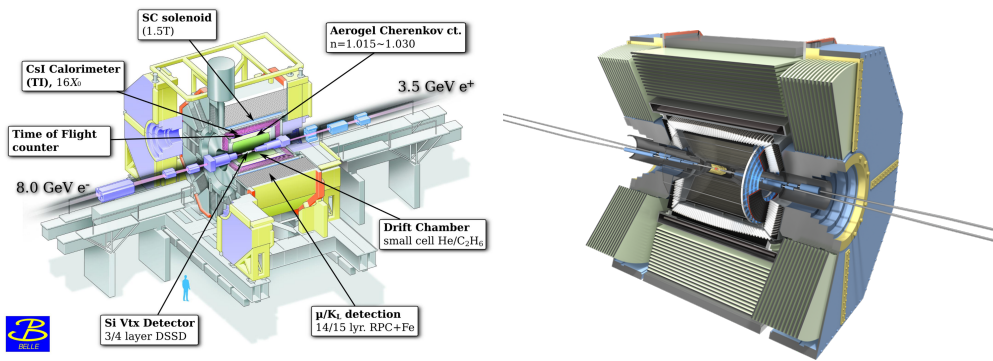
Figure 3.1: The detectors Belle (left) [14] and Belle II (right) [13] with cut section showing the individual sub-detectors.

The detectors are composed of a vertex detector, Central Drift Chamber (CDC) as the primary tracking device, particle identification system based on Cherenkov light emission (and time-of-flight estimation at Belle), Electromagnetic Calorimeter (ECL), superconducting solenoid with 1.5 T field and K_L and Muon detector (KLM), housed in the magnetic flux return yoke.

While all sub-detectors have been upgraded in some way, several subsystems were replaced completely in Belle II. This includes mainly the new vertex detector, but also the CDC is entirely new, as well as the particle identification systems and the end-cap ECL. This is also illustrated in Fig. 3.2, where the replaced parts are highlighted. The individual sub-detectors and the upgrade are discussed in

the following sections in detail, as well as the upgrade of the trigger and data acquisition system.
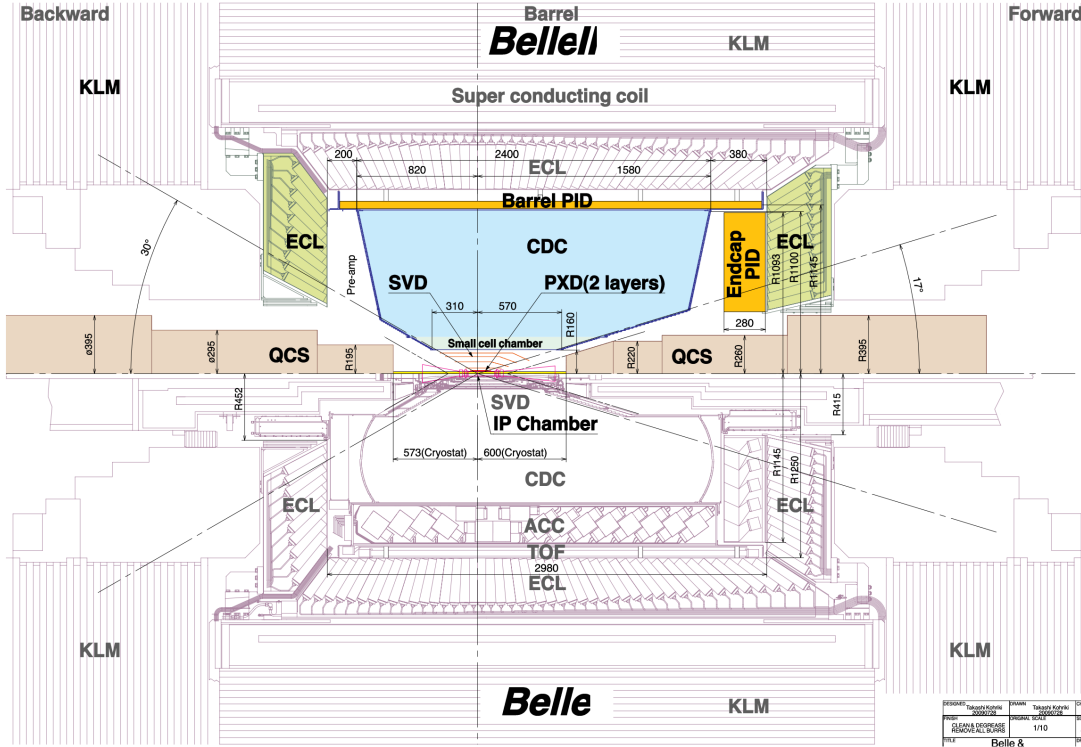
Figure 3.2: Schematic cross-section of the Belle II (top) and Belle (bottom) detectors. The replaced parts are highlighted in color [10].

3.1 The Vertex Detectors of Belle and Belle II

The vertex detector plays a crucial role in precision time-dependent CP-violation or lifetime measurements. Its primary role is the reconstruction of the decay vertices of B mesons. Thus it must be located as close as possible to the interaction point, simultaneously being able to withstand the high radiation environment. It must also be lightweight, with minimal radiation length to not disturb particle trajectories, while still generating sufficient signal for very high detection efficiency.

The detection method is therefore based on silicon semiconductor technology. The basic principle lies in a pn-junction, which is reverse-biased by a large enough voltage[1] to entirely deplete its bulk volume made from an n-type semiconductor. This volume is then sensitive to the passage of ionizing particles. As they cross the depleted volume, electron-hole pairs are created (one per each 3.6 eV of deposited energy) and immediately separated by the bias voltage. This process results in a current flowing over the junction. This current can be amplified, measured, and converted into a digital signal. Decoding the position of detection sub-elements like a strip or pixel allows to read out the position of the particle's trajectory intersection with the bulk. Usually, also the information about the respective

---

[1]Typical values are in tens of Volts, but depend, e.g., on sensor irradiation.

deposited charge can be retrieved. Various algorithms then can be applied to improve the estimate of the intersection using signals from neighboring channels.

At Belle, the Double-sided Silicon Strip Detector (DSSD) technology is employed. Originally three detection layers (SVD1) were used, but after problems with radiation hardness, the majority of the Belle dataset is taken with an upgraded SVD2 with four layers. Belle recorded $152 \times 10^6 B\overline{B}$ pairs in the SVD1 configuration (SVD1 experiment) and the remaining $620 \times 10^6 B\overline{B}$ pairs with the upgraded SVD2 vertex detector (SVD2 experiment). At Belle II, the Silicon Vertex Detector (SVD) with four layers and DSSD technology is used as well, but this time completely redesigned to accommodate the new inner-most detector: the Pixel Detector (PXD).

3.1.1 SVD1 and SVD2 at Belle

The Belle SVD1 detector was replaced by the SVD2 in 2001. With about twice more modules, the readout channels from several modules needed to be combined to fit into the existing readout. This then required full tracking to be used for position determination in SVD2. The first of the four SVD2 layers was also placed closer to the IP with a radius of 20 mm instead of 30 mm in SVD1. The parameters of the two configurations are compared in Table 3.1. In Fig. 3.3, the SVD2 detector is visualized.
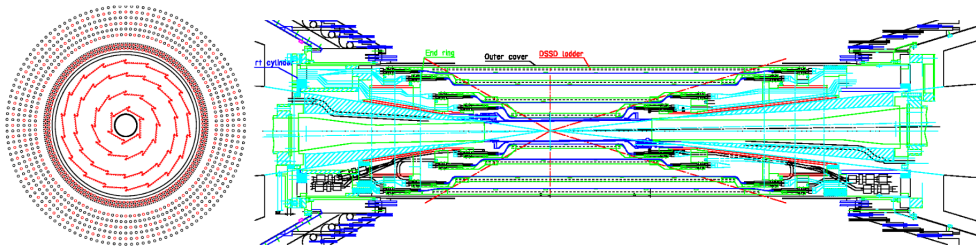
Figure 3.3: Front (left) and side (right) view illustration of the SVD2 detector. The front view also includes several layers of the Belle CDC wires [15].

The upgrade also led to improved impact parameter resolutions. In Fig. 3.4, these are shown as a function of pseudo-momentum disentangling angular dependence of multiple scattering effects (see Sec. 17.1.3 for more details). Split cosmic ray tracks were utilized for this estimate. The final resolution for SVD2 in the $R - \phi$ and z-direction can be parametrized as [16] [2]

$$\sigma_{R-\phi} = 17.4 \oplus 34.3/(p\beta \sin^{3/2}(\theta))\mu m$$
$$\sigma_z = 26.3 \oplus 32.9/(p\beta \sin^{5/2}(\theta))\mu m, \tag{3.1}$$

where p is particle momentum in units of GeV/c and β its velocity in units of c. The first term determines the actual detector resolution, while the second term parametrizes the material budget and related multiple-scattering effects depending on the (pseudo) momentum.

---

[2] $\oplus$ denotes a quadratic sum, such that $\sigma = a \oplus b$ means $\sigma = \sqrt{a^2 + b^2}$

	Belle SVD1	Belle SVD2	Belle II SVD	unit
Layers	3	4	4	
Layer radius	30/45.5/60.5	20/43.5/ 70/88.8	38/80/ 115/140	mm
Ladders in layer	8/10/14	6/12/18/18	8/10/14/17	
Sensors in ladder	2/3/4	2/3/5/6	2/3/4/5	
Num. of modules	102	246	187	
Sensor width	32	25.6 (33.28 for L4)	57.6 (38.4 for L3) 38.4–57.6 slanted	mm
Sensor length	54.5	76.8 (74.75 for L4)	122.8	mm
Sensor thickness	300	300	320 (300 slanted)	μm
Pitch in $R - \phi$	25	50 (65 for L4)	75 (50 for L3) 50–75 slanted	μm
Pitch in z	84	75 (73 for L4)	240 (160 for L3)	μm

Table 3.1: Comparison of parameters among the Belle SVD1 and SVD2 and the Belle II SVD detectors [2, 10].

3.1.2 Belle II SVD

The Belle II SVD detector is illustrated in Fig. 3.5. The most obvious change is the addition of non-rectangular slanted sensors in the forward region for the outer three layers. These improve the vertex resolution by limiting the effective amount of material crossed by particles from the IP. The inner and outer radius of 38 mm and 140 mm, is determined by the outer PXD and inner CDC radius, respectively. At Belle II, the SVD has an additional function of extrapolating tracks into the PXD, allowing the tracking to operate even in conditions of high luminosity background. SVD is also crucial for reconstructing long-lived particles, like K_S^0, which decay outside of the PXD. The fine time resolution of SVD (~ 3 ns) can also be used to suppress off-time background hits.

The sensors with a thickness of $300 - 320$ μm have a large area, close to current industrial production limits, and are organized in four layers and 49 ladders, slightly overlapping in $R - \phi$. The pitch is $50 - 75$ μm in $R - \phi$ and $160 - 240$ μm in z. The main parameters are summarized and also compared to Belle SVD in Table 3.1. Specific parameters are used for the inner SVD layer (L3) and the slanted forward sensors. The slanted sensors have a trapezoidal shape and variable pitch size in $R - \phi$.

3.1.3 Belle II Pixel Detector and its Upgrade

With its first layer just 14 mm from the interaction point, the pixel detector has to face several challenges in the high background environment caused mainly by QED processes at very high luminosities. For precise B vertex reconstruction, a very low material budget is a must, together with a sufficient radiation tolerance. All these requirements are fulfilled by the DEPFET (DEPleted Field Effect Transistor) technology chosen for the Belle II pixel detector. DEPFET was invented by Josef Kemmer and Gerhard Lutz in 1987 [18], and its development and
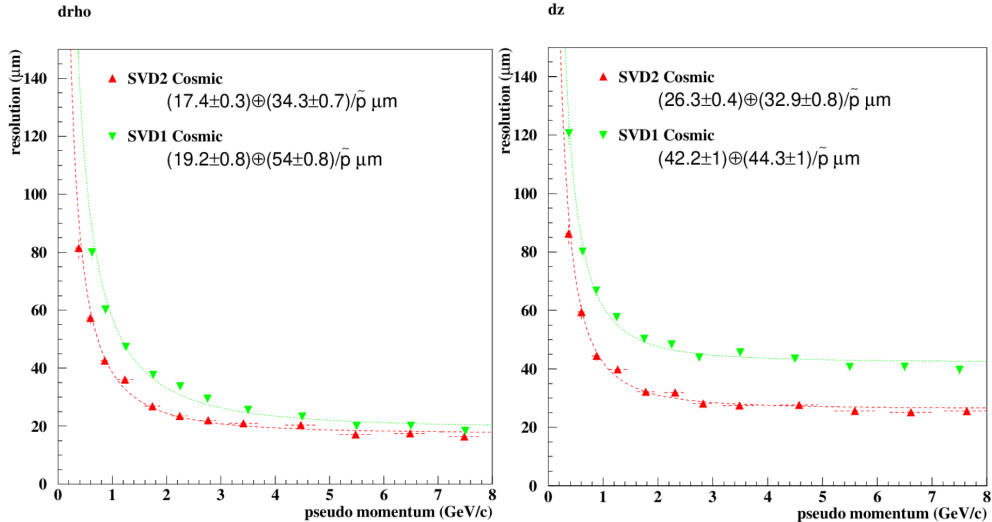
Figure 3.4: Impact parameter resolution determined by cosmic ray tracks for Belle SVD1 and SVD2 vertex detector configuration in $R - \phi$ (left) and z (right) [16].

fabrication are maintained by the Semiconductor Laboratory of the Max Planck Institute for Physics.

DEPFET pixels, illustrated in Fig. 3.6, integrate signal detection and amplification directly onto the sensitive depleted n bulk in the lithographic process. A Field Effect Transistor (FET) is fabricated just over a small Internal Gate (IG) region. The IG is about 1 μm under the FET of each pixel. It is created using enhanced n-type doping surrounded by a p-type doped region. When electron-hole pairs are created in the depleted volume by a traversing particle, holes drift to the bottom $p+$ contact, while electrons are forced to accumulate in the internal gate region, which forms a potential minimum thanks to the complex doping profile. Electrons are kept in the IG until a pulse voltage is applied at $n+$ clear regions just next to the FET. This overcomes the potential barrier and sucks the electrons out of the IG, making the pixel ready for subsequent detection.

The readout of the signal from each pixel is initiated by applying a voltage across the transistor. The current over the transistor is proportional to the charge stored in the IG. This readout is non-destructive and can be repeated as the charge is kept in the IG until the clear signal. Thanks to a low capacitance of the IG, low noise operation can be achieved at room temperatures with significant internal amplification, leading to currents of approximately 500 pA per each collected electron in the IG.

The high signal-to-noise ratio (over 30) permits the fabrication of very thin devices with a high detection efficiency. The Belle II PXD uses sensors with active areas thinned to 75 μm, embedded in a non-thinned 525 μm silicon support frame from the same wafer. This frame hosts switcher chips along the sensor length, steering the readout of pixel signal in a rolling shutter approach. Another set of chips is housed on the balcony of the sensor, outside of detector acceptance, which provides signal digitization and initial processing. The readout of the entire matrix, having up to 1600 rows, can be finished within 20 μs.

The originally proposed Belle II PXD is visualized in Fig. 3.6. It was designed
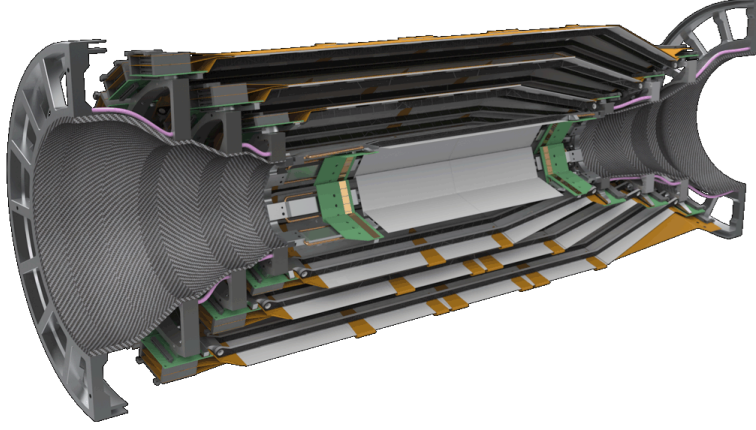
Figure 3.5: 3D visualization of the Belle II SVD detector, showing a side view of its four layers and the support structures. The forward sensors (at $+z$) are located on the right [17].
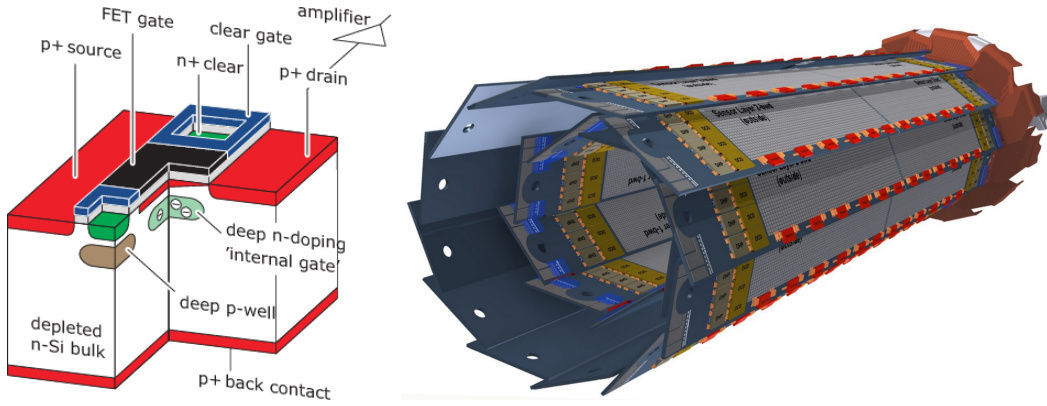
Figure 3.6: Schematic cross-section of the DEPFET pixel (left) and visualization of the complete Belle II PXD (right) [10, 19].

with two layers at a radius of 14 mm and 22 mm with 40 sensors in total. The sensors in the forward direction are segmented, with a smaller pitch size of 55 μm in the z-direction closer to the IP. The pitch size in $R-\phi$ is 50 μm. The parameters of the sensors are summarized in Table 3.2. Due to problems in ladder assembly resulting in a limited number of sensors available at the construction time, only the first layer of the PXD was completed. Only four sensors are installed from the second layer at ladder positions to cover a lower quality sensor in the first layer. This sensor finally turned out to be defective and is disabled. Thus the vertex performance of the current PXD is expected to be slightly worse than in design simulations. This important characteristic of the current PXD will be evaluated and compared to Monte Carlo (MC) simulation expectations in the alignment part of this thesis, see Sec. 17.1.3. At this point, let us show the expected design performance of the complete vertex detector in Fig. 3.7. The critical point is a significant improvement to the achieved Belle resolution by about a factor of two. An impatient reader can find the corresponding comparison for real Belle II data after all our alignment efforts at the very end of this thesis in Fig. 20.1.

The incomplete PXD will be entirely replaced by its original design configu-
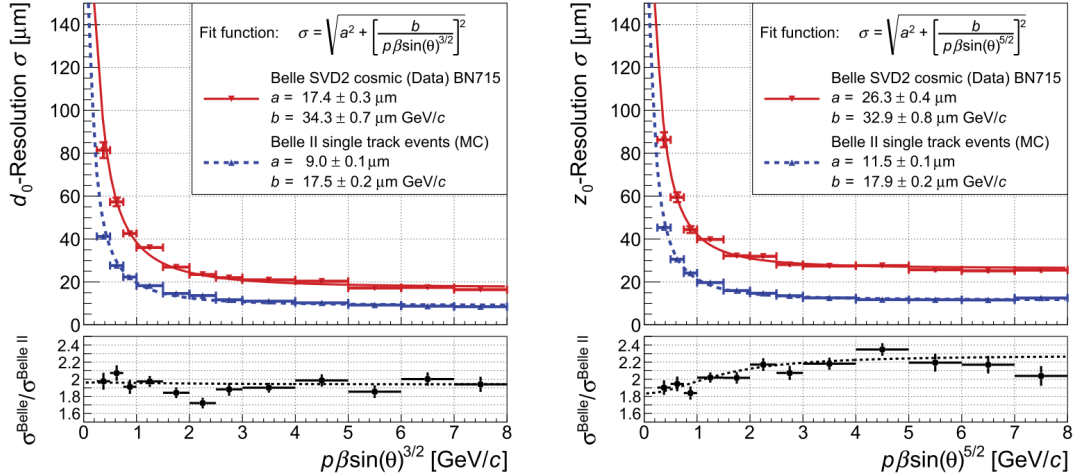
Figure 3.7: Impact parameter resolution in $R - \phi$ (left) and z (right) compared for Belle cosmic data and design Belle II simulation with particle gun generating single muons from the IP. In each bin, the resolution is estimated as a σ of a Gaussian fit to central 90% of the distribution of difference of track parameters of the split cosmic trajectory (Belle) or difference of reconstructed track parameter from the generated one (Belle II) [11].

layer	radius (mm)	ladders	sensors/ ladder	total sensors	pitch $R - \phi$ (μm)	pitch z (μm)	width (mm)	length (mm)
1	14	8	2	16 (15)	50	55, 60	12.5	44.80
2	22	12 (2)	2	24 (4)	50	70, 85	12.5	61.44

Table 3.2: Main parameters of the PXD and its sensors. The numbers in parenthesis refer to the initial Belle II PXD configuration with only a partially equipped second layer and a disabled sensor in the first layer. Numbers separated by comma refer to the backward and forward parts of the segmented forward modules, respectively [10].

ration during an extended shutdown, currently scheduled for 2022 or 2023. The shutdown period will include minor upgrades to other detector or accelerator parts. Besides slightly improved vertex resolution, the complete second layer will be needed to keep high tracking efficiency in the increasing background conditions at high luminosities.

3.2 Central Drift Chamber

The CDC is the primary tracking device of the Belle II detector. Besides track recognition, it is necessary to measure particle momenta in the 1.5 T magnetic field. CDC provides signals for the trigger and is used for particle identification via dE/dx measurements.

The CDC is a wire chamber filled with a gas mixture of helium (50%) and ethane (50%). Charged particles ionize the gas, and a large potential difference between field and sense wires leads to a fast separation of electrons and ions. Electrons drift to the sense wires (made from gold-plated tungsten) with a very
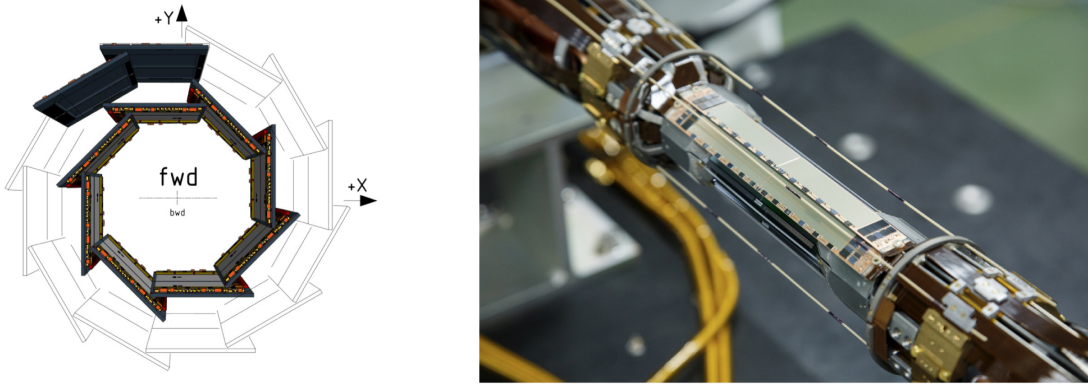
Figure 3.8: Schematic drawing (left) of the currently installed PXD detector and photography (right) of the actual device before integration into the SVD in 2018 [20].

small diameter of 30 μm, leading to very strong electric fields. Electrons are accelerated with these fields and ionize additional gas molecules, resulting in a charge avalanche. This process results in an electronic pulse registered by the readout electronics, with amplitude proportional to the charge deposited in each cell. Fig. 3.9 (right) illustrates such cells for the Belle II CDC. Each cell is formed from a sense wire surrounded by eight field wires, generating the desired electric potential profile. The cells are organized into layers and super-layers with alternating axial and stereo wires. Axial wires point along the z-coordinate, while the stereo wires are slightly rotated by a stereo angle of several tens of mrad to allow for a crude determination of the z-position (about 1–2 mm) of the track and thus its polar angle to infer the longitudinal momentum component.
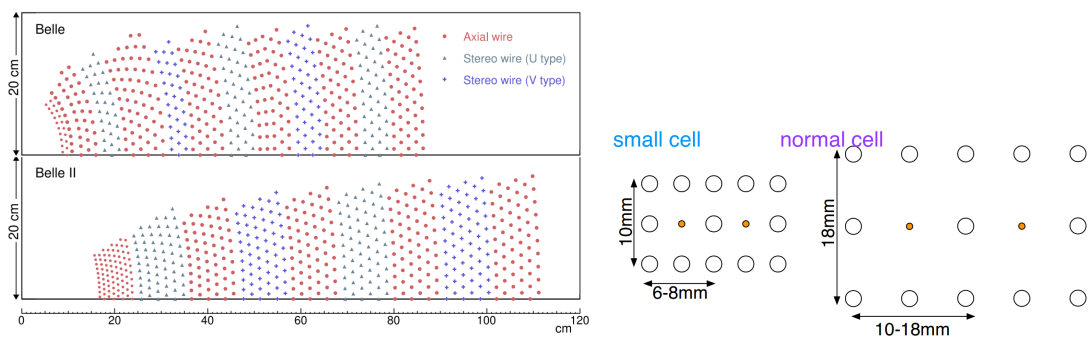
Figure 3.9: Comparison of super-layer, layer, and wire layout in the Belle and Belle II CDC (left) [21] and detail of the configuration of the small inner cell and normal cell with sense wire (orange) and field wires in the Belle II CDC (right).

The wire configuration is compared for the Belle and Belle II CDC in Fig. 3.9 (left) and the main parameters of the two drift chambers are summarized in Table 3.3. The Belle II CDC is equipped with a larger small-cell chamber to ensure good tracking in the much larger background. The spatial $R-\phi$ resolution is approximately 130 μm and 100 μm in the best region for the Belle and Belle II CDC, respectively. The relative transverse momentum resolution of the Belle
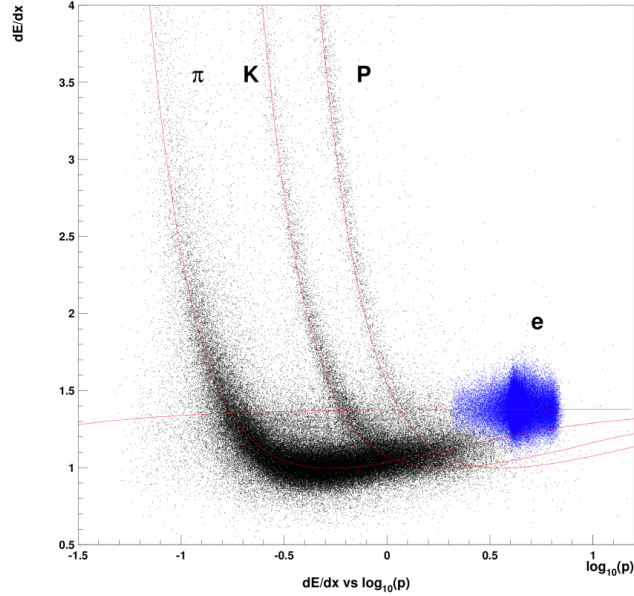
Figure 3.10: Energy losses dE/dx as a function particle momentum as measured in the Belle CDC [14]; p is in units of GeV/c.

	Belle CDC	Belle II CDC	unit
Radius of inner cylinder	77	160	mm
Radius of outer cylinder	880	1130	mm
Radius of innermost sense wire	88	168	mm
Radius of outermost sense wire	863	1111.4	mm
Number of layers	50	56	
Number of sense wires	8400	14336	
Diameter of sense wires	30	30	μm

Table 3.3: Main parameters of the Belle and Belle II CDC [10].

CDC with the SVD was measured to be

$$\frac{\sigma(p_t)}{p_t} = 0.127\% \, p_t \oplus 0.321\% \beta \tag{3.2}$$

where β is particle velocity in units of c and p_t is in units of GeV/c. The Technical design report [10] assumes

$$\frac{\sigma(p_t)}{p_t} = 0.1\% \, p_t \oplus 0.3\% \beta \tag{3.3}$$

can be achieved in combination with the vertex detector for the Belle II CDC. We will return to evaluating this important performance characteristic in the alignment part, see Sec. 17.5.

The total collected charge of the CDC hits associated with a track, and its estimated momentum can be used for particle identification via ionization losses.

Different particle species have distinct momentum dependence on their energy losses. Despite a sizeable stochastic spread of the energy losses, in some momentum regions, the distinctions can be unambiguous, as illustrated for Belle CDC dE/dx measurements in Fig. 3.10.

3.3   Particle Identification System

The primary purpose of the subsystem for particle identification is to separate kaons from pions, but when its information is combined with CDC dE/dx, it contributes to a general particle identification scheme, assigning likelihoods to different particle hypotheses. Particle identification sub-detectors have been completely replaced and redesigned for the Belle II detector, partially to host the new CDC with a larger radius but mainly to withstand higher occupancy and improve radiation tolerance and physics performance.

At Belle, the Aerogel Cherenkov Counter (ACC) and Time of Flight counter (TOF), labeled in Fig. 3.2, form the particle identification system. The ACC uses Cherenkov light emitted by charged particles traveling faster than the speed of light in a medium, which depends on its refractive index. The ACC works as a threshold Cherenkov counter providing K/π separation in the momentum range from 1.2 to 3.5 GeV/c.

In the Belle TOF, the flight time to reach the detector for particles coming from the IP is measured by modules with plastic scintillators attached to photomultipliers, providing a time resolution of 100 ps, also used as a fast timing signal for the trigger. With known flight time, momentum, and flight length (~ 1.2 m), the mass of the particle can be estimated, which is illustrated in Fig. 3.11. TOF provides K/π separation for momenta < 1.2 GeV/c.

The achieved combined (with CDC) kaon identification efficiency is about 85% while keeping lower than 10% pion misidentification rate at Belle [15], as also shown in Fig. 3.11.

At Belle II, particle identification in the forward end-cap region is achieved by proximity-focusing Aerogel Ring-Imaging Cherenkov detector (ARICH). The Belle TOF is replaced by the Time-Of-Propagation counter (TOP) in the barrel part. The TOP uses large quartz bars, where internally reflected Cherenkov photons are registered by a photomultiplier screen at the bar's end, as shown in Fig. 3.12. At Belle II, the expected (average) combined Belle II kaon efficiency is over 90% with less than 10% pion fake rate reaching to lower momenta than Belle, see Fig. 3.12 (right).

3.4   Electromagnetic Calorimeter

The primary purpose of the Electromagnetic Calorimeter (ECL) is to identify and reconstruct photons (especially those from π^0 decays) and electrons and their energy thanks to the generation and registration of electron showers in scintillator crystals. The crystals have a large electromagnetic radiation length ($\sim 16X_0$ for both Belle and Belle II) and serve both as an active medium for shower development and generation of scintillation light. As heavier particles (than electrons) have smaller specific ionization losses (and do not undergo bremsstrahlung) at
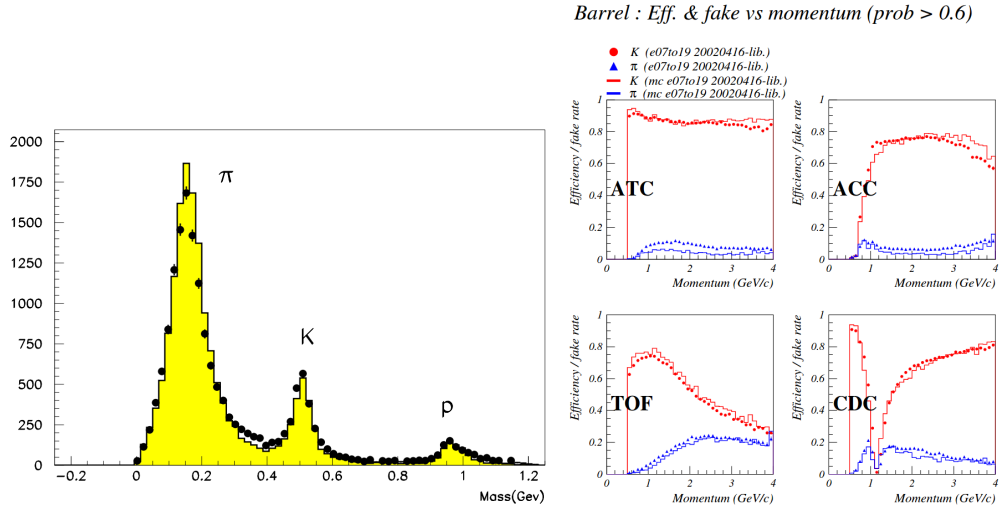
Figure 3.11: Left: Mass distribution reconstructed by the Belle TOF for particles with momentum less than 1.2 GeV/c for MC (yellow histogram) and data (black points). Right: Efficiency and fake rate for kaons at the Belle experiment separately for the ACC, TOF, and CDC and combined performance (denoted as ATC) [15]. The drop at ~ 1.1 GeV/c for the CDC can be understood from Fig. 3.10, where dE/dx measurements loose any separation power to distinguish pions and kaons around this momenta.
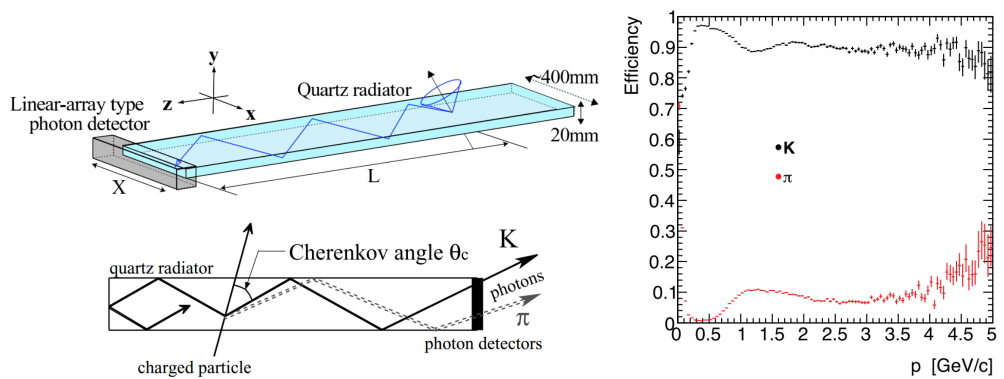
Figure 3.12: Left: Illustration of TOP quartz bar and the mechanism of Cherenkov light propagation to the photo-detectors [10]. Right: Simulated kaon efficiency and fake rate (for pions) for the combined Belle II particle identification [11].

31

given momenta, using momentum information from the CDC, the ECL can distinguish electrons from other particles by measuring E/p. The barrel part of the ECL is the same for Belle and Belle II; only the read-out is upgraded. CsI(Tl) crystals used for the whole Belle ECL are replaced by pure CsI crystals at Belle II in the forward end-cap part.
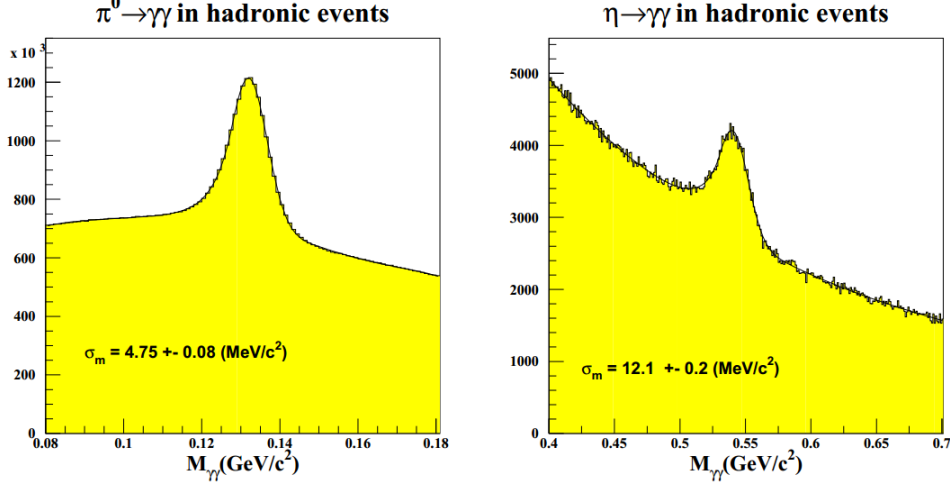
Figure 3.13: The reconstructed invariant mass of photon pairs for Belle data, showing clear peaks of decays of π^0 (left) and η (right) [2].
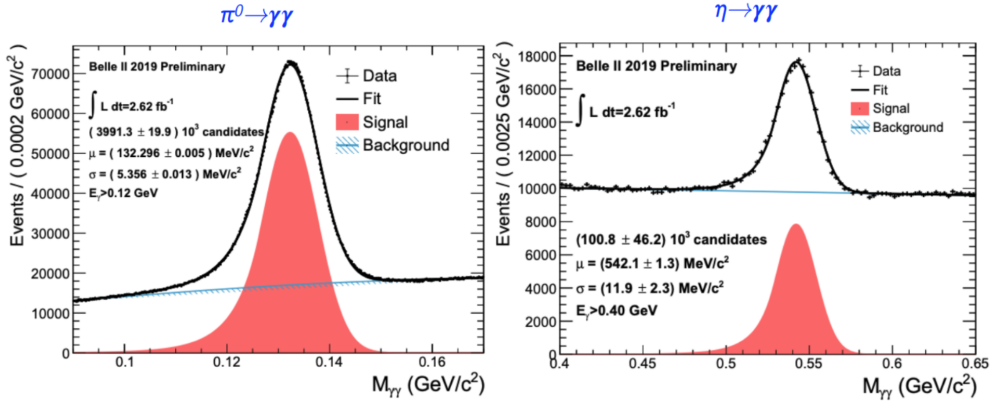
Figure 3.14: The reconstructed invariant mass of photon pairs for Belle II data, showing clear peaks of decays of π^0 (left) and η (right) [22].

The key calorimeter performance characteristic is energy resolution. This is demonstrated specifically in reconstruction of neutral particles like π^0 or η for Belle in Fig. 3.13 and for Belle II in Fig. 3.14. At Belle, the measured energy resolution of the ECL was

$$\frac{\sigma_E}{E} = \left(1.34 + \frac{0.066}{E} + \frac{0.81}{E^{1/4}}\right)\%, \tag{3.4}$$

while for the Belle II ECL, the expected resolution reads [10]

$$\frac{\sigma_E}{E} = \left(1.2 + \frac{0.2}{E} + \frac{1.6}{E^{1/4}}\right)\%, \tag{3.5}$$

where E is units of GeV.

a)

Ground plane	0.25 mm Mylar / 0.035 mm Copper
Dielectric foam	7 mm
Cathode plane	0.035 mm Copper / 0.25 mm Mylar
+HV	3.00 mm
Gas gap	2.00 mm
-HV	3.00 mm
Insulator	0.5 mm Mylar
+HV	3.00 mm
Gas gap	2.00 mm
-HV	3.00 mm
Cathode plane	0.25 mm Mylar / 0.035 mm Copper
Dielectric foam	7 mm
Ground plane	0.035 mm Copper / 0.25 mm Mylar
	31.6 mm total

b) Internal spacers — Gas input, +HV, Conducting Ink, -HV, Gas output — 220 cm — RPCs

c) 440 cm — Iron — 167-275 cm

Figure 3.15: Cross-section of the active layer of the RPC (a), a top view of the RPC with internal spacers for segmentation (b), and one block of two adjacent modules with the iron layer making one octant of the barrel KLM (c) [10].

3.5 K_L and Muon Detector

For both Belle and Belle II, the outermost sub-detector, located outside of the superconducting solenoid, is designed to detect muons[3] and long-lived hadrons, namely K_L^0 mesons. The K_L and Muon (KLM) detector is a sampling calorimeter divided into a barrel (BKLM) and end-cap (EKLM) part, with alternating 15 (14 in each end-cap) layers of sensitive elements and 4.7 cm thick iron plates, which also serve as a return yoke for the magnetic field flux. While muons will lose energy primarily due to ionization, K_L^0 can shower hadronically in the KLM material with a hadronic interaction length of 3.9 λ_0 (or ECL with 0.8 λ_0) while leaving no trace in CDC, allowing their identification for isolated KLM clusters.

At Belle, the whole KLM was instrumented with glass-electrode Resistive Plate Chambers (RPC). Ionized gas from hadron showers or muons in the gas gap leads to a discharge between the electrode plates at high voltage, registered as a signal in a segmented RPC module. One such RPC module with its layered structure is shown in Fig. 3.15.

For Belle II, the end-cap KLM and the first two BKLM layers are equipped with fast plastic scintillator strips instead of RPCs to keep good efficiency also in higher backgrounds, primarily caused by neutrons (e.g., from showers initiated by radiative Bhabha scattering)[11].

3.6 Trigger System and Data Acquisition

At a B-Factory, only a fraction of events registered by the detector is of interest. The trigger system must provide a very fast decision after the beam crossing to select events relevant to physics or detector calibration. The decision logic is divided into multiple layers with increasing complexity and latency. These events are then processed in the Data Acquisition System (DAQ), which reads out the

---

[3]with $p_t > 600$ MeV/c to reach it in the magnetic field

sub-detector information based on the decision logic output, builds the event, and sends it to an online computing farm for event monitoring and storage.

At Belle, the trigger starts with a hardware Level 1 (L1) trigger, which combines information from the sub-detectors looking for distinct signatures, like multiple charged tracks, typical for hadronic events. Its information is further passed to the Level 3 (L3) software trigger. A Level 4 trigger is applied offline to further reduce the number of events for long-term storage. The final decision on whether to keep or discard the event is performed by the Global Decision Logic (GDL). The efficiency of the Belle trigger for hadronic events was higher than 99 %. The trigger rate (for all triggered events) at regular operation was about 200 Hz, but over 500 Hz was stably achieved.
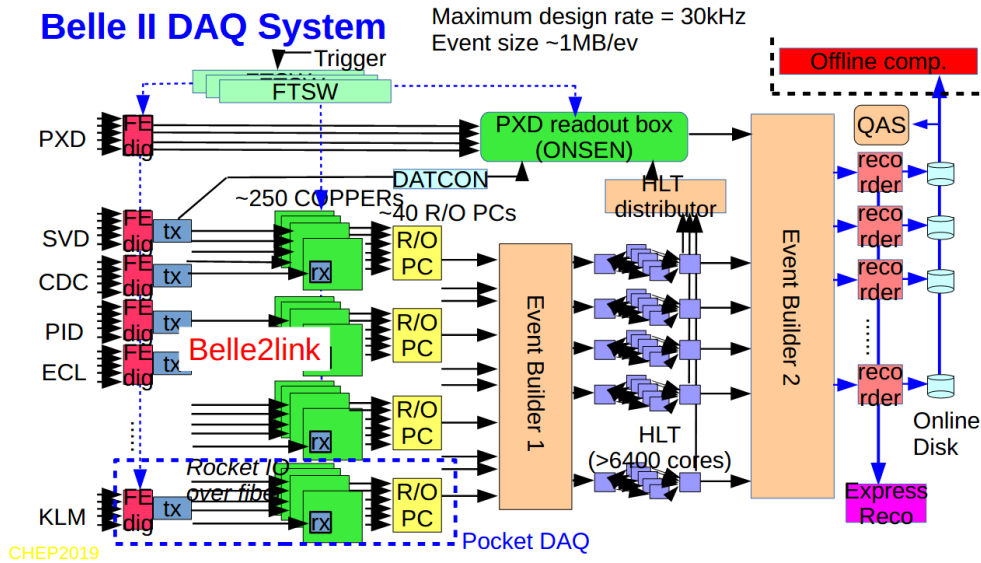
Figure 3.16: Schematics of the Belle II data acquisition system [23].

At Belle II, the trigger system is designed to withstand a trigger rate as high as 30 kHz. In addition, special selections open possibilities e.g. for dark sector (axion) searches via dedicated single (three) photon trigger [11]. The Belle II trigger is composed of a hardware L1 trigger, followed by the first event builder and software High-Level Trigger (HLT), reducing the event rate to about 10 kHz. In addition, HLT must identify the Regions Of Interest (ROI) in the PXD to allow for significant data reduction. After the HLT decision and inclusion of the PXD hits into the final event is performed at the second event builder, events are written to a disc array. A fraction of events is reconstructed by an express reconstruction chain to provide fast monitoring of online data quality. A (simplified) scheme of the Belle II DAQ is shown in Fig. 3.16.

Part II

Belle Physics Analysis

CHAPTER 4

Theoretical Introduction

What is the difference between left and right? Such an innocent question (surprisingly?) boils down to asking: "Why do we exist at all?" As will be briefly shown in the following sections, the symmetry between left and right is directly related to the symmetry between matter and antimatter. Antimatter, being first predicted theoretically by Dirac in 1928 [24], should be an exact equivalent of the standard matter. However, once these two interact, they annihilate typically into photons. According to our current knowledge, just after the Big Bang, matter and antimatter should have been balanced. However, with an exact symmetry between the two, their mutual interaction should lead to a universe very different from this one and almost certainly without us. Most of the antimatter seems to have disappeared. Literally, as far as we can see [1], only matter is present in macroscopic amounts. This suggests that there must be some underlying mechanism that violates matter-antimatter symmetry, known as *CP* violation, or charge-parity violation. Even a small *CP* violation could generate the present matter-antimatter unbalance, but all the known sources of *CP* asymmetry in the SM are too tiny. This motivates the search for yet undiscovered *CP*-violating phenomena sourced by physics beyond the SM.

In sections 4.1 and 4.2, we will look into the left-right and matter-antimatter symmetry, a related phenomenon of time reversal, and briefly summarize two major experiments which established the violation of these fundamental symmetries of nature. The present general formulation is given in Sec. 4.3 and after developing the theory of neutral B meson mixing and related *CP*-violation phenomena in Sec. 4.4, the SM prediction for our decay of interest is discussed in Sec. 4.5. The basics of experimental measurement principles can be found in Sec. 4.6. Finally, after some additional discussion of our particular decay category $B \to \eta_c(K_S^0 K^+ \pi^-)K$ in Sec. 4.7, the previous experimental measurements are discussed in Sec. 4.8, also in the context of new physics motivations for measurement of $\sin 2\phi_1$ in $B^0 \to \eta_c K_S^0$. As a last part of the theoretical introduction, the measurement methodology is briefly discussed in Sec. 4.9.

The body of the analysis part is separated into four main chapters. In Chapter 5, the signal event selection, detection efficiency, and continuum background suppression are described, and the other sources of backgrounds are briefly discussed. As we build up the data model in Chapter 6, the signal and background components are discussed in more detail before proceeding to a branching fraction measurement. A significant part of this chapter will evaluate the main systematic error stemming from signal and (non-resonant) background interference. In

---

[1]with our best instruments looking for possible sources of annihilation signals or antiparticles in cosmic radiation.

Chapter 7, we extend the data model to the time difference and flavor dimensions. After *CP*-violation measurement for the control channel and a number of analysis validations and cross-checks, we proceed to the final *CP*-violation measurement in Chapter 8 and conclude the analysis in Chapter 9.

4.1 Discrete Symmetries of Nature

Symmetries play a very important role in physics. Continuous symmetries imply conserved quantities according to the Noether theorem. This leads to the law of energy, momentum, and angular momentum conservation due to the general invariance of physical laws with respect to time, translation, and rotation, respectively. Abstract continuous symmetries are the basis for the formulation of the SM as a non-abelian gauge theory with local $U(1) \times SU(2) \times SU(3)$ symmetry. The three basic discrete symmetries can be divided into two already recognized in classical physics and one of purely quantum nature:

- **Time reversal** T, or time conjugation changes time $t \to -t$. It is well known that most laws of physics are generally invariant under time reversal, which leads to the famous problem of the arrow of time. Time usually enters in the form of derivative $\frac{d}{dt}$, and thus T leads to the reversal of motion. For position $\boldsymbol{x}$, momentum $\boldsymbol{p}$ and angular momentum $\boldsymbol{L} = \boldsymbol{x} \times \boldsymbol{p}$, we have

$$\boldsymbol{x} \xrightarrow{T} \boldsymbol{x} \qquad \boldsymbol{p} \xrightarrow{T} -\boldsymbol{p} \qquad \boldsymbol{L} \xrightarrow{T} -\boldsymbol{L}. \qquad (4.1)$$

- **Parity transformation** P switches sign of the coordinates in position vectors $\boldsymbol{x} = (x, y, z) \to -\boldsymbol{x} = (-x, -y, -z)$, which changes right-handed coordinate system into a left-handed one. This corresponds to a mirror reflection with additional rotations (and possibly a translation depending on the reflection plane). As symmetry for translations and rotations is conserved, the question of mirror symmetry can be reduced to parity invariance. One has

$$\boldsymbol{x} \xrightarrow{P} -\boldsymbol{x} \qquad \boldsymbol{p} \xrightarrow{P} -\boldsymbol{p} \qquad \boldsymbol{L} \xrightarrow{P} \boldsymbol{L} \qquad (4.2)$$

and can see that $\boldsymbol{L}$ is an example of an axial vector, which is invariant under parity. A product of axial and ordinary (polar) vector yields a pseudoscalar—a number defined in such a way that it changes sign under parity, as for example, the helicity $h = \boldsymbol{s} \cdot \boldsymbol{p} / |\boldsymbol{s} \cdot \boldsymbol{p}|$, which is a projection of particle spin $\boldsymbol{s}$ (which behaves as angular momentum $\boldsymbol{L}$, i.e. $\boldsymbol{s} \xrightarrow{P} \boldsymbol{s}$) onto its momentum.

Parity is also an intrinsic property of most particles meaning these are eigenstates of the parity operator: $\hat{P}|\psi\rangle = P_\psi|\psi\rangle = \pm|\psi\rangle$. Dirac equation predicts that fermions and anti-fermions have opposite parity and by convention $P_f = +1$ for fermions (e^-, p^+, n^0 or quarks). Bound states of a fermion and anti-fermion have parity $(-1)^{L+1}$, where L is their orbital momentum state. So, for example, the two bound states of a charm quark and anti-quark highly relevant for this thesis, η_c, and J/ψ, have $L = 0$ and thus odd parity ($P = -1$).

- **Charge conjugation** C does not have any classical nor non-relativistic quantum-mechanical analog. It is related to the fact that all particles have their corresponding antiparticles, which are related by a transformation changing the sign of all $U(1)$ charges, represented by additive quantum numbers: electric charge, lepton number, flavor, the third component of isospin and so on [25].

  Thus for a particle, the charge conjugation yields its antiparticle:

  $$C|\psi\rangle = |\overline{\psi}\rangle. \tag{4.3}$$

  Eigenstates of C with a well-defined intrinsic C-parity can only be (some) neutral particles (e.g. $C(\gamma) = -1$, $C(\pi^0) = +1$). Bound states of a fermion and anti-fermion have $C = (-1)^{L+S}$, where L is their orbital momentum state and S is the total spin. For example the pseudoscalar meson η_c with $S = 0$ (the quarks have antiparallel spins in antisymmetric combination) has $C(\eta_c) = +1$, while the vector meson J/ψ (with symmetric spin part of the wave function and $S = 1$) has $C(J/\psi) = -1$.

Each of these symmetries can be violated (not conserved in particle decays) separately, as well as combinations of them. However, based on very general requirements for a reasonable quantum field theory, in particular Lorentz invariance, the combined symmetry CPT should not be broken, as stated by the CPT theorem [25]. It follows that if, for example, the combined symmetry CP is broken, then T must be violated in order to preserve the CPT symmetry. Both C and P are conserved in the electromagnetic and strong interactions, and these symmetries were originally thought to be fundamental symmetries of nature. This assumption had to be abandoned in the light of the experimental results discussed in the following section.

4.2 Experimental Path to Violation of P, C and CP

The first experimental hints of parity violation came from the $\tau - \theta$ puzzle, in which particles nowadays known as charged kaons decayed into final states with two or three pions, which have different parity ($P(\pi) = -1$). The puzzle was resolved by proposing that the weak interaction, necessarily involved in decays of strange particles, does not conserve parity [26]. The proposed direct experimental test, which would not involve decays of strange particles (poorly understood at that time), was conducted by C. S. Wu in 1957 [27]. Until this experiment, the distinction between left and right was a mere human[2] convention. The weak decays of strongly polarized ^{60}Co nuclei clearly showed that β electrons are emitted in a direction opposite to the spin of the nucleus, in proportion to the polarization level[3]. The established parity (and charge) violation in the weak interactions came

---

[2]While the left-right symmetry is broken (almost) consistently in several natural phenomena, these are (most likely) not universal for life beyond Earth (if there is any).

[3]One could relate left and right (for example) to the direction of current in winding of the polarization magnet coil.

out to be, in fact, maximal. This manifests itself phenomenologically in the fact that only left-handed neutrinos and right-handed anti-neutrinos exist in nature[4]. Similarly, only left-handed particles (and right-handed antiparticles) participate in the charged-current weak interactions mediated by intermediate $W^{\pm}$ bosons.

While weak interactions violate P and C symmetry, the combined CP symmetry was originally believed to be conserved. Thus, the correct *mirror image* of physical laws needs to switch left with right and particles with antiparticles[5]. However, according to our current knowledge, no significant amounts of antimatter are present in the observable universe, so this symmetry must be violated, too. It was first pointed out by Sakharov in 1967 [28], that CP violation at the fundamental level is one of the necessary conditions to generate the observed abundance of matter over antimatter. At that time, CP violation had been, however, already established experimentally, to the surprise of most theorists [29].

Violation of CP was first discovered in 1964 in studies of neutral kaon decays. Neutral kaons, produced at that time in strong interactions, are not their own antiparticles, as not all of their additive quantum numbers are zero, namely the strangeness defined by (three times) the charge of the constituent strange quark ($K^0 = d\overline{s}$, $\overline{K}^0 = \overline{d}s$). This leads to a remarkable phenomenon of neutral meson mixing (which does not require CP violation), where a particle can turn into its antiparticles and vice versa. We will discuss this topic in detail in the context of B mesons later; see Sec. 4.4. CP was assumed to be fully conserved in weak interactions, which are the only decay mode that allows kaons to change strangeness to permit their decay. We have[6]

$$CP|K^0\rangle = |\overline{K}^0\rangle \tag{4.4}$$

and see that $|K^0\rangle$ is not a CP eigenstate. From an analogous relation for $\overline{K}^0$, one can, however, easily construct CP eigenstates as

$$\begin{aligned}
|K_1\rangle &= \frac{1}{\sqrt{2}}(|K^0\rangle + |\overline{K}^0\rangle) \\
|K_2\rangle &= \frac{1}{\sqrt{2}}(|K^0\rangle - |\overline{K}^0\rangle).
\end{aligned} \tag{4.5}$$

Thus $CP|K_1\rangle = +|K_1\rangle$ and $CP|K_2\rangle = -|K_2\rangle$ and only $|K_1\rangle$ should be able to decay into two pions ($CP|\pi\pi\rangle = +|\pi\pi\rangle$). The two states also differ radically in their respective lifetimes as K_2 must decay into three pions ($CP|\pi\pi\pi\rangle = -|\pi\pi\pi\rangle$), leaving only a very limited phase space in this very fortunate case of neutral kaons. These are nearly the true mass eigenstates, known as K_S^0 (K short) and K_L (K long) for kaons. For B mesons, this situation will be quite different due to their high mass and many possible decay modes.

---

[4]In the limit of zero neutrino mass.

[5]Thus an experimenter in a distant world from antimatter (using anti-^{60}Co and emitted positrons) would arrive at the opposite definition of left and right according to our instructions to replicate Wu's experiment.

[6]Up to a phase. Kaons as well as pions have an odd parity and the C parity phase can be chosen such that $C|K^0\rangle = -|\overline{K}^0\rangle$. Opposite conventions result in switching the labels for the states with very close masses or lifetimes and redefinition of their differences, which are the actual observable quantities [29].

In 1964, while trying to set an upper limit, the process $K_L^0 \to \pi^+\pi^-$ was actually discovered with a small branching fraction ($\sim 2 \times 10^{-3}$). This small *CP* violation requires that the actual K_S^0 and K_L^0 mesons (which are true mass eigenstates with definite masses and lifetimes) are, in fact, mixtures of two *CP* eigenstates, parametrized by a small complex mixing parameter ϵ as[7]:

$$
\begin{aligned}
|K_S^0\rangle &= \frac{1}{\sqrt{1+|\epsilon|^2}}(|K_1\rangle - \epsilon|K_2\rangle) \\
|K_L^0\rangle &= \frac{1}{\sqrt{1+|\epsilon|^2}}(|K_2\rangle + \epsilon|K_1\rangle)
\end{aligned}
\tag{4.6}
$$

Thus, for example, the real K_S^0 is mostly the *CP*-even state K_1, but with a small admixture of the *CP*-odd state K_2.

At this point, the theorist took the lead and tried to incorporate *CP* violation into the evolving Standard Model. In the following years, the strangeness of kaons was embodied in the framework of flavor physics, adding tastes of charm and beauty quarks to the big picture and completing the three generations of leptons and quarks with quite a late top quark observation in 1995 [30]. Interestingly, the third generation of quarks was postulated theoretically shortly after observing *CP* violation in neutral kaons. We will discuss this briefly in the next section, where the theory will be explained already in the context of the contemporary Standard Model of particle physics.

4.3   *CP* Violation in the Standard Model

The Standard Model incorporates three generations of fermions (leptons and quarks), the force-carrying vector bosons (photon, intermediate $W^\pm$ and Z^0 bosons, and gluons), and the scalar Higgs boson into the framework of quantum field theory. The particle content of the SM is shown in Fig. 4.1. The bottom quark is also known as the beauty[8] quark. We prefer to use this name or just b quark. The flavor quantum number distinguishes the six different quarks with an up-type and down-type quark in each generation. While the flavor is conserved in the electromagnetic and strong interactions, weak interactions allow quarks to transform into each other.

These transitions are governed by the charged-current part of the interaction Lagrangian between (left-handed) down and up-type quarks fields [32]:

$$
\mathcal{L}_{\text{CC}}^{(\text{quark})} = \frac{g}{2\sqrt{2}}\begin{pmatrix}\overline{u} & \overline{c} & \overline{t}\end{pmatrix}\gamma^\mu(1-\gamma_5)\underbrace{\begin{pmatrix}V_{ud} & V_{us} & V_{ub}\\ V_{cd} & V_{cs} & V_{cb}\\ V_{td} & V_{ts} & V_{tb}\end{pmatrix}}_{\boldsymbol{V}_{\text{CKM}}}\begin{pmatrix}d\\ s\\ b\end{pmatrix}W_\mu^+ + \text{h.c.},
\tag{4.7}
$$

---

[7]We will later utilize a different form of this expression using two complex parameters p_K and q_K related to ϵ, where $|K_S^0\rangle = p_K|K^0\rangle + q_K|\overline{K}^0\rangle$ and $|K_L^0\rangle = p_K|K^0\rangle - q_K|\overline{K}^0\rangle$

[8]According to author's knowledge, the Belle experiment name refers to the beauty (flavor) via the fictional character *Belle* in *Beauty and the Beast*. In one of the commissioning phases, the BEAST (Belle Exorcism for A STable experiment) detector was installed to measure accelerator backgrounds.

Standard Model of Elementary Particles
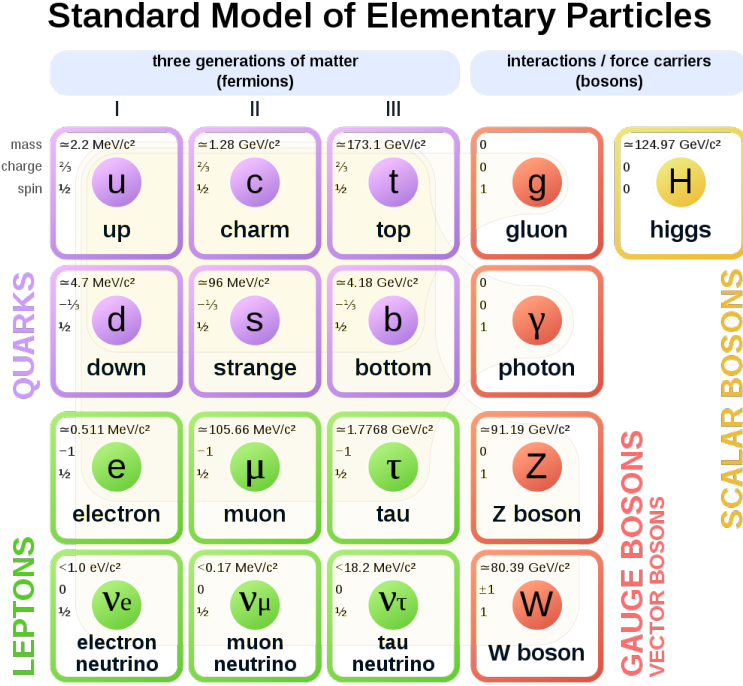
Figure 4.1: Fundamental particles of the Standard Model [31].

where g is the weak coupling constant, γ^μ are Dirac gamma matrices, W_μ^+ is the charged intermediate vector boson field, and $\boldsymbol{V}_{\mathrm{CKM}}$ is a unitary matrix with coefficients setting the mixing probability between each two up and down quark types. The CKM matrix emerges as a consequence of the diagonalization of independent mass matrices for the primordial up- and down-type quark fields, which appear after the Higgs field acquires a vacuum expectation value. Because the charged-current weak interaction mixes left-handed up- and down-type quarks, a non-trivial matrix ($\boldsymbol{V}_{\mathrm{CKM}}$) appears after transformation into physical quark fields.

In 1972, at the time when only three quarks were known, it was pointed out by Kobayashi, and Maskawa [33] that three generations of quarks could allow for *CP* violation in a 3×3 mixing matrix, which incorporates the original Cabibbo mixing angle $\theta_c \sim 13°$ from the three-quark times. With nine free parameters of a unitary matrix, of which five can be removed by a suitable choice of phases of the quark fields, four parameters of the CKM matrix need to be determined. However, only three angles are needed to define a real-valued unitary (orthogonal) matrix, so the CKM matrix must be complex. The additional single complex phase of the CKM matrix gives rise to all *CP*-violation phenomena in the quark sector.

Instead of the standard parametrization of the CKM matrix using three mixing angles and one *CP*-violating phase, a parametrization by Wolfenstein, which expands $\boldsymbol{V}_{\mathrm{CKM}}$ in terms of $\lambda = \sin\theta_c \sim 0.22$ [34] is more suitable for our needs. Using its three parameters $A \sim 0.79$, $\rho \sim 0.141$, $\eta \sim 0.357$, the CKM matrix reads

$$\boldsymbol{V}_{\mathrm{CKM}} = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \qquad (4.8)$$

The Wolfenstein parametrization is unitary to all orders in λ, and one can see that it reveals a very distinctive hierarchy of the elements. The major diagonal terms of order one characterize the most probable transitions between quarks of the same generation. The transitions between the first and second generation are of order λ and for the second and third generation of λ^2. This is also illustrated diagrammatically in Fig. 4.2 (a). The *CP*-violating phase at this order of λ is only contained in transitions between the first and third generation with elements V_{td} and V_{ub}, proportional to λ^3.
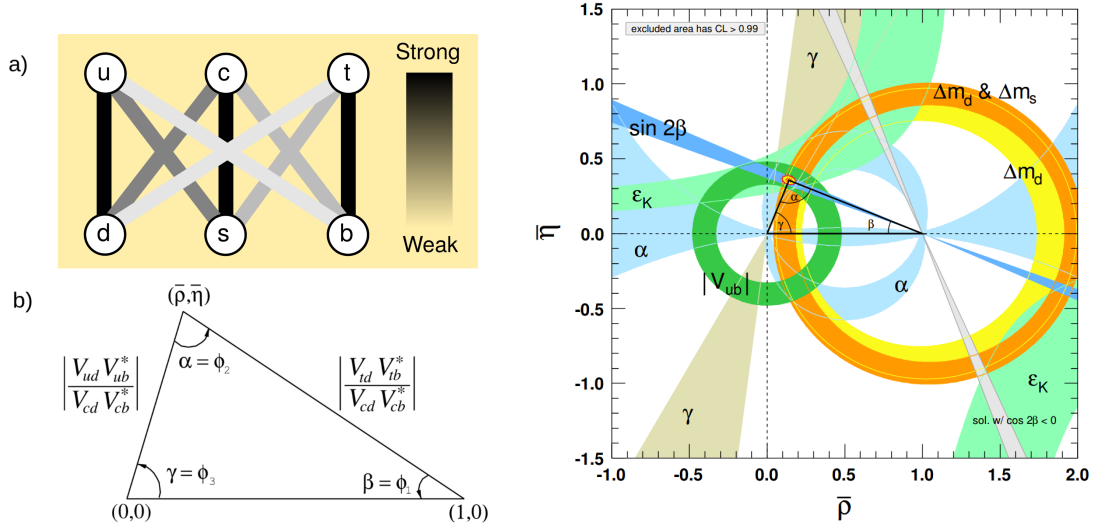
Figure 4.2: Left: Diagrammatic illustration of quark transition probabilities (a) and the unitary triangle (b) with angles and lengths of sites denoted using the standard normalized representation in the complex plane [35]. Right: The current experimental constraints on sides and angles of the unitary triangle [34].

The unitarity of the CKM matrix leads to several relations among its elements. Of particular relevance are those constraining the scalar products of any two rows and columns as [29]:

$$\sum_i V_{ji} V_{ki}^\star = 0 = \sum_i V_{ij} V_{ik}^\star, \qquad j \neq k. \tag{4.9}$$

Six such relations emerge in the form of a sum of three complex numbers equal to zero. Each can be visualized as a triangle in the complex plane, all of an equal area related to an invariant expression[9] for the total amount of *CP* violation from the CKM matrix. Only one of these triangles, the one related to *B* meson decays, has sides of comparable size ($\sim \lambda^3$) and thus promises generally large observable *CP*-violation effects. This triangle is usually called The Unitary Triangle (UT) and is defined by the second relation in Eq. 4.9 with $j = d, k = b$ (product of the first and third CKM column) as

$$V_{ud} V_{ub}^\star + V_{cd} V_{cb}^\star + V_{td} V_{tb}^\star = 0. \tag{4.10}$$

---

[9]Before the normalization below, area(any triangle) = $J/2$, where $J = |\Im(V_{ij} V_{kl} V_{il}^\star V_{kj}^\star)| \sim 3 \times 10^{-5}$ [34] is the Jarlskog invariant [29].

It is common to normalize this expression by division by an almost[10] real number $V_{cd}V_{cb}^{\star}$, such that one side has unit length, and the triangle can be visualized as in Fig. 4.2. The angles of the unitary triangle are defined as

$$\beta = \phi_1 = \arg\left(-\frac{V_{cd}V_{cb}^{\star}}{V_{td}V_{tb}^{\star}}\right)$$
$$\alpha = \phi_2 = \arg\left(-\frac{V_{td}V_{tb}^{\star}}{V_{ud}V_{ub}^{\star}}\right) \qquad (4.11)$$
$$\gamma = \phi_3 = \arg\left(-\frac{V_{ud}V_{ub}^{\star}}{V_{cd}V_{cb}^{\star}}\right),$$

where the notation β, α, γ is common at BaBar, while Belle is using ϕ_1, ϕ_2, ϕ_3.

It is evident that the CKM matrix elements appear in plenty of phenomena related to quark decays. The four parameters of the CKM matrix lead to a large number of predictions, which can be, in principle, probed by experiments. The theoreticians' challenge is finding measurements with small theoretical uncertainties, mainly stemming from non-perturbative QCD effects and hadronization. The idea is to over-constrain the UT from many measurements. Experiments can independently probe the sides and the angles, checking for their consistency. Any inconsistency could mean deviations from unitarity and constitute strong indications for NP effects. The contemporary experimental status of the UT is shown in Fig. 4.2, where at the current precision, all measurements indeed agree, and the triangle closes at the apex. However, not all measurements are shown in this figure, and small discrepancies already exist, which means that some deviation from SM predictions could be confirmed with more precise measurements. As it is not a priory known where to look for new physics, it is desirable to study the phenomena in many possible ways. The decays of B mesons produced at the B-Factory are an excellent playground for such studies. In the following sections, we will explicitly look at the measurement of the $\phi_1 = \beta$ angle, which is the flagship of the B-Factory physics program. This particular class of measurements drives many B-Factory design decisions.

4.4 Time-Dependent *CP* Violation in *B* Decays

We have already briefly mentioned that some neutral particles can mix between matter and anti-matter in the case of neutral kaons. In this section, this effect will be discussed deeper in the context of heavy mesons, which contain a beauty quark or anti-quark and a first-generation quark ($\overline{d}$ or d for neutral B). This mixing will be exploited to access *CP*-violating effects, which can lead to time-dependent decay rates different for B^0 and $\overline{B}^0$, demonstrating the (tiny) difference between matter and anti-matter in the SM.

4.4.1 Mixing of Neutral *B* Mesons

The theory of neutral particle mixing starts with quite a general formulation using non-relativistic quantum mechanical treatment of the system represented by

---

[10]Up to $\mathcal{O}(\lambda^4)$ according to the Wolfenstein parameterization. The slight complexity of $V_{cd}V_{cb}^{\star}$ means that the triangle is also rotated a bit.

two-dimensional Hilbert vector space. For the neutral B mesons, the two possible states of definite flavor are $|B^0\rangle$ and $|\overline{B}^0\rangle$. These are, however, not the mass eigenstates, for which the time evolution is rather simple. Under assumptions of *CPT* symmetry, the time evolution and decay of the neutral B meson wave function $|\psi\rangle = \psi_1(t)|B^0\rangle + \psi_2(t)|\overline{B}^0\rangle$ is governed by the time-dependent Schroedinger equation[11] $i\hbar\frac{\partial}{\partial t}\begin{pmatrix}\psi_1(t)\\\psi_2(t)\end{pmatrix} = \mathcal{H}_{\text{eff}}\begin{pmatrix}\psi_1(t)\\\psi_2(t)\end{pmatrix}$ with an effective Hamiltonian [29]:

$$\mathcal{H}_{\text{eff}} = \boldsymbol{M} - \frac{i}{2}\boldsymbol{\Gamma} = \begin{pmatrix}M & M_{12}\\M_{12}^\star & M\end{pmatrix} - \frac{i}{2}\begin{pmatrix}\Gamma & \Gamma_{12}\\\Gamma_{12}^\star & \Gamma\end{pmatrix} \tag{4.12}$$

which is not hermitian (because of the i before $\boldsymbol{\Gamma}$)—the probability is not conserved as the B meson eventually decays. The mass eigenstates of a definite lifetime can be obtained by diagonalization of $\mathcal{H}_{\text{eff}}$. The eigenvalues are:

$$\mu_{1,2} = M - \frac{i}{2}\Gamma \pm \sqrt{\left(M_{12} - \frac{i}{2}\Gamma_{12}\right)\left(M_{12}^\star - \frac{i}{2}\Gamma_{12}^\star\right)}. \tag{4.13}$$

The corresponding eigenvectors (in general non-orthogonal as $\mathcal{H}_{\text{eff}}$ is not hermitian) can be written as

$$\begin{aligned}|B_1\rangle &= p|B^0\rangle + q|\overline{B}^0\rangle\\|B_2\rangle &= p|B^0\rangle - q|\overline{B}^0\rangle,\end{aligned} \tag{4.14}$$

where $|p|^2 + |q|^2 = 1$ and

$$\frac{q}{p} = \sqrt{\frac{M_{12}^\star - \frac{i}{2}\Gamma_{12}^\star}{M_{12} - \frac{i}{2}\Gamma_{12}}}. \tag{4.15}$$

For the eigenvectors, the standard exponential time evolution results from the solution of the time-dependent Schroedinger equation, where the decay is achieved by the complex part of the eigenvalues:

$$|B_{1,2}(t)\rangle = e^{-i\mu_{1,2}t}|B_{1,2}\rangle \tag{4.16}$$

with $|B_{1,2}\rangle$ being the initial state. The masses of the eigenstates can be obtained as $m_{1,2} = \Re(\mu_{1,2})$ and the lifetimes as $\Gamma_{1,2} = -2\Im(\mu_{1,2})$. We can now invert the relation 4.14 to get

$$\begin{aligned}|B^0\rangle &= \frac{1}{2p}(|B_1\rangle + |B_2\rangle)\\|\overline{B}^0\rangle &= \frac{1}{2q}(|B_1\rangle - |B_2\rangle)\end{aligned} \tag{4.17}$$

at time $t = 0$. The time evolution for the flavor eigenstates is finally obtained by letting the mass eigenstates to evolve in time $|B_{1,2}\rangle \to |B_{1,2}(t)\rangle$, which gives

$$\begin{aligned}|B^0(t)\rangle &= g_+(t)|B^0\rangle + \frac{q}{p}g_-(t)|\overline{B}^0\rangle\\|\overline{B}^0(t)\rangle &= g_+(t)|\overline{B}^0\rangle + \frac{p}{q}g_-(t)|B^0\rangle,\end{aligned} \tag{4.18}$$

where

$$g_\pm(t) = \frac{1}{2}\left(e^{-i(m_1 - \frac{i}{2}\Gamma_1)t} \pm e^{-i(m_2 - \frac{i}{2}\Gamma_2)t}\right). \tag{4.19}$$

---

[11]$\hbar = 1$ in natural units mostly used in this thesis.

The time evolution can be further simplified for the case of B mesons by neglecting the lifetime difference[12] and setting $\Gamma_1 = \Gamma_2 = \Gamma = \hbar/\tau_{B^0}$, where $\tau_{B^0} \sim 1.519$ ps [36]. Using average mass $\overline{m} = (m_1 + m_2)/2$ and the mass difference[13] $\Delta m = m_2 - m_1 \sim 0.507$ ps^{-1}, $g_\pm(t)$ becomes

$$
\begin{aligned}
g_+(t) &= e^{-\frac{\Gamma}{2}t} \cos\left(\frac{\Delta m}{2}t\right) \\
g_-(t) &= ie^{-\frac{\Gamma}{2}t} \sin\left(\frac{\Delta m}{2}t\right)
\end{aligned}
\tag{4.20}
$$

where a factor of $e^{-i\overline{m}t}$, irrelevant for further discussion, is omitted from the expressions on the r.h.s. The final form of the oscillations finally emerges by evaluating $|g_\pm(t)|^2$ to obtain the probability of observing a decay of the same or opposite flavor at time $t > 0$ by starting from a pure flavor state at $t = 0$. For an initial state of B^0

$$
\begin{aligned}
|\langle B^0|B^0(t)\rangle|^2 &= \frac{1}{2}e^{-\Gamma t}(1 + \cos(\Delta m t)) \\
|\langle \overline{B}^0|B^0(t)\rangle|^2 &= \left|\frac{q}{p}\right|^2 \frac{1}{2}e^{-\Gamma t}(1 - \cos(\Delta m t))
\end{aligned}
\tag{4.21}
$$

while for an initial $\overline{B}^0$, we get

$$
\begin{aligned}
|\langle \overline{B}^0|\overline{B}^0(t)\rangle|^2 &= \frac{1}{2}e^{-\Gamma t}(1 + \cos(\Delta m t)) \\
|\langle B^0|\overline{B}^0(t)\rangle|^2 &= \left|\frac{p}{q}\right|^2 \frac{1}{2}e^{-\Gamma t}(1 - \cos(\Delta m t)).
\end{aligned}
\tag{4.22}
$$

The oscillation probability is shown in Fig. 4.3. Thus the oscillations only demonstrate the mass difference. CP violation in mixing, which manifests as $P(B^0 \to \overline{B}^0) \neq P(\overline{B}^0 \to B^0)$ (see below), only occurs for $|q/p| \neq 1$. But for B mesons, this number is very close to unity [35], despite having quite a large phase. To reveal this phase, we need to multiply the states from the left by a state to which both flavors can decay to employ the interference effect.

4.4.2 Types of CP Violation

In the time evolution of neutral B mesons, three basic types of CP violation can be defined when B^0 and $\overline{B}^0$ decay to the same final state $|f\rangle$. Let us assume f is a CP eigenstate[14], that is $f = \overline{f}$ and $CP|f\rangle = \pm|f\rangle$. Denoting the amplitudes of pure flavor state transition to f as:

$$
\begin{aligned}
\langle f|B^0\rangle &= A \\
\langle f|\overline{B}^0\rangle &= \overline{A},
\end{aligned}
\tag{4.23}
$$

---

[12]This is the main difference from the oscillations of neutral kaons where the lifetimes of the two states are radically different.

[13]For B mesons, B_1 is identified with the lower mass state B_L and B_2 with the higher mass state B_H. The mass difference is also often denoted by Δm_d to emphasize that we mean B_d mesons—with b and d quarks.

[14]For $f \neq \overline{f}$, there are four possibly distinct amplitudes $A_f, A_{\overline{f}}, \overline{A}_f, \overline{A}_{\overline{f}}$ which will make the expression more complicated. It is however not relevant for further discussion
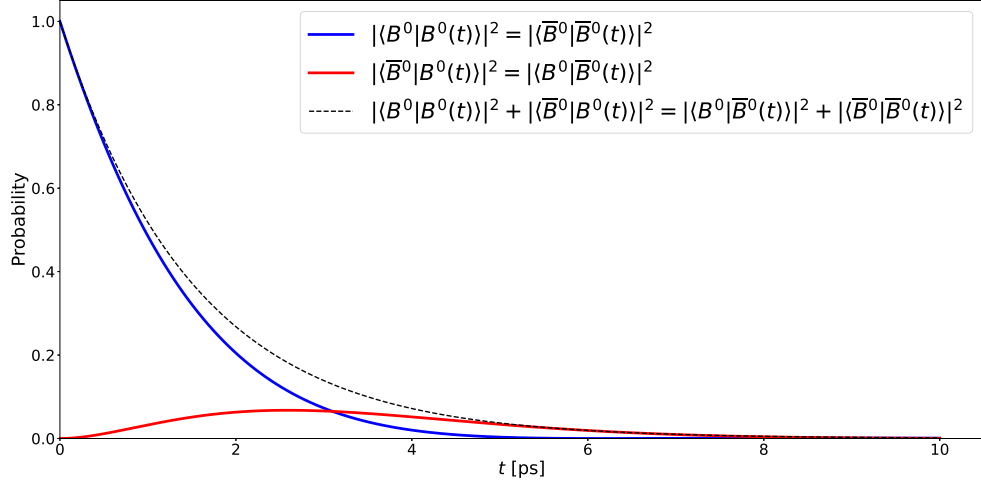
Figure 4.3: Probability to find the same flavor (blue) or opposite, oscillated flavor (red) neutral B meson at time t when starting from a pure flavor state (B^0 or $\overline{B}^0$) at $t = 0$. The black dashed curve shows the total exponential decay rate.

the time-dependent amplitudes for the decay of B meson and its antiparticle to f are obtained from Eq. 4.18 as

$$\langle f|B^0(t)\rangle = A\,(g_+(t) + \lambda g_-(t))$$
$$\langle f|\overline{B}^0(t)\rangle = \overline{A}\,(g_+(t) + \frac{1}{\lambda}g_-(t)), \tag{4.24}$$

where we have used the standard definition for the complex parameter λ[15], usually denoted as λ_{CP}

$$\lambda = \lambda_{CP} = \frac{q}{p}\frac{\overline{A}}{A}. \tag{4.25}$$

Three different types of CP violation can now be discussed based on the magnitude and phase of λ, respective q/p or $\overline{A}/A$ (where multiple types can be present simultaneously):

- **Direct CP violation** in decay occurs in case of $|\overline{A}/A| \neq 1$. This is the only type of CP violation, which can be also observed in charged B meson decays, which cannot mix. It can result from an interference of multiple diagrams with different weak and strong phases.

- **Indirect CP violation** in mixing follows from $|q/p| \neq 1$. This type of CP violation was initially observed in the neutral kaon mixing. It results in different probabilities for $K^0 \to \overline{K}^0$ and $\overline{K}^0 \to K^0$. In the kaon system, this effect is observable thanks to the large lifetime difference of the mass eigenstates K_S^0 and K_L^0. For B mesons, this type of CPV is irrelevant.

- ***CP* violation in interference of mixing and decay** happens for $\Im(\lambda) \neq 0$. This is the time-dependent CP violation from the title of this thesis. It is caused by interference of direct decay amplitude $B^0 \to f$ and amplitude where the B meson first changes its flavor before it decays ($B^0 \to \overline{B}^0 \to f$).

---

[15]This is *not* the Wolfenstein parameter $\lambda = \sin\theta_c$. Which one is meant should be obvious from the context.

The principle of observability of $\Im(\lambda)$ in the interference of mixing and decay is worth noting. Let us assume for a moment that we can write $\lambda = e^{i\phi}$ and that $|A| = |\overline{A}|$, which will be justified later for the case of B mesons. At the time $t = 0$, $g_-(0) = 0$, and the squares of the two amplitudes would be equal. At time $t = \pi/\Delta m$, $g_+(\pi/\Delta m) = 0$ and $g_-(\pi/\Delta m) = ie^{-\Gamma t/2}$. Thus the amplitudes now differ by a factor of $e^{2i\phi}$. By squaring the two amplitudes, there is no difference again and thus no CP violation. We can see this factor is a complex number whose argument is related to the weak phase, which has an opposite sign for B^0 and $\overline{B}^0$. However, we need two phases, of which one switches its sign and one does not. The second invariant phase factor, whose role is played by the strong phase for the direct CP violation, is provided by the wave function phase from the unitary time evolution. This is also illustrated in Fig. 4.4, where the squares of the two interfering diagrams are illustrated for B^0 and $\overline{B}^0$. The maximal CP violation occurs[16] at time $t = \pi/\Delta m/2 \sim 3$ ps. At this moment, under our simplified assumptions, one can write

$$\left\langle f \middle| B^0\left(t = \frac{\pi}{2\Delta m}\right)\right\rangle = e^{-\frac{\pi\Gamma}{4\Delta m}}\frac{\sqrt{2}}{2}(1 + ie^{+i\phi})$$
$$\left\langle f \middle| \overline{B}^0\left(t = \frac{\pi}{2\Delta m}\right)\right\rangle = e^{-\frac{\pi\Gamma}{4\Delta m}}\frac{\sqrt{2}}{2}(1 + ie^{-i\phi}).$$

(4.26)

For these two complex numbers, their absolute values are now, however, not identical, and the complex phase in λ fully reveals in their difference, which is proportional to $\sin\phi$.
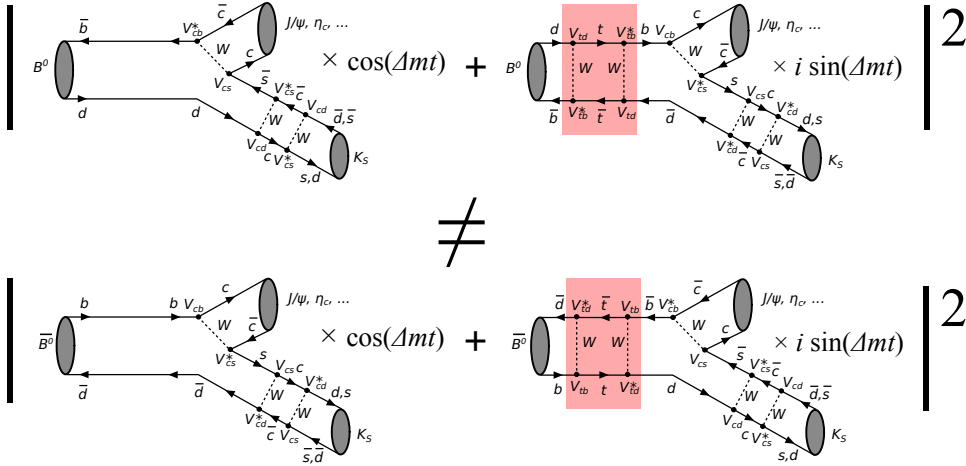
Figure 4.4: Illustration of Feynman diagrams for a direct B meson decay to $(c\bar{c})K_S$ interfering with neutral meson oscillation (red box) followed by the decay for an initial pure B^0 state (top) and CP conjugated $\overline{B}^0$ state (bottom). Due to the additional phases from unitary time evolution, the squares of the two absolute values are generally not equal, and the difference depends on t.

The time-dependent decay rate of an initially pure B^0 state is obtained by

---

[16]Further periods are much less interesting as the B mesons quickly decay

squaring the corresponding time-dependent amplitudes in Eq. 4.24 as

$$\Gamma(B^0 \to f) \equiv \Gamma(B^0 \to f; t) = |\langle f|B^0(t)\rangle|^2 = |A|^2[g_+(t) + \lambda g_-(t)][g_+^\star(t) + \lambda^\star g_-^\star(t)].$$
(4.27)

After some calculations [17] , one gets

$$\Gamma(B^0 \to f) =$$
$$|A|^2 \frac{e^{-\Gamma t}}{2}(1 + |\lambda|^2)\Big[1 - \frac{|\lambda|^2 - 1}{1 + |\lambda|^2}\cos(\Delta mt) - \frac{2\Im(\lambda)}{1 + |\lambda|^2}\sin(\Delta mt)\Big].$$
(4.28)

The equivalent expression for $|\langle f|\overline{B}^0(t)\rangle|^2$ can be obtained by the replacement[18] $\lambda \to \frac{1}{\lambda}$:

$$\Gamma(\overline{B}^0 \to f) =$$
$$|\overline{A}|^2 \frac{1}{|\lambda|^2}\frac{e^{-\Gamma t}}{2}(1 + |\lambda|^2)\Big[1 + \frac{|\lambda|^2 - 1}{1 + |\lambda|^2}\cos(\Delta mt) + \frac{2\Im(\lambda)}{1 + |\lambda|^2}\sin(\Delta mt)\Big].$$
(4.29)

4.4.3   Time-Dependent *CP* Asymmetry for *B* Mesons

The expressions for decay rates can be further simplified, as the q/p factor, which comes from $B^0 - \overline{B}^0$ oscillations and the box diagram, further discussed in Sec. 4.5, is close to unity in magnitude. Thus for B mesons, we can substitute

$$|\lambda|^2 = \frac{|\overline{A}|^2}{|A|^2}$$
(4.30)

in Eq. 4.28 and 4.29, resulting in a striking similarity between the two equalities. By taking their difference and dividing it by the average as

$$a(t) = \frac{\Gamma(\overline{B}^0 \to f) - \Gamma(B^0 \to f)}{\Gamma(\overline{B}^0 \to f) + \Gamma(B^0 \to f)},$$
(4.31)

the time-dependent asymmetry $a(t)$ reveals in an experimentally and theoretically more suitable form, where many unknowns and uncertainties cancel. One gets

$$a(t) = \frac{|\lambda|^2 - 1}{|\lambda|^2 + 1}\cos(\Delta mt) + \frac{2\Im(\lambda)}{|\lambda|^2 + 1}\sin(\Delta mt).$$
(4.32)

The *CP* violation is thus fully described by two parameters defined as

$$\mathcal{A}_{CP} = \frac{|\lambda|^2 - 1}{|\lambda|^2 + 1}$$
$$\mathcal{S}_{CP} = \frac{2\Im(\lambda)}{|\lambda|^2 + 1},$$
(4.33)

where $\mathcal{A}_{CP}$ measures the direct *CP* asymmetry in the decay and $\mathcal{S}_{CP}$ measures the time-dependent asymmetry in the interference of mixing and decay.

---

[17]Note that $\lambda^\star - \lambda = -2i\Im(\lambda)$
[18]Using $\Im(\frac{1}{\lambda}) = \frac{-\Im(\lambda)}{|\lambda|^2}$

4.5 Time-Dependent *CP* Violation in $b \to c\bar{c}s$ Decays

In this section, the *CP*-violation parameters $\mathcal{A}_{CP}$ and $\mathcal{S}_{CP}$ will be related to fundamental SM parameters, namely one of the unitary angles. The case of $b \to c\bar{c}s$ transitions, originally recognized as golden modes to access the angle ϕ_1, will be used. The most famous process $B^0 \to J/\psi K_S^0$ is an example of decay into a pure *CP* eigenstate common to B^0 and $\overline{B}^0$. This discussion is, however, general for all $c\bar{c}$ mesons, in particular η_c.

Two amplitudes represented by lowest-order Feynman diagrams, leading to the necessary interference, are illustrated in Fig. 4.5. The process $B^0 \to c\bar{c}K_S^0$ on the left interferes with $B^0 \to \overline{B}^0 \to c\bar{c}K_S^0$ (on the right), with additional $B^0 \to \overline{B}^0$ oscillation, which proceeds via a box diagram, whose example is shown in 4.5 (a). The additional box diagram (c) represents the $\overline{K}^0 \leftrightarrow K^0$ oscillation, but only schematically—to illustrate how a superposition of flavor eigenstates emerges to reach a mass (and to good accuracy also a *CP*) eigenstate K_S^0. Needless to say, the $\overline{K}^0 \leftrightarrow K^0$ oscillation is both a complication and a blessing [29]. Without them, we will not have a final *CP* eigenstate. Fortunately, corrections arising from *CP*-violation effects in kaons are small, and in the typical experimental setup, where K_S^0 are registered through $\pi^+\pi^-$ decays by a vertex detector up to a distance of several flight lengths of K_S^0, the occasional $K_L^0 \to \pi^+\pi^-$ decays can be ignored [37].
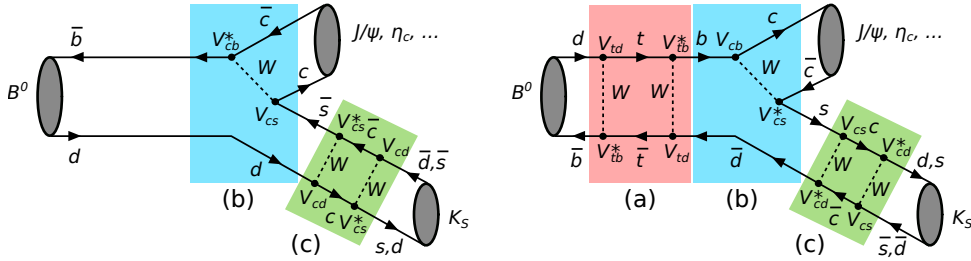
Figure 4.5: Examples of interfering diagrams for $B^0 \to (c\bar{c})K_S^0$ decays. The tree-level decay (b) might be preceded by a $B^0 \to \overline{B}^0$ oscillation (a). The green box diagram (c) represents neutral kaon mixing, necessary to reach the final *CP* eigenstate.

The parameter encoding the *CP*-violation effects is defined as (see Eq. 4.25)

$$\lambda_{c\bar{c}K_S^0} = \frac{q}{p}\frac{\overline{A}_{c\bar{c}K_S^0}}{A_{c\bar{c}K_S^0}} = \frac{q}{p}\frac{\langle c\bar{c}K_S^0|\overline{B}^0\rangle}{\langle c\bar{c}K_S^0|B^0\rangle}, \tag{4.34}$$

where q/p comes from $B^0 - \overline{B}^0$ oscillations. The term $A_{c\bar{c}K_S^0}$ has two contributions. The first is shown (on the left) in Fig. 4.5 (b) and is proportional to the product of CKM elements $V_{cb}^\star V_{cs}$. The emerging strange meson is, however, not yet common to B^0 and $\overline{B}^0$. In the diagram, we have an $\overline{s}$-quark going from the weak vertex V_{cs} (b) and thus K^0 meson, which results from the B^0 decay. That is $B^0 \to (c\bar{c})K^0$ and $\overline{B}^0 \to (c\bar{c})\overline{K}^0$. To get a K_S^0 meson, we have to multiply this by the transition

amplitude $\langle K_S^0 | K^0 \rangle$. A similar treatment for a diagram starting with $\overline{B}^0$, also shown in Fig. 4.4 (bottom), leads to

$$\lambda_{c\bar{c}K_S^0} = \eta_{CP} \frac{q}{p} \frac{V_{cs}^\star V_{cb} \langle K_S^0 | \overline{K}^0 \rangle}{V_{cb}^\star V_{cs} \langle K_S^0 | K^0 \rangle}, \tag{4.35}$$

where the sign factor of η_{CP} encodes the CP eigenvalue of the final state, i.e.

$$\eta_{CP}(J/\psi K_S^0) = \eta_{CP}(\eta_c K_S^0) = -1.$$

Note that $CP(\eta_c) = -1$, while $CP(J/\psi) = +1$ but for J/ψ the B (scalar) decays to a vector (J/ψ) and scalar (K_S^0)[19] and thus, these must be in a state with orbital momentum $L = 1$ (giving one minus factor to the parity of the final state).

The treatment of the oscillation part is quite similar for B and K mesons. Like for B (Eq. 4.14), we can rewrite part of Eq. 4.6 as

$$|K_S^0\rangle = p_K |K^0\rangle + q_K |\overline{K}^0\rangle. \tag{4.36}$$

Instead of 'bra' vector, one needs the corresponding 'ket' vector, which is equivalent to $\langle K_S^0 |$ under the assumption of CP conservation in the neutral kaon oscillations[20]. Then one can write $\langle K_S^0 | K^0 \rangle = p_K^\star$ and $\langle K_S^0 | \overline{K}^0 \rangle = q_K^\star$ [38]. The coefficients p_K and q_K have the same meaning as p and q but are evaluated in the neutral kaon system. Assuming $|M_{12}| \gg |\Gamma_{12}|$, one can write Eq. 4.15 as[8]

$$\frac{q}{p} = \sqrt{\frac{M_{12}^\star}{M_{12}}}. \tag{4.37}$$

This means that q/p is given by the elements of the mass matrix in Eq. 4.12, which connect different flavor eigenstates. These correspond to the amplitudes of $B^0 \to \overline{B}^0$ obtained from the box diagram. For B's, the t-quark dominates in the oscillations; other contributions can be neglected. Collecting the CKM elements for $B^0 \to \overline{B}^0$ then gives $M_{12} \propto (V_{td} V_{tb}^\star)^2$. Therefore for the B mesons

$$\frac{q}{p} = \frac{V_{td} V_{tb}^\star}{V_{tb} V_{td}^\star}. \tag{4.38}$$

In fact, this term brings already the full phase difference (other smaller phases will cancel) and in the SM, $\frac{q}{p} \simeq e^{-2i\phi_1}$ [29]. For the neutral kaons, the charm contribution is the most relevant for oscillations (as V_{ts} is of order λ^2 while V_{tb} is close to unity for B mesons). This can be written in analogy to the beauty sector to a good approximation [25] as

$$\frac{q_K}{p_K} = \frac{V_{cd} V_{cs}^\star}{V_{cs} V_{cd}^\star}. \tag{4.39}$$

Combining the results together, we get

$$\lambda_{c\bar{c}K_S^0} = \eta_{CP} \frac{V_{td} V_{tb}^\star}{V_{tb} V_{td}^\star} \frac{V_{cs}^\star V_{cb}}{V_{cb}^\star V_{cs}} \frac{V_{cd}^\star V_{cs}}{V_{cs}^\star V_{cd}} = \eta_{CP} \frac{V_{tb}^\star V_{td}}{V_{cb}^\star V_{cd}} \frac{V_{cb} V_{cd}^\star}{V_{tb} V_{td}^\star}. \tag{4.40}$$

---

[19]To a very good accuracy $CP(K_S^0) = +1$, which neglects CP violation in the neutral kaon system.

[20]Otherwise one should use the reciprocal basis [38], which will lead to the same final result for $\lambda_{c\bar{c}K_S^0}$ in case $|q_K/p_K| \simeq 1$, which is true to order of 10^{-3}.

This complex number is in the form of $z/z^\star$, so $|\lambda_{c\bar{c}K_S^0}|^2 = 1$, and thus the direct CP asymmetry parameter (defined in Eq. 4.33) $\mathcal{A}_{c\bar{c}K_S^0} = 0$ in the SM to a good approximation. This allows us to write $\lambda_{c\bar{c}K_S^0}$ as a complex unit and obtain its imaginary part as $\Im(\lambda_{c\bar{c}K_S^0}) = \sin(\arg(\lambda_{c\bar{c}K_S^0}))$. We get[21]

$$\Im(\lambda_{c\bar{c}K_S^0}) = \mathcal{S}_{c\bar{c}K_S^0} = -\eta_{CP}\sin(2\phi_1), \tag{4.41}$$

where ϕ_1 is angle of the unitary triangle defined in Eq. 4.11 and illustrated in Fig. 4.2.

In summary, the SM predicts for the CP-violation parameters measured in $B^0 \to \eta_c K_S^0$

$$\begin{aligned} \mathcal{A}_{CP} &= 0 \\ \mathcal{S}_{CP} &= \sin 2\phi_1 \end{aligned} \tag{4.42}$$

within sub-percent theoretical uncertainty. Moreover, these parameters should be (up to the sign different for CP-odd and CP-even final states) identical for all $b \to c\bar{c}s$ transitions, irrespective of the involved charmonium resonance. Note that there is a two-fold ambiguity $\phi_1 \to \pi/2 - \phi_1$, which can be resolved by other measurements [14].

4.5.1 Trees and Penguins

In the previous discussion, it was assumed that the decay corresponding to part (b) in Fig. 4.5 could be described by a single amplitude which represents the tree level $b \to c\bar{c}s$ decay. At the one-loop level, so-called penguin diagrams have to be considered. An example of a gluonic penguin diagram for our process is shown next to the tree-level one in Fig. 4.6, specifically for the case of η_c charmonium[22].
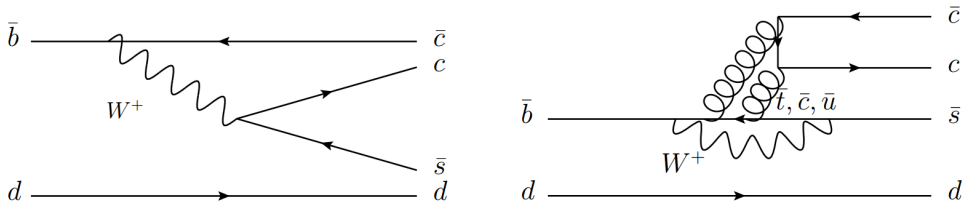
Figure 4.6: Feynmann diagrams for $\bar{b} \to \bar{c}c\bar{s}$ decay [40]. Left: Dominant color-suppressed tree diagram. Right: Strong penguin diagram with the exchange of gluons. For an electroweak penguin diagram, the gluons are replaced by (single) γ or Z, but these contributions can be neglected.

Similar diagrams where the gluons are replaced by a (single) Z^0 boson or γ (electroweak penguins) can be neglected [41]. The three possible virtual quarks $\bar{t}, \bar{c}, \bar{u}$ in the loop result in three contributions to the penguin amplitude. The total amplitude can be written as [25]

$$A_{c\bar{c}K^0} = V_{cb}V_{cs}^\star(T + P_c - P_t) + V_{ub}V_{us}^\star(P_u - P_t) \tag{4.43}$$

---

[21] Using $\arg(x/x^\star) = 2\arg(x)$, $\arg(x^\star) = -\arg(x)$ and $\arg(-x) = \arg(x) + \pi$ in the definition of ϕ_1/β.

[22] For J/ψ, three gluon exchange is necessary due to charge-conjugation invariance [39]

where T is the amplitude of the tree diagram and P_t, P_c, and P_u represent the amplitudes of the penguin diagrams, where the CKM elements are factored out. From the Wolfenstein parameterization in Eq. 4.8, one can see there is a large phase in V_{ub}, coming from the u quark loop and P_u, which is different from the phase from the box diagram. At the same time, the combination of CKM elements $V_{ub}V_{us}^\star \propto \lambda^4$, while $V_{cb}V_{cs}^\star \propto \lambda^2$, so P_u is CKM-suppressed. The P_u contribution introduces a theoretical uncertainty in the correspondence of $\mathcal{S}_{c\bar{c}K_S^0}$ and $\sin 2\phi_1$. The size of this effect depends on the relative size of P_u. It is estimated that[29]

$$\left| \frac{\lambda^2 P_u}{T} \right| < 1 \ \% \tag{4.44}$$

which results in sub-percent theoretical uncertainty of $\sin 2\phi_1$ determination, called penguin pollution.

These considerations are usually made with the assumption of no additional interference[23]. However, by selecting particular final states, the number of interfering diagrams can be larger when the final state resulting from $c\bar{c}$ together with the K_S^0 from B meson can be, for example, reached by a (non-flavor-specific) direct B decay into final state particles. This will be the case of our analyzed $\eta_c \to K_S^0 K^+ \pi^-$ decay, where an important background source stems from $b \to s$ transitions, which must proceed via a loop, where the u quark contribution has a different weak phase than the tree-level decay. Or the final state may be reached via a $b \to u$ tree directly. Evaluation of such diagrams is very difficult due to non-perturbative QCD calculations. This can lead to more significant uncertainties in $\sin 2\phi_1$, which can be considered theoretical. The measurement of the CP-violation parameters $\mathcal{A}_{CP}$ and $\mathcal{S}_{CP}$ is still fully valid, albeit these values might then be understood as effective, where background interference effects are not decoupled.

4.6 Measurement of $\sin 2\phi_1$ at B-Factories

The actual observation of the time-dependent CP violation is facing a number of experimental challenges, which have been successfully resolved. First, a clean and copious sample of neutral B mesons is needed. This is achieved by colliding electrons and positrons at a specific center-of-mass energy. The cross-section for production of hadrons in e^+e^- collisions around 10 GeV is shown in Fig. 4.7 (left), where several Υ resonances are denoted. These resonances are bound states of b and $\bar{b}$ quark, also called bottomonium. The spectrum of bottomonium is shown in Fig. 4.7 (right). The mass of $\Upsilon(4S)$ resonance is just above the threshold for $B\overline{B}$ production, and this resonance almost exclusively (in $> 96 \ \%$ of cases [36]) decays into B^+B^- or $B^0\overline{B}^0$. The kinetic energy of the B mesons is minimal in the CM frame compared to their mass as the energy release of $\Upsilon(4S)$ decay is only about 20 MeV.

---

[23]Tag-side interference is often considered [8] and will also be included as a systematic error in our analysis.
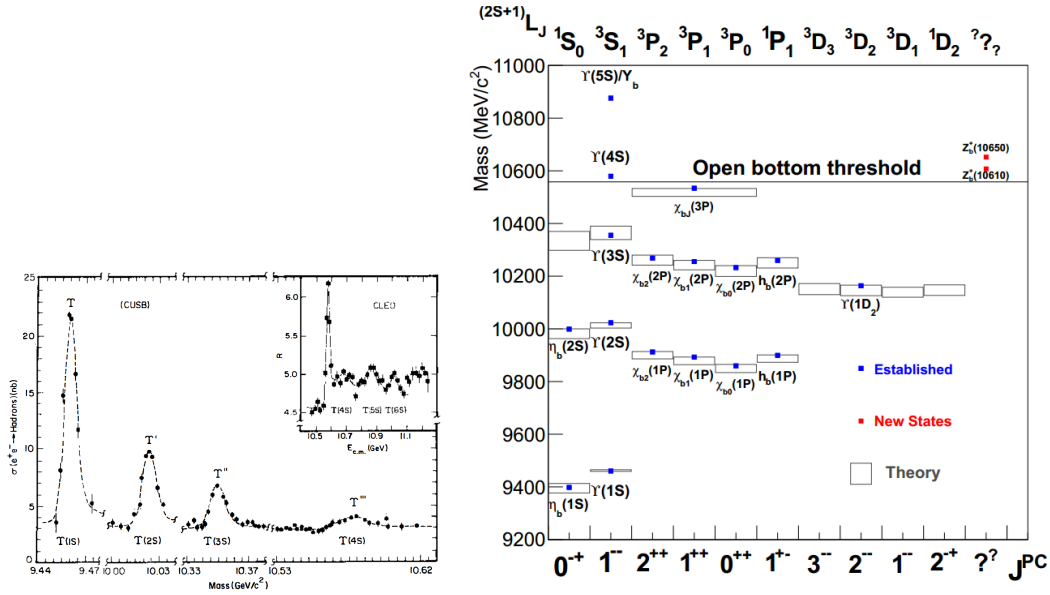
Figure 4.7: Cross section of electron-positron annihilation to hadrons with peaks corresponding to bottomonium resonances (left) [42], and the spectrum of bottomonium (right), with states sorted by mass and their quantum numbers [8].

4.6.1 Coherent B Mesons and the EPR Paradox

The $\Upsilon(4S)$ decay to B mesons proceeds via strong interaction, and thus it conserves both C and P separately. From the quantum numbers of spin-one particle $\Upsilon(4S)$, $J^{PC}(\Upsilon(4S)) = 1^{--}$, it follows that in the decay $\Upsilon(4S) \to B^0\overline{B}^0$, the two B mesons with spin zero ($J^P(B) = 0^-$) must be in a p-wave state with orbital angular momentum quantum number $L = 1$. Upon exchange of the two B mesons, this yields a factor of $(-1)^{L=1} = -1$, i.e., the angular part of the wave function is antisymmetric. Einstein-Bose statistics dictates that the total wave function describing a state of identical bosons must be symmetric. Thus the two (oscillating) bosons cannot be identical at any time while they transform into each other. This is an example of an entangled state which can be compactly written as

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\Big(|B^0\rangle|\overline{B}^0\rangle - |\overline{B}^0\rangle|B^0\rangle\Big), \tag{4.45}$$

which is correctly C and P[24] odd (and thus CP even) like the initial state of $\Upsilon(4S)$.

Next, one should measure the decay rates as a function of time since $\Upsilon(4S)$ decay in the rest frames of the respective B mesons, separately for each flavor. This brings two complications due to a short B lifetime (about 1.5 ps) and the fact that we are interested primarily in decays into a CP eigenstate, common to B^0 and $\overline{B}^0$. Thus by the very definition, we cannot tell the flavor of the B meson, which decays into our signal. At this point, the advantage of the initial entangled state will come to the rescue. By inserting the time evolution for the ket vectors from Eq. 4.18 and using Eq. 4.20, the wave function of the system of the two

---

[24]Note that the two bosons fly to opposite directions, so spatial inversion generates exactly the same (or mathematically equivalent) linear combination, but with a minus sign.

mesons, first[25] at time t_1 and second at time t_2, can be written as

$$|\Psi(t_1, t_2)\rangle = \frac{1}{\sqrt{2}}e^{-\frac{\Gamma}{2}(t_1+t_2)}\left[\cos\left(\frac{\Delta m \Delta t}{2}\right)\left(|B^0\rangle|\overline{B}^0\rangle - |\overline{B}^0\rangle|B^0\rangle\right)\right.$$
$$\left. + i\sin\left(\frac{\Delta m \Delta t}{2}\right)\left(\frac{p}{q}|B^0\rangle|B^0\rangle - \frac{q}{p}|\overline{B}^0\rangle|\overline{B}^0\rangle\right)\right], \tag{4.46}$$

where $\Delta t = t_2 - t_1$. Until one of the mesons decays, $\Delta t = 0$ and exactly one B^0 and one $\overline{B}^0$ is present in the system. Despite both oscillating, this oscillation is exactly synchronous. Let's assume, for example, that the first B decays at time $t_1 = t_{\text{tag}}$ into a final state f_{tag}, which reveals[26] its flavor as B^0 with amplitude $A_{\text{tag}} \equiv \langle f_{\text{tag}}|B^0\rangle = \langle\overline{f}_{\text{tag}}|\overline{B}^0\rangle$. At this moment, the wave function of the system collapses, and the flavor of the other particle must be $\overline{B}^0$ at the same instant, irrespective of their distance. This is the essence of the Einstein-Podolsky-Rosen (EPR) paradox [43] and the key principle allowing experimental observation of flavor through the second meson at the very moment when the time evolution from a pure $\overline{B}^0$ state starts. If this second meson then decays at time $t_2 = t_{CP}$ into a CP eigenstate f_{CP}, its wave function[27] just before this moment is

$$|\overline{B}^0(t_{\text{CP}})\rangle = \frac{A_{\text{tag}}}{\sqrt{2}}e^{-\frac{\Gamma}{2}(t_{\text{tag}}+t_{CP})}\left[\cos\left(\frac{\Delta m \Delta t}{2}\right)|\overline{B}^0\rangle + i\frac{p}{q}\sin\left(\frac{\Delta m \Delta t}{2}\right)|B^0\rangle\right] \tag{4.47}$$

and the amplitude of the process is $\langle f_{CP}|\overline{B}^0(t_{CP})\rangle$. This, together with the second case, where the flavor of the tag-side meson is $\overline{B}^0$, has the same form as Eq.4.18 with a replacement $t \to \Delta t = t_{CP} - t_{\text{tag}}$. Indeed, the resulting asymmetry can be again written as

$$a(\Delta t) = \frac{\Gamma(f_{CP}, f_{\text{tag}}) - \Gamma(f_{CP}, \overline{f}_{\text{tag}})}{\Gamma(f_{CP}, f_{\text{tag}}) + \Gamma(f_{CP}, \overline{f}_{\text{tag}})} = \mathcal{A}_{CP}\cos(\Delta m\Delta t) + \mathcal{S}_{CP}\sin(\Delta m\Delta t), \tag{4.48}$$

where $\Gamma(f_{CP}, f_{\text{tag}}) = |\langle f_{CP}|\langle f_{\text{tag}}|\Psi(t_{CP}, t_{\text{tag}})\rangle|^2$; equivalently for $\Gamma(f_{CP}, \overline{f}_{\text{tag}})$.

Technically, to extract maximum information and include reconstruction effects properly, the probability distributions for finding a given tag-side B meson flavor with time difference Δt are used instead of the asymmetry Eq. 4.48, which is then only constructed ad hoc. Using $\Gamma = 1/\tau_B$, where τ_B is the B^0 (or potentially B^+) lifetime, this distribution, properly normalized over $\Delta t \in (-\infty, +\infty)$ and the two flavors, can be written as

$$\mathcal{P}_{\text{phys}}(\Delta t, q) = \frac{e^{-|\Delta t|/\tau_B}}{4\tau_B}\left\{1 + q \times [\mathcal{A}_{CP}\cos(\Delta m\Delta t) + \mathcal{S}_{CP}\sin(\Delta m\Delta t)]\right\}, \tag{4.49}$$

---

[25]For example, the meson, which is going to the upper hemisphere.

[26]This outcome is random. The flavor is not predetermined.

[27]The total wave function after the tag side decay, which performs a quantum mechanical measurement by projecting out one of the eigenstates using $|f_{\text{tag}}\rangle\langle f_{\text{tag}}|$ (or $|\overline{f}_{\text{tag}}\rangle\langle\overline{f}_{\text{tag}}|$), is proportional to $\left[\cos\left(\frac{\Delta m \Delta t}{2}\right)|f_{\text{tag}}\rangle|\overline{B}^0\rangle + i\frac{p}{q}\sin\left(\frac{\Delta m \Delta t}{2}\right)|f_{\text{tag}}\rangle|B^0\rangle\right]$. This expression is now separable, and one can talk about two individual particles and wave functions in their respective Hilbert spaces. The state $|f_{\text{tag}}\rangle$ is, however, now composed of different particles (so it does not belong to the original product of Hilbert spaces to describe the two-particle system) and also evolves in time. Just after the decay, at $\Delta t = 0$, the state is proportional to a pure $|f_{\text{tag}}\rangle|\overline{B}^0\rangle$ and the second meson is in state $\overline{B}^0$ and starts mixing.

where q is given by the true tag-side flavor, $q = +1(-1)$ for $B_{\text{tag}} = B^0 (B_{\text{tag}} = \overline{B}^0)$, and $\Delta t = t_{CP} - t_{\text{tag}}$ can be both positive and negative.

4.6.2 Measuring Time with Vertex Detectors

One needs to measure the difference between the decay times of the B mesons, which seems very difficult, as they are produced nearly at rest after the $\Upsilon(4S)$ decay. However, if the original $\Upsilon(4S)$ moves very fast, the daughter B's can travel a measurable distance. This is the reasoning behind the asymmetric accelerator. With the center of mass of the collision moving in the laboratory frame, the produced $\Upsilon(4S)$ meson is boosted along the electron beam direction with $\beta\gamma = 0.425$ at Belle (0.28 at Belle II). Thanks to relativistic time dilation, the B mesons live a bit longer in the laboratory frame, further slightly increasing the distance traveled before their decay. Ignoring their small kinetic energy in the CM frame, the distance Δz between the two B decay vertices along the boost is related to the proper time difference Δt in the center of mass system of the particles as

$$\Delta t \cong \frac{\Delta z}{\beta\gamma c}. \tag{4.50}$$

The boost leads to an average distance between B^0 and $\overline{B}^0$ decay vertices of about 200 μm at Belle (130 μm at Belle II) along the z-direction, which can be resolved with high-precision semiconductor vertexing technology. An illustration of a typical decay chain with a golden channel $B^0 \to J/\psi K_S^0$ is shown in Fig. 4.8.

Many factors enter the experimental resolution of Δt, mostly through Δz, which must be determined from a vertex fit to the B decay products. These effects can be studied on simulations as a function of $\delta t = \Delta t - \Delta t_{\text{true}}$ and the estimated uncertainty $\sigma_{\Delta t}$. Note that Δt_{true} is the Δt from previous sections, while in this paragraph, Δt is the reconstructed proper-time difference. The uncertainty $\sigma_{\Delta t}$ follows from error propagation in tracking and vertexing and is determined for each particular event. There are more such variables affecting the resolution, for example, the flight direction of the B meson in the CM frame θ_B^{CM}, which is neglected in the kinematic approximation 4.50. If the resolution function $\mathcal{R}(\delta t; \sigma_{\Delta t}, \dots)$ [8] is determined, the probability density can be evaluated by convolution as

$$\mathcal{P}_{\text{rec}}(\Delta t, q; \sigma_{\Delta t}, \dots) = \mathcal{P}_{\text{phys}}(\Delta t, q) \otimes \mathcal{R}(\Delta t; \sigma_{\Delta t}, \dots)$$
$$= \int_{-\infty}^{+\infty} \mathcal{P}_{\text{phys}}(\Delta t_{\text{true}}, q) \mathcal{R}(\Delta t - \Delta t_{\text{true}}; \sigma_{\Delta t}, \dots) \, d\Delta t_{\text{true}}, \tag{4.51}$$

which is also a function of the event-dependent variables used to parameterize the resolution function. In other words, this is a conditional probability, and the total probability density function (PDF) can only be constructed once PDFs for the event-dependent variables (like $\sigma_{\Delta t}$) are specified. This also means it is not trivial to plot such a distribution—one needs to assume some distributions for the (conditional) event-dependent variables. Thus in all plots where fit results (utilizing Δt resolution functions) and data are compared in the next chapters, the event variables come from the data and are plugged into the probability density functions, evaluated at each event. The total distribution is thus a sum of many such functions (as many as there are events in the dataset).
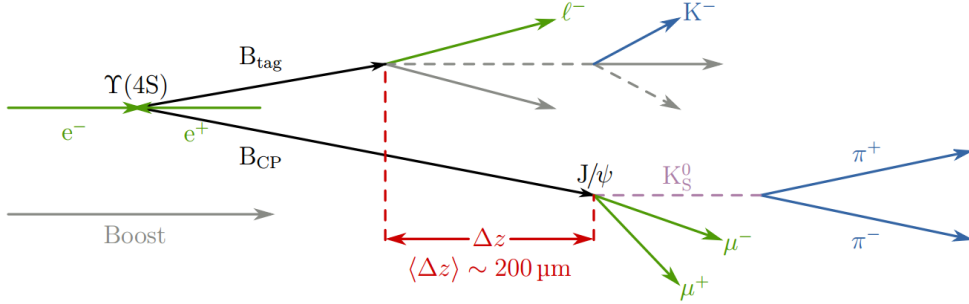
Figure 4.8: Illustration of the time difference and flavor measurement at Belle showing the decay chain for the golden channel, where neutral meson B_{CP} decays into $J/\psi \to \mu^+\mu^-$ and $K_S^0 \to \pi^+\pi^-$. The second meson emerging from $\Upsilon(4S)$, denoted as B_{tag} is used to tag the flavor of the second B, in this case through the high momentum lepton ℓ^-. The decay time difference of the two B mesons is inferred from the longitudinal distance Δz of their decay vertices [14].

4.6.3 Flavor Tagging

In Fig. 4.8, on the upper side (the tag side), the meson B_{tag} decays in a typical flavor-specific way into a single high-energy lepton. These decays predominantly proceed via a semileptonic $b \to c\ell^-\overline{\nu}_l$ transition, where the lepton ℓ carries the sign of the charge of the original b quark and thus $B_{\text{tag}} = (b\overline{d}) = \overline{B}^0$ in the figure. This allows us to identify the flavor with large confidence, but generally not unambiguously, as, in addition, there are other tagging modes, which can have a much lower probability of correct flavor assignment. For example, the lepton may be a secondary lepton originating from a cascade decay $b \to W^-c(c \to s\ell^+\nu)$. In this case, the charge of the lepton is opposite to the b quark charge, but its momentum spectrum is also much softer [8]. The procedure of flavor assignment based on the decay characteristic is called flavor tagging. Flavor tagging generally follows after signal reconstruction. The fully reconstructed signal-side particles are excluded, and the rest of the event is analyzed by the tagging algorithm.

At Belle, the output of the flavor tagging procedure is the estimated flavor charge $q = \pm1$, $q(B^0) = +1, q(\overline{B}^0) = -1$ of the tag-side B-meson and the expected flavor dilution factor r, representing the likelihood of correct flavor determination [8]. This factor ranges from zero (no flavor information) to one (unambiguous flavor assignment). The signed probability

$$q\,r = \frac{N(B^0) - N(\overline{B}^0)}{N(B^0) + N(\overline{B}^0)} \tag{4.52}$$

is obtained from a lookup table determined from MC simulation and calibrated by self-tagging decays on data. The average mistag probability $w = (w_{B^0} + w_{\overline{B}^0})/2$ is related to the dilution factor $(1-2w)$, which reduces the observed CP and mixing asymmetries from physical ones. As mistag probability can differ between flavors, $\Delta w = w_{B^0} - w_{\overline{B}^0}$ is further being determined. The difference between tagging efficiencies for B^0 and $\overline{B}^0$ is neglected at Belle. These tagging performance parameters are determined separately for seven tagging-quality categories (ranges of r). The categories and the corresponding average mistag averages and differences

for Belle data can be found in Table 5.1. With imperfect flavor determination and resolution effects, the probability density 4.49 changes to

$$\mathcal{P}(\Delta t, q) = \frac{e^{-|\Delta t|/\tau_B}}{4\tau_B}\Big\{1 - q\Delta w$$
$$+ q(1 - 2w)[\mathcal{S}_{CP}\sin(\Delta m\Delta t) + \mathcal{A}_{CP}\cos(\Delta m\Delta t)]\Big\} \otimes \mathcal{R}(\Delta t),$$
(4.53)

where q is the reconstructed (measured) B_{tag} flavor, and Δt is the reconstructed proper-time difference. In practice, the resolution function $\mathcal{R}$ (and thus $\mathcal{P}$) depends on a number of conditional event-dependent variables, like the Δz resolution from tracking. The reconstructed distribution is compared to the true physical probability density in Fig. 4.9. It should be noted that the *raw* asymmetry is further diluted by backgrounds. A proper model for the signal and background components, the resolution function(s), and the flavor tagging imperfections permit *unfolding* of the underlying physics parameters using a maximum likelihood fit.
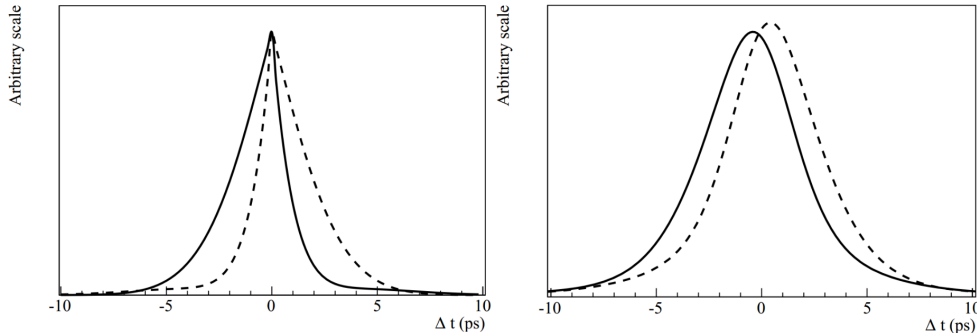
Figure 4.9: Physical distribution (left) of the time-dependent decay rate as a function of the proper time difference for B^0-tagged (solid line) and $\overline{B}^0$-tagged events (dashed line) and the observed distribution (right) after considering imperfect tagging and resolution effects. The *CP*-violation parameters are set to $\mathcal{S}_{CP} = 0.7$ and $\mathcal{A}_{CP} = 0$ and assuming $\eta_{CP} = +1$ (*CP*-even final state) [8].

4.7 Pseudoscalar Charmonium $\eta_c(1S)$ in $B \to \eta_c K$ with $\eta_c \to K\overline{K}\pi$

Bound states of a charm quark and anti-quark belong to a class of neutral mesons called charmonium, with its most famous member being J/ψ, discovered in 1974 [44]. The J/ψ meson, as well as many other states, shown in Fig. 4.10 together with their quantum numbers, are allowed to decay into two leptons; with a very distinctive experimental signature. The pseudoscalar η_c meson is, in fact, the lowest-lying ground s-state charmonium, but with $J^{PC} = 0^{-+}$, it cannot decay into two leptons due to angular momentum and C conservation. This makes η_c experimentally much more difficult to observe. The current world average for η_c mass is (2983.9 ± 0.5) MeV/c^2 [36]. The width $\Gamma_{\eta_c} = (32.0 \pm 0.8)$ MeV is significantly larger than for J/ψ, $\Gamma_{J/\psi} = (92.9 \pm 2.8)$ keV. The measured η_c

parameters differ significantly among different experiments and production and decay modes [8]. The smallest errors are coming from two-photon production of η_c, but even here, the results can be affected by interference with a non-resonant direct production of the $K\overline{K}\pi$ final state, through which η_c is usually observed.
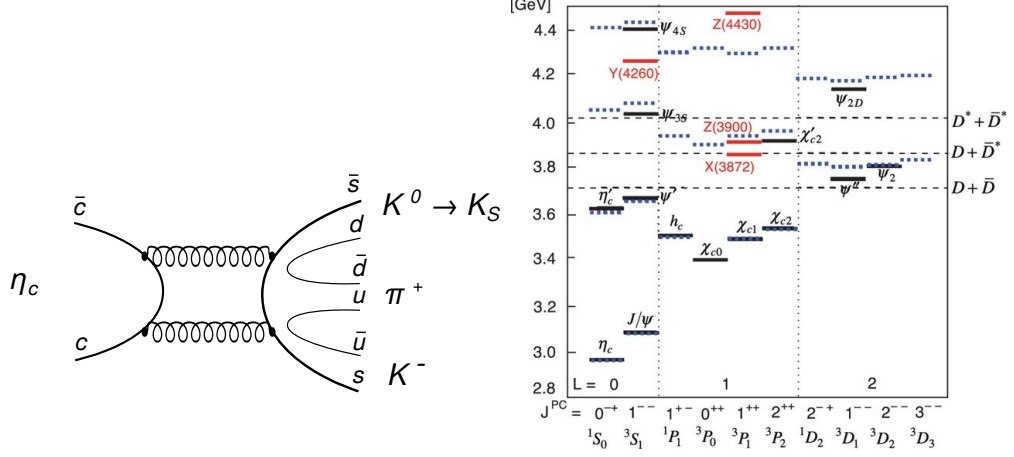
Figure 4.10: Example of Feynman diagram for a typical η_c decay, in this case to $K_S^0 K^- \pi^+$ (left) and the spectrum of charmonium (right) sorted according to mass and quantum numbers of the states [45].

The η_c charmonium primarily decays hadronically via annihilation into two gluons, suppressed by the Zweig (OZI) rule, as illustrated in Fig. 4.10. One of the dominant η_c decays is $\eta_c \to K\overline{K}\pi$. This notation includes several individual decay modes related by isospin symmetry. As isospin quantum numbers of η_c are $I = 0$, $I_3 = 0$, these must be conserved in a strong decay and in the final state. The three possible isospin projections for π with $I_3 = -1, 0, +1$ correspond to π^-, π^0 and π^+, respectively. The K states with $I_3 = -\frac{1}{2}, +\frac{1}{2}$ are K^0, K^+, and $K^-, \overline{K}^0$ for $\overline{K}$, respectively. The coefficients of combinations matching total $I = I_3 = 0$ and zero total charge are given by the Clebsh-Gordan coefficients as [46]

$$|K\overline{K}\pi\rangle = + \frac{1}{\sqrt{3}}|\overline{K}^0 K^+ \pi^-\rangle + \frac{1}{\sqrt{3}}|K^0 K^- \pi^+\rangle$$
$$- \frac{1}{\sqrt{12}}|K^+ K^- \pi^0\rangle - \frac{1}{\sqrt{12}}|K^- K^+ \pi^0\rangle \qquad (4.54)$$
$$- \frac{1}{\sqrt{12}}|K^0 \overline{K}^0 \pi^0\rangle - \frac{1}{\sqrt{12}}|\overline{K}^0 K^0 \pi^0\rangle$$

and thus, for decays without experimentally challenging π^0 in the final state, we have

$$|\langle K^0 K^- \pi^+ | K\overline{K}\pi\rangle|^2 = |\langle \overline{K}^0 K^+ \pi^- | K\overline{K}\pi\rangle|^2 = \frac{1}{3} \qquad (4.55)$$

This, however, still includes both half K_S^0 and half K_L^0, which are observed instead of K^0 and $\overline{K}^0$, where again K_L^0 is not so easy to detect and reconstruct. This leaves the most accessible mode for η_c reconstruction, for which the branching ratio will be related to $\mathcal{B}(\eta_c \to K\overline{K}\pi)$ as

$$\mathcal{B}(\eta_c \to K_S^0 K^\pm \pi^\mp) \equiv \mathcal{B}(\eta_c \to K_S^0 K^+ \pi^-) + \mathcal{B}(\eta_c \to K_S^0 K^- \pi^+) = \frac{1}{3}\mathcal{B}(\eta_c \to K\overline{K}\pi) \qquad (4.56)$$

The notation means that both charge-conjugated variants are considered, and this will be implied and not written explicitly anymore. We will only keep using $K_S^0 K^- \pi^+$ but mean both $(K_S^0 K^\pm \pi^\mp)$, as usual[28]. The current world average is [36]

$$\mathcal{B}(\eta_c \to K\overline{K}\pi) = (7.3 \pm 0.5)\%. \tag{4.57}$$

Note that we can just multiply the branching fraction by $\frac{1}{2}$ to get values for K_S^0 (K_L^0) with good accuracy if we effectively integrate over all possible K_S^0 decay times[29]. Then only K_L^0 survives after several $c\tau_{K_S^0}$ and thus one detects half of the original K^0 (or $\overline{K}^0$ from $\overline{B}^0$) via the decay $K_S^0 \to \pi^+\pi^-$, neglecting the very small pollution from the CP violating decays $K_L^0 \to \pi^+\pi^-$.

The decays $B \to c\bar{c}K$ are tree-dominated (see Sec. 4.10), and color-suppressed, resulting in relatively small branching fractions. The current world averages for the two lowest charmonia are [36]

$$
\begin{aligned}
\mathcal{B}(B^0 \to \eta_c K^0) &= (7.9 \pm 1.2) \times 10^{-4} \\
\mathcal{B}(B^0 \to J/\psi K^0) &= (8.73 \pm 0.32) \times 10^{-4} \\
\mathcal{B}(B^+ \to \eta_c K^+) &= (10.9 \pm 0.9) \times 10^{-4} \\
\mathcal{B}(B^+ \to J/\psi K^+) &= (10.10 \pm 0.29) \times 10^{-4}.
\end{aligned}
\tag{4.58}
$$

To get $\mathcal{B}(B^0 \to \eta_c K_S^0)$ and $\mathcal{B}(B^0 \to J/\psi K_S^0)$, one should again multiply the branching fractions by $\frac{1}{2}$. The charged decay $B^+ \to \eta_c K^+$ was also used [47] to extract η_c mass and width in presence of coherent interfering non-resonant background from $B^+ \to (K_S^0 K^+ \pi^-)K^+$ decays mediated by the penguin diagrams [8] (see also Sec. 4.10). While the effects of penguin pollution are estimated to be negligible or at least acceptable at the current precision [41] for CP-violation measurements, the effects on width can be significant [29]. Interference effects lead to modification of the resonance line shape and measured width, which can bias the apparent signal yield.

4.8 Previous Measurements and Physics Motivations

The mixing-induced ($\mathcal{S}$) and direct ($\mathcal{A}$) CP violation was studied in a number of charmonia states, see Fig. 4.11 and Fig. 4.12. The current world average for $\sin 2\phi_1$ is [34]

$$\sin 2\phi_1 = 0.699 \pm 0.017 \tag{4.59}$$

In the $\eta_c K_S^0$ decay mode, the latest BaBar result [48]

$$
\begin{aligned}
\mathcal{S}_{\eta_c K_S^0} &= +0.925 \pm 0.160 \pm 0.057_{\text{syst}} \\
\mathcal{A}_{\eta_c K_S^0} &= -0.080 \pm 0.124 \pm 0.029_{\text{syst}},
\end{aligned}
\tag{4.60}
$$

obtained with 465×10^6 $B\overline{B}$ pairs represents the current most precise measurement for η_c charmonium. Note that BaBar is using an opposite definition for the sign

---

[28]From time to time and in some figures, we repeat explicit form.

[29]This is done by default in the reconstruction having a vertex detector with a sufficient radius.

of the direct CP-violation parameter and denotes it as $\mathcal{C} = -\mathcal{A}$. At Belle, only a result with 151×10^6 $B\overline{B}$ pairs is available:

$$\mathcal{S}_{\eta_c K_S^0} = 1.126^{+0.27}_{-0.39} \pm 0.06_{\text{syst}}. \tag{4.61}$$

This analysis is a direct continuation of the effort started by Z. Drásal, who used the full Belle dataset to measure $\mathcal{S}_{\eta_c K_S^0}$, $\mathcal{A}_{\eta_c K_S^0}$ in the relatively clean decay channel $\eta_c \to p\bar{p}$ with the result:

$$\begin{aligned} \mathcal{S}_{\eta_c K_S^0} &= 0.68^{+0.38}_{-0.46} \pm 0.13_{\text{syst}} \\ \mathcal{A}_{\eta_c K_S^0} &= 0.00^{+0.23}_{-0.31} \pm 0.08_{\text{syst}} \end{aligned} \tag{4.62}$$

published in his doctoral thesis [40].

	Parameter: sin(2β) ≡ sin(2φ₁)			
Mode	**BaBar**	**Belle**	**LHCb**	**Average**
	N(BB)=465M	N(BB)=772M	∫Ldt=3 fb⁻¹	
J/ψK$_S$ (η$_{CP}$=-1)	0.657 ± 0.036 ± 0.012	0.670 ± 0.029 ± 0.013	0.731 ± 0.035 ± 0.020 0.83 ± 0.08 ± 0.01$^{(*)}$	0.695 ± 0.019 (0.018$_{stat-only}$)
J/ψK$_L$ (η$_{CP}$=+1)	0.694 ± 0.061 ± 0.031	0.642 ± 0.047 ± 0.021	-	0.663 ± 0.041 (0.037$_{stat-only}$)
J/ψK^0	0.666 ± 0.031 ± 0.013	-	0.75 ± 0.04	0.690 ± 0.018 (0.017$_{stat-only}$)
ψ(2S)K$_S$ (η$_{CP}$=-1)	0.897 ± 0.100 ± 0.036	0.738 ± 0.079 ± 0.036	0.84 ± 0.10 ± 0.01	0.817 ± 0.056 (0.053$_{stat-only}$)
ψ(nS)K^0	-	-	0.760 ± 0.034	0.701 ± 0.017 (0.016$_{stat-only}$)
χ$_{c1}$K$_S$ (η$_{CP}$=-1)	0.614 ± 0.160 ± 0.040	0.640 ± 0.117 ± 0.040	-	0.632 ± 0.099 (0.094$_{stat-only}$)
η$_c$K$_S$ (η$_{CP}$=-1)	0.925 ± 0.160 ± 0.057	-	-	-
J/ψK*^{0} (K*^{0} → K$_S$π^0) (η$_{CP}$= 1-2|A$_\perp$|²)	0.601 ± 0.239 ± 0.087	-	-	
All charmonium	0.687 ± 0.028 ± 0.012	0.667 ± 0.023 ± 0.012	0.760 ± 0.034	0.698 ± 0.017 (0.016$_{stat-only}$)
χ$_{c0}$K$_S$ (η$_{CP}$=+1)	0.69 ± 0.52 ± 0.04 ± 0.07 $^{(**)}$ N(BB)=383M	-	-	-
J/ψK$_S$, J/ψ → hadrons (η$_{CP}$=+1)	1.56 ± 0.42 ± 0.21 $^{(***)}$ N(BB)=88M	-	-	-
All charmonium (incl. χ$_{c0}$K$_S$ etc.)	0.691 ± 0.031 (0.028$_{stat-only}$)	0.667 ± 0.023 ± 0.012	0.760 ± 0.034	0.699 ± 0.017 (0.016$_{stat-only}$)

Figure 4.11: Current experimental status of $\mathcal{S} = \sin 2\phi_1$ measurement in $b \to c\bar{c}s$ for modes with different charmonia [49].

	Parameter: C=−A (if not stated otherwise)			
Mode	**BaBar**	**Belle**	**LHCb**	**Average**
	N(BB)=465M	N(BB)=772M	∫Ldt=3 fb⁻¹	
J/ψK$_S$	0.026 ± 0.025 ± 0.016	0.015 ± 0.021 $^{+0.023}_{-0.045}$	−0.038 ± 0.032 ± 0.005$^{(†)}$ 0.12 ± 0.07 ± 0.02$^{(*)}$	0.000 ± 0.020 (0.014$_{stat-only}$)
J/ψK$_L$	−0.033 ± 0.050 ± 0.027	−0.019 ± 0.026 $^{+0.041}_{-0.017}$	-	−0.023 ± 0.030 (0.023$_{stat-only}$)
J/ψK^0	0.016 ± 0.023 ± 0.018	-	−0.014 ± 0.030	−0.007 ± 0.018 (0.012$_{stat-only}$)
ψ(2S)K$_S$	0.089 ± 0.076 ± 0.020	−0.104 ± 0.055 $^{+0.027}_{-0.047}$	−0.05 ± 0.10 ± 0.01	−0.019 ± 0.048 (0.041$_{stat-only}$)
ψ(nS)K^0	-	-	−0.017 ± 0.029	−0.008 ± 0.017 (0.012$_{stat-only}$)
χ$_{c1}$K$_S$	0.129 ± 0.109 ± 0.025	0.017 ± 0.083 $^{+0.026}_{-0.046}$	-	0.066 ± 0.074 (0.066$_{stat-only}$)
η$_c$K$_S$	0.080 ± 0.124 ± 0.029	-	-	-
J/ψK*^{0} (K*^{0} → K$_S$π^0)	0.025 ± 0.083 ± 0.054	-	-	
All charmonium	0.024 ± 0.020 ± 0.016	−0.006 ± 0.016 ± 0.012	−0.017 ± 0.029	−0.005 ± 0.015 (0.012$_{stat-only}$)
χ$_{c0}$K$_S$ (η$_{CP}$=+1)	−0.29 $^{+0.53}_{-0.44}$ ± 0.03 ± 0.05 $^{(**)}$	-	-	-
All charmonium (incl. χ$_{c0}$K$_S$)	0.023 ± 0.025 (0.020$_{stat-only}$)	−0.006 ± 0.016 ± 0.012	−0.017 ± 0.029	−0.005 ± 0.015 (0.012$_{stat-only}$)

Figure 4.12: Current experimental status of $\mathcal{C} = -\mathcal{A}$ measurement in $b \to c\bar{c}s$ for modes with different charmonia [49].

In contrast to the vector charmonium J/ψ, η_c is a pseudoscalar. Potential NP scenarios might show up as a difference between observed CP violation among

different (vector and pseudoscalar or axial vector) charmonia if parity and CP are violated by NP contributions [50]. The spectrum of charmonia is shown (along with some un-conventional states) in Fig. 4.10. One can see that only pseudoscalar states are η_c and $\eta_c(2S)$ (denoted as η_c'). The available statistics for $\eta_c(2S)$ are very limited. Therefore the η_c charmonium is the only mode that is reasonably accessible experimentally with the current dataset to test the universality of $\sin 2\phi_1$ among vector and pseudoscalar charmonium.

The NP contribution at scale $\mathcal{O}(10\%)$ leading to a difference between $\sin 2\phi_1$ from $B^0 \to \eta_c K_S^0$ and vector charmonia proposed in Ref. [50] are not excluded by the current BaBar measurement itself and are smaller than the current uncertainty. Therefore an independent measurement of CP violation in $B^0 \to \eta_c K_S^0$, with uncorrelated instrumental errors, is highly desirable. Even with a comparable statistical precision, the constraint on the potential deviation from SM prediction can be tightened. Needless to say the consideration in Ref. [50] were motivated by observed possible deviations of $\sin 2\phi_1$ for $B^0 \to \phi K_S^0$ from charmonium measurements. Since that time (2003), these measurements are now consistent with the SM predictions (current world average for $\sin 2\phi_1$ in $B^0 \to \phi K_S^0$ is $0.74^{+0.11}_{-0.13}$ [49]), and only much smaller deviations are allowed instead of initially assumed NP corrections of order one. In the considered physics model, a new type of Z boson denoted as Z' with effective $\bar{s}Z'b$ couplings would enter via penguin $b \to s$ contributions. SM-like couplings $\bar{q}Z'q$ are assumed, with vector and axial currents, which have different coupling constants. This difference then could allow observing NP contributions most comfortably in the $c\bar{c}$ system, as it has the biggest spread of couplings among axial and vector currents [50]. Nowadays, Z'-like particle is of interest due to its leptonic interaction, being one of the possible explanations for the non-universality of leptons observed in $b \to s\ell^+\ell^-$ transitions at LHCb [12]. This shows that the selected topic is still of interest and related to an active area of research, despite constraints leading to the original motivations have been significantly tightened in the last two decades.

4.9   Methodology

This analysis is based on a very similar methodology and tools as in Ref. [40], but using $\eta_c \to K_S^0 K^\pm \pi^\mp$ decay channel with a larger branching fraction. In fact, Z. Drásal finished the analysis of $\eta_c \to K_S^0 K^+ \pi^-$ (charge conjugation is implied) up to all selection optimizations and continuum and peaking background suppression. This work builds on many tools developed by Z. Drásal for his analysis, which he kindly provided, along with skimmed and reconstructed Belle Monte Carlo (MC) simulation and real data and all preliminary results. The decay channel $\eta_c \to K_S^0 K^+ \pi^-$ promises smaller statistical uncertainty than $\eta_c \to p\bar{p}$, but the analysis is more challenging due to a higher level and more complicated composition of the backgrounds. From preliminary toy MC studies, the final uncertainty of CP-violation parameters was expected to be very close to the BaBar result, both for statistical and systematic uncertainty.

In most Belle analyses, only CP-violation parameters are floating in a fit to the Δt distribution with an event-by-event probability obtained from a one or more-dimensional fit, typically in beam-constrained mass (M_{bc}) and energy

difference (ΔE) distributions (see Eq. 5.9 for definitions). This fit is performed to extract fractions of signal and background(s), which are then fixed in the fit for *CP*-violation parameters. The *CP* fit is then usually performed in a much tighter signal window (in M_{bc} and ΔE), where the respective fractions of signal and background need to be (once) calculated by integrating the model (without Δt dimension) probability density functions.

This analysis aims to perform the fit simultaneously in the dimensions that allow us to extract signal and background fractions and in the time and flavor dimensions $\Delta t \times q$. The primary motivation is to handle appropriately the systematic uncertainties that naturally occur in the traditional approach, where, in the end, one has to vary all fixed parameters used in the *CP* fit. The number of such parameters is reduced in our approach as they are floated in the final fit. Thus the uncertainty of such parameters is automatically propagated to the statistical errors of the *CP*-violation parameters. This seems especially important in the case of the peaking background contribution, which is often fixed from MC studies. In our case, a significant part of the background comes from $b \to u, d, s$ transitions, where MC cannot be used for reliable estimates of the total composition.

In general, the generic MC data samples are an approximation. The branching fractions set in large configuration files (decay files) used for particle decays in generators are often outdated or only leading order theoretical predictions; some decay channels may be missing, and so forth. The generic $B\overline{B}$ MC sample is much better understood (validated) as it is a necessary ingredient to perform signal simulations, where one of the B mesons decays generically. The generic $B \to u, d, s$ MC sample is also called rare-B sample as these processes generally have even lower branching fractions than $b \to c$. We only draw rough estimates and qualitative conclusions from the generic MC samples and compare these to results from data, where possible. The aim is to estimate the fraction of the total peaking background from a direct fit to the data, albeit with sizeable statistical error and potentially significant fit bias. Keeping the fraction floating in the final fit allows directly taking such effects into account in the (statistical) error estimation for the parameters of interest.

In addition, and in contrast to the BaBar and previous Belle analysis, M_{bc} is not used as a fit variable but for the best B-candidate selection. This complicates the analysis due to a higher level of signal cross-feed, which needs to be modeled and estimated from MC. The other potential variable for the best candidate selection is the invariant mass of the $K_S^0 K^\pm \pi^\mp$ three-body final state or quality of the η_c vertex fit, which may be, however, correlated with the determined B vertex used for Δt measurement and can cause potential bias. Therefore the same strategy as for $\eta_c \to p\bar{p}$ [40] is employed, with M_{bc} used to choose the best B candidate in each event, and the reconstructed invariant mass of the η_c candidate and energy difference ΔE as fit variables. We call this a 2D model (and fit) and build upon it the full complete simultaneous model in $\Delta E \times M(K_S^0 K^+ \pi^-) \times \Delta t \times q \times l$, where q is the B_{tag} flavor (± 1) and $l \in \{0 \ldots 6\}$ is an index labeling bins in the dilution factor r (likelihood of correct flavor assignment), which we call a 4D model (fit). Only three dimensions are continuous. The discrete dimensions $q \times l$ can be counted as zero, one or two more dimensions formally. However, the model for the q dimension is trivial.

We follow the standards for blind analysis at Belle. In this approach, every detail of the analysis and data model must be defined before the final fit for CP-violation parameters. Thus after event selection and background suppression, the simpler 2D model is studied first. Due to the two different vertex detector configurations at Belle, SVD1 ($152 \times 10^6 B\overline{B}$ pairs) and SVD2 ($620 \times 10^6 B\overline{B}$ pairs), all analysis steps and parameters are extracted for both experiments separately. We also employ a control, charged mode $B^\pm \to \eta_c(K_S^0 K^\pm \pi^\mp)K^\pm$. All analysis steps are first probed on the control mode before looking at the signal data. Only once the 2D model is verified by measuring the branching fractions for the control and signal mode, we move to time-dependent studies and simultaneous fits to SVD1 and SVD2 datasets. The time-dependent CP-violation measurement is first exercised on the control mode, with null results expected. Only once the data model is completely defined and validated can the analysis be un-blinded, and the final measurement of CP violation in the signal mode can be performed.

Before un-blinding the analysis, an internal Belle note needed to be prepared and reviewed by a dedicated review committee. This part of the thesis is mostly based on this internal note but leaves out some details of individual intermediate fits. This is to save space and limit the amount of almost identical plots, which stem from the fact that each result needs to be extracted for the control and signal mode, as well as for the SVD1 and SVD2 experiments.

CHAPTER 5

Event Reconstruction and Background Suppression

This chapter will summarize the steps to obtain the final set of candidates for maximum likelihood fits. In real experimental data, $\Upsilon(4S)$ meson decays are only a minority of registered events. Besides trigger selections, further requirements are posed by specialized physics skims. In particular, the `HadronB(J)` skim is applied for all data in this thesis, which retains $\sim 99\%$ of $B\overline{B}$ events and reduces non-hadronic backgrounds to a negligible level [51]. For Monte Carlo (MC) simulations, event reconstruction, and analysis, the Belle Analysis Software framework (basf) based on GEANT3 [52] simulation toolkit was used. It should be noted that all this work was already done by Z. Drásal and provided as skimmed and reconstructed data. Thus this chapter is mostly included for consistency and describes the previous work, which was reviewed and in some cases repeated, but in the end, almost all the selections in this chapter have been already defined and optimized before the analysis was taken over by the author.

First, the signal reconstruction from final state particles is discussed, followed by the definition of fit variables, signal and sideband regions, the best B-candidate selection, flavor tagging, vertex, and proper-time difference reconstruction. The following sections summarize the continuum suppression method, vetos for peaking backgrounds, and the mix of remaining backgrounds, in particular $b \to c$-induced transitions. The final signal efficiency (and expected number of signal events) and all selection criteria are summarized at the end of this chapter in Tables 5.5 and 5.4.

5.1 Event Selection

All final state particles in this analysis are charged tracks. The decay chain

$$
\begin{aligned}
B^0 \to &\eta_c K_S^0 \\
&\eta_c \to K_S^0 K^+ \pi^-
\end{aligned}
\tag{5.1}
$$

illustrated in Fig. 5.1 is reconstructed gradually from the final state particles. Note that charge conjugation is always implied and usually not written explicitly. As a natural control sample, we employ the charged B decay

$$
B^+ \to \eta_c K^+,
\tag{5.2}
$$

where no CP-violation effects are expected (charged B meson cannot mix, and no direct CP violation is expected). The control (charged) mode is used for overall
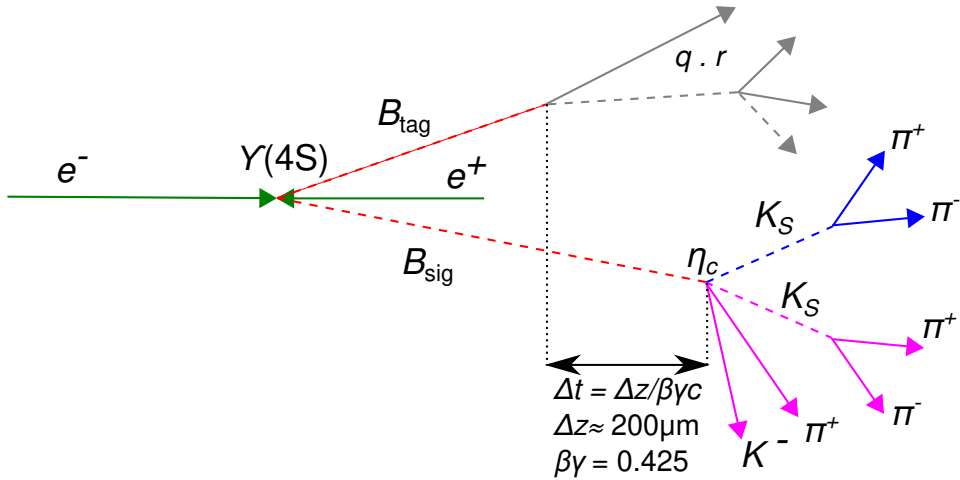
Figure 5.1: Illustration of a signal event with decay of primary $\Upsilon(4S)$ emerging from the e^+e^- collision into a signal B_{sig} followed by $B^0_{\text{sig}} \to \eta_c K^0_S, \eta_c \to K^0_S K^- \pi^+$, and tag B_{tag} meson. Paths of neutral particles are dashed, and there is no path for the very short-lived $\Upsilon(4S)$ and η_c resonances.

validation of the analysis with real data before un-blinding and for extraction of data/MC calibration factors for the measured signal (neutral) mode.

We reconstruct K^0_S using the

$$K^0_S \to \pi^+\pi^- \tag{5.3}$$

decay mode (branching ratio $\sim 70\%$). The K^0_S candidates are reconstructed by `V0finder` which combines oppositely charged tracks with invariant mass within ± 30 MeV/c^2 of $m_{K^0_S}$ without any additional selection [53]. To greatly reduce the background, a so-called `goodKs` cut is applied, and we keep candidates with invariant mass in the range

$$482 \text{ MeV}/c^2 < M(\pi^+\pi^-) < 514 \text{ MeV}/c^2, \tag{5.4}$$

which corresponds to approximately four times $\pm\sigma$ of the K^0_S mass distribution ($\sigma = 3.74 \pm 0.05$ MeV/c^2). This range is based on the results of an independent study of $K^0_S \to \pi^+\pi^-$ reconstruction [53]. In this study, the K^0_S mass peak was fitted using a double Gaussian (with the same mean and widths σ_1, σ_2) and first-order Chebyshev polynomial to account for background. The resulting sigma used in this study is defined as $\sigma = \sqrt{(A_1\sigma_1^2 + A_2\sigma_2^2)/(A_1 + A_2)}$, where A_i are amplitudes of the respective Gaussians.

For the tracks of charged pions and kaons, the impact parameters relative to the calibrated IP position in the transverse (dr) and longitudinal (dz) direction are required to satisfy

$$\begin{aligned}
|dr| &< 0.5 \text{ cm, and} \\
|dz| &< 3 \text{ cm.}
\end{aligned} \tag{5.5}$$

At least one of the tracks must have at least two SVD hits in the z direction and at least one SVD hit in the $R - \phi$ direction (to properly reconstruct the B_{sig} vertex). To distinguish pions from kaons, a combined information from particle identification system (from ACC and TOF) and dE/dx (from CDC) is used to

form 'binary' likelihood ratios $\mathcal{L}_{i/j} = \mathcal{L}_i/(\mathcal{L}_i + \mathcal{L}_j)$, where $\mathcal{L}_k$ is likelihood that a particle is of type k. We require:

$$\mathcal{L}_{K/\pi} \geq 0.6 \text{ for kaons,} \\ \mathcal{L}_{K/\pi} < 0.9 \text{ for pions}$$ (5.6)

and probability of electron hypothesis (from ECL)

$$p(e) < 0.95$$ (5.7)

for both kaons and pions.
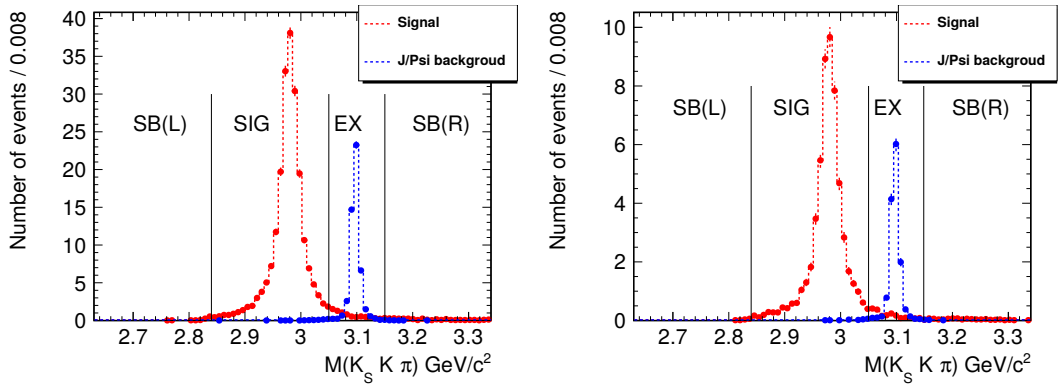
Figure 5.2: Distribution of the reconstructed η_c invariant mass obtained from ten streams (one stream corresponds to the recorded data luminosity) of generic $B\overline{B}$ MC for control mode (left) and signal mode (right). Only correctly reconstructed signal is shown (red) together with background (blue) stemming from J/ψ decays from $B^+ \to K^+ J/\psi$ (left - charged mode) and $B^0 \to K_S^0 J/\psi$ (right - neutral mode). The involved decays are $J/\psi \to K_S^0 K^+ \pi^-$, $J/\psi \to K^{*+} K^-$ or $J/\psi \to K^{*0} K_S^0$. Decays with mis-reconstructed J/ψ (where kaon from J/ψ is switched with a kaon from B meson) are not included (these contribute to the ΔE peaking background, flat in $M(K_S^0 K^+ \pi^-)$). The histograms are scaled to the nominal integrated luminosity, and both experiments (SVD1 and SVD2) are combined. Vertical lines separate the left and right sideband regions (SB), the signal region (SIG), and the excluded region (EX) for the J/ψ decays.

The η_c resonance is reconstructed by combining K_S^0, K^+ and π^- candidates, which fall into an invariant mass window given as

$$2.84 \text{ GeV}/c^2 < M(K_S^0 K^+ \pi^-) < 3.05 \text{ GeV}/c^2.$$ (5.8)

The right cut is chosen to avoid interference with J/ψ and its prominent peak in the reconstructed invariant mass spectrum, which is also avoided by the right sideband region as depicted in Fig. 5.2. The boundaries for the signal window in $M(K_S^0 K^\pm \pi^\mp)$ have been studied on ten streams (one stream corresponds to the recorded data luminosity) of generic $B\overline{B}$ (with signal strength not adjusted to the current world average). As can be seen in Fig. 5.2, the signal distribution is very wide, but its width is still being underestimated in our MC samples, which are generated with a smaller width for the η_c resonance; see also Table 6.1.

Considering the larger value from PDG, a four sigma window around the central value of η_c mass would correspond to a left cut of about 2.85 GeV/c^2. We made it looser by 10 MeV/c^2 in an attempt to improve fit stability and background shape determination (in the presence of signal) in the $M \equiv M(K_S^0 K^+ \pi^-)$ variable. It would be helpful to have some (even limited) range where no signal is expected. However, for this analysis, such efforts are complicated by the non-negligible amount of mis-reconstructed signal events, not shown in the figure but discussed in detail in Sec. 6.1.4.

What concerns the right-side cut value, it was chosen with a safety margin, which results in an expected contribution of only several J/ψ background events (less than 2%) in the signal region. The reason for such a conservative cut is a complete absence of any modeling of the interference of J/ψ and η_c, which can modify the line shape. This is expected to be a concern, so we prefer a small loss of signal events in favor of avoiding the need to take the interference with J/ψ into account.

At this point, the candidate B_{sig} is created by combining the η_c and K_S^0 ($K^{\pm}$ for control mode) candidates' reconstructed four-momenta. The reconstructed B_{sig} meson can be described by two almost orthogonal kinematic observables used at Belle: beam-constrained mass M_{bc} and energy difference ΔE. These are defined as:

$$M_{\text{bc}} = \sqrt{(E_{\text{beam}}^{\text{CM}})^2 - (p_B^{\text{CM}})^2}$$
$$\Delta E = E_B^{\text{CM}} - E_{\text{beam}}^{\text{CM}}, \tag{5.9}$$

where $E_{\text{beam}}^{\text{CM}}$ is the beam energy in the CM frame ($E_{\text{beam}}^{\text{CM}} = (E_{e^+}^{\text{CM}} + E_{e^-}^{\text{CM}})/2$) and E_B^{CM} (p_B^{CM}) is the reconstructed energy (momentum) of B_{sig} in the CM frame. While ΔE peaks at zero for correctly reconstructed B mesons, M_{bc} peaks at the B meson mass. We define the signal region by

$$5.271 \text{ GeV}/c^2 < M_{\text{bc}} < 5.290 \text{ GeV}/c^2. \tag{5.10}$$

We use M_{bc} for best candidate selection among all reconstructed B_{sig}. The average multiplicity[1] is 1.049 (1.089) for $B^{\pm}$ (B^0), see Fig. 5.3. Thus the best candidate has the minimum difference $|M - m_B|$, where $m_B = (5279.25 \pm 0.26)$ MeV/c^2 [34] [2] (the exact value used in the analysis is 5279.15 MeV/c^2).

With M_{bc} used for the best candidate selection, we are left with two sensible observables to characterize the kinematics of our signal events, which we use as (2D) fit variables:

- energy difference ΔE, for which we set a fit window

$$-0.1 \text{ GeV} < \Delta E < +0.1 \text{ GeV}, \tag{5.11}$$

- reconstructed invariant mass of the η_c candidate, further denoted as $M(K_S^0 K^+ \pi^-)$ or M, in the full signal window defined above:

$$2.84 \text{ GeV}/c^2 < M \equiv M(K_S^0 K^+ \pi^-) < 3.05 \text{ GeV}/c^2 \tag{5.12}$$

---

[1]We have later also checked the multiplicity for real data, giving 1.041 (1.079) for the charged (neutral) mode in good agreement with the simulation.

[2]The small difference of 0.3 MeV of neutral and charged B meson mass is far smaller than detector resolution, so the exact value is irrelevant and within uncertainties of the PDG inputs.
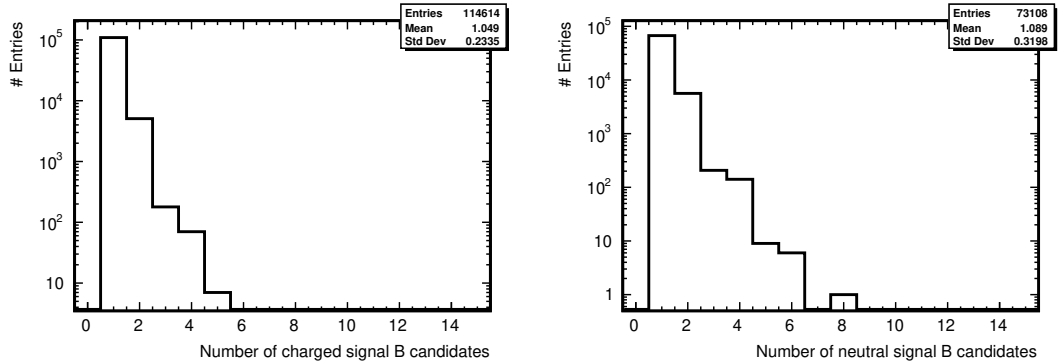
Figure 5.3: *B* candidate multiplicity determined from signal MC for charged (left) and neutral (right) mode. MC data from both experiments (SVD1 and SVD2) are combined in these plots.

In addition, we define two sideband regions, later denoted as $M_{\rm bc}$ sideband and $M(K^0_S K^+ \pi^-)$ sideband, set as follows:

$$M_{\rm bc} \text{ sideband} : 5.200 \text{ GeV}/c^2 < M_{\rm bc} < 5.265 \text{ GeV}/c^2$$
$$M(K^0_S K^+ \pi^-) \text{ sideband} : M \in (2.63; 2.84) \text{ GeV}/c^2 \cup (3.13; 3.34) \text{ GeV}/c^2 \quad (5.13)$$

At this point, we have a single signal B candidate ($B_{\rm sig}$) fully reconstructed. The rest of the event is used to determine the flavor of the second B ($B_{\rm tag}$) and its vertex from the remaining tracks. At Belle, this is handled by the flavor tagging procedure [54]. The output of the flavor tagging is the flavor of the tag-side B meson and a dilution factor $r \in [0, 1]$, which quantifies the likelihood of a correctly determined flavor, see also Sec. 4.6.3. The tagging is calibrated in seven separate r-bins, defined in Table 5.1, for which the difference between wrong tag probabilities Δw and mistag probability w are separately determined for MC and data (using self-tagging B decays). All these values are taken from a dedicated Belle study [54] and are also nicely summarized in Ref. [40]. For the best sensitivity to CP-violation parameters, PDFs and yields are parametrized for each r-bin separately (with common CP parameters), and the fit is performed simultaneously in all r-bins.

Because for the measured (neutral) mode, the K^0_S (with non-negligible flight length) cannot be used for vertex reconstruction, we do not use its counterpart $K^\pm$ for the control mode as well. Therefore in both cases, the vertex of the final reconstructed $B_{\rm sig}$ is entirely defined by the reconstructed vertex position of η_c (which is again determined only by charged tracks from kaon and pion) with additional information about calibrated IP position included via an IP-tube constraint (in the $x - y$ plane) in the vertex fit. The IP constraint is smeared in the $x - y$ plane by 21 μm to account for the small transverse B movement. To ensure a good proper time reconstruction, we require for the obtained goodness-of-fit h [55], and estimated reconstruction error in the z-coordinate for both, B_{sig} and B_{tag}

$$h \equiv (\chi^2_{\rm w/o\ IP})/\text{NDF} < 50$$
$$\sigma_z < 200 \ \mu\text{m} \quad (5.14)$$

r-bin index l	Range of r_l	SVD1 w_l	SVD1 Δw_l	SVD2 w_l	SVD2 Δw_l
0	$0.000 - 0.100$	0.5000	$+0.0000$	0.5000	$+0.0000$
1	$0.100 - 0.250$	0.4189	$+0.0570$	0.4188	-0.0088
2	$0.250 - 0.500$	0.3299	$+0.0126$	0.3193	$+0.0104$
3	$0.500 - 0.625$	0.2339	-0.0148	0.2229	-0.0109
4	$0.625 - 0.750$	0.1706	-0.0006	0.1632	-0.0186
5	$0.750 - 0.875$	0.0998	$+0.0089$	0.1041	$+0.0017$
6	$0.875 - 1.000$	0.0229	$+0.0047$	0.0251	-0.0036

Table 5.1: Definition of r-bins and data-driven values for average wrong-tag fractions w_l and wrong tag fraction differences Δw_l per every r-bin index l. More details, including uncertainties of w_l and Δw_l, as well values used for MC, can be found in Ref. [40].

for multi-track vertices, and

$$\sigma_z < 500 \ \mu\text{m} \tag{5.15}$$

for single-track vertices. Here $\chi^2_{\text{w/o IP}}$ (NDF) is the χ^2 (number of degrees of freedom) for a vertex fit without the IP constraint.

The difference in the decay times of B_{sig} and B_{tag} can be inferred from their distance in z due to the boosted center-of-mass system and very small kinetic energy of the daughter B mesons in $\Upsilon(4S)$ rest frame. Within a good approximation, the proper time difference can be obtained after a boost to the CM frame as

$$\Delta t \approx \frac{z_{\text{sig}} - z_{\text{tag}}}{c\beta\gamma}, \tag{5.16}$$

where $\beta\gamma$ is approximately 0.425 at KEKB. We require all events to satisfy

$$|\Delta t| < 70 \ \text{ps} \tag{5.17}$$

to remove very far outliers due to mis-reconstruction.

5.2 Continuum Background Suppression

The dominant background in the analysis originates from continuum processes $e^+e^- \to q\bar{q}$, where quarks other than b are involved ($q = u, d, s, c$). While in the case of $e^+e^- \to \Upsilon(4S) \to B\overline{B}$, the B mesons are nearly at rest in the CM frame, the lighter quarks tend to have significant momentum. This is reflected in the topology of the events. While B decays tend to be spherically symmetric, the continuum events have a jet-like structure. To characterize this qualitative statement quantitatively, we construct several event-shape variables from the event information related to Fox Wolfram moments [56]. The l-th Fox-Wolfram moment is defined as

$$H_l = \sum_{i,j} \frac{|p_i||p_j|}{E^2} P_l(\cos\theta_{ij}), \tag{5.18}$$

where P_l is the l-th Legendre polynomial, E is the total visible energy in the event, θ_{ij} is the opening angle between particle i and j and p_i is momentum of particle i. The sum runs over all particle indices[3] i, j from charged tracks in the event. All quantities shall be expressed in the CM frame. We utilize the following ten discriminating variables:

- The thrust angle $\cos\theta_{\text{thr}}$, defined as a scalar product of the thrust vector formed from the momenta of particles of the signal B meson and the rest of the event.

- Reduced Fox-Wolfram moment R_2, defined as

$$R_2 = H_2/H_0 \tag{5.19}$$

is a classical discriminating variable which could be sufficient to separate B decays from $q\bar{q}$ background in simple analyses. The distributions of R_2 for signal and background events are shown in Fig. 5.4.

- Eight modified Fox-Wolfram moments, H_l^{so} and H_l^{oo}, $l = 0, 1, 2, 3$, where the first letter in the superscript s (o) means that indices i run over particles assigned to signal (other) B meson. Equivalently for the second superscript and index j.

The modified moments of type H_l^{ss} are not used because some significant and generally larger correlations have been found with observables used as fit variables or for the best candidate selection. This can be seen in Fig. 5.5, where correlation coefficients are computed for the event-shape variables and the fit variables, as well as the best-candidate selection variable M_{bc}. We have also checked correlations to the Δt variable and found no correlations larger than 1.3%. The unused, modified Fox-Wolfram moments H_l^{ss} are shown in the last four rows. Any significant correlations with fit variables might generate additional systematics when continuum suppression cut (see below) is varied. One can see that the selected ten variables have correlations smaller than 2%. The only significant correlation is about 10% for H_0^{so} and M_{bc}, which is, however, not problematic, as M_{bc} is not used as a fit variable.

For each event, all the above ten topology variables' values, denoted as x_i, $i = 1..10$ are combined in the Fisher linear discriminant function, defined as

$$\mathcal{F} = \sum_{i=1}^{10} w_i x_i, \tag{5.20}$$

where w_i are weights that provide optimal separation between continuum background (combination of generic Belle u, d, s and charm continuum MC was used) and signal MC. The distribution of the optimized value of $\mathcal{F}$ is fitted with a sum of two bifurcated Gaussians for the signal or a single bifurcated Gaussian for the background:

$$\begin{aligned}
\mathcal{P}_{sig}(\mathcal{F}) &= f\mathcal{G}(\mathcal{F}; \mu^{sig}, \sigma_L^{sig}, \sigma_R^{sig}) + (1-f)\mathcal{G}(\mathcal{F}; \mu^{sig}, \sigma_L^{sig}k_L, \sigma_R^{sig}k_R) \\
\mathcal{P}_{bkg}(\mathcal{F}) &= \mathcal{G}(\mathcal{F}; \mu^{bkg}, \sigma_L^{bkg}, \sigma_R^{bkg}).
\end{aligned} \tag{5.21}$$
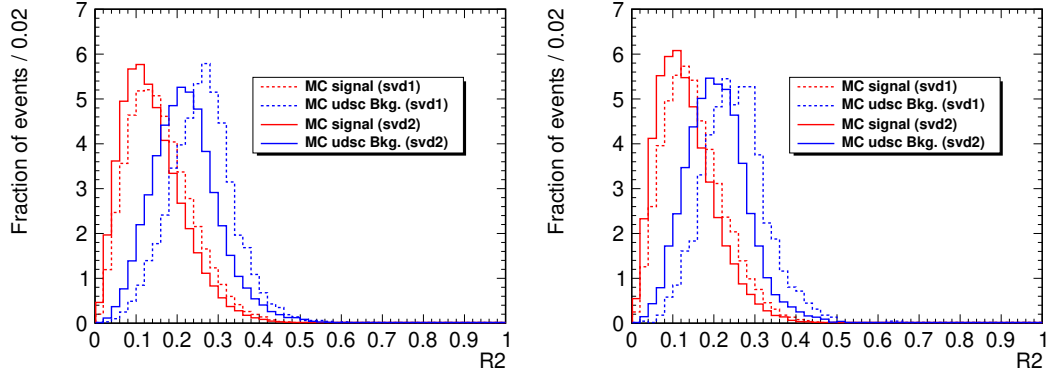
Figure 5.4: Distribution of the reduced Fox-Wolfram moment R_2 for the charged (left) and neutral mode (right) MC signal (red) and MC u, d, s, c background (blue). SVD1 (SVD2) experiment is shown using a dashed (solid) line.

The fitted PDFs of $\mathcal{F}$ and distributions for MC u, d, s, c background and signal MC samples are illustrated in Fig. 5.6.

In addition, the angle of B momentum in the CM frame with respect to the z-axis, $\cos \theta_B$, flat for continuum background and following $(1 - \cos^2 \theta_B)$ for $\Upsilon(4S)$ decays, is fitted with PDFs

$$\mathcal{P}_{\text{sig}}(\cos \theta_B) = \mathcal{N}(1 - c^{\text{sig}} \cos^2 \theta_B)$$
$$\mathcal{P}_{\text{bkg}}(\cos \theta_B) = \mathcal{N}(1 + c^{\text{bkg}} \cos \theta_B). \tag{5.22}$$

The distributions and fitted PDFs of $\cos \theta_B$ for MC signal and background are shown in Fig. 5.7. We have also checked the correlations of $\cos \theta_B$ to the fit variables. The only significant correlation is about 7% to Δt. With our loose cut on the continuum suppression (see below), this correlation cannot cause problems.

The Fisher weights w_i are determined, and parameters of the PDFs in signal (background) likelihoods $\mathcal{L}_{\text{sig}}$ ($\mathcal{L}_{\text{bkg}}$) are defined and fitted separately for the SVD1 and SVD2 experiment and charged and neutral mode using the MC signal and continuum samples.

The PDFs describing distributions of $\mathcal{F}$ and $\cos \theta_B$ for signal and background are then combined into the signal and background likelihoods $\mathcal{L}_{\text{sig}} = \mathcal{P}_{\text{sig}}(\mathcal{F})\mathcal{P}_{\text{sig}}(\cos \theta_B)$ and $\mathcal{L}_{\text{bkg}} = \mathcal{P}_{\text{bkg}}(\mathcal{F})\mathcal{P}_{\text{bkg}}(\cos \theta_B)$, respectively. In the event selection, these are evaluated for each candidate event, and a likelihood ratio is formed as

$$\mathcal{LR} = \frac{\mathcal{L}_{\text{sig}}}{\mathcal{L}_{\text{bkg}} + \mathcal{L}_{\text{sig}}}. \tag{5.23}$$

Its distribution for MC signal and continuum events is shown in Fig. 5.8. The likelihood ratio conveniently aggregates the likeliness of a signal-like event from the event-shape and topology information in a single variable, on which we can impose a cut. This cut can be optimized in several possible ways, but most of the analysis is done using the lowest reasonable value $\mathcal{LR} > 0.2 = CS_{\text{cut}}$, which brings a very high signal efficiency, while the background is still suppressed significantly. For the charged mode, it selects 96.2 (96.2) % of the signal and rejects 75.2 (69.6)

---

[3]both combinations, (i,j) and (j,i), including $i = j$, are taken.
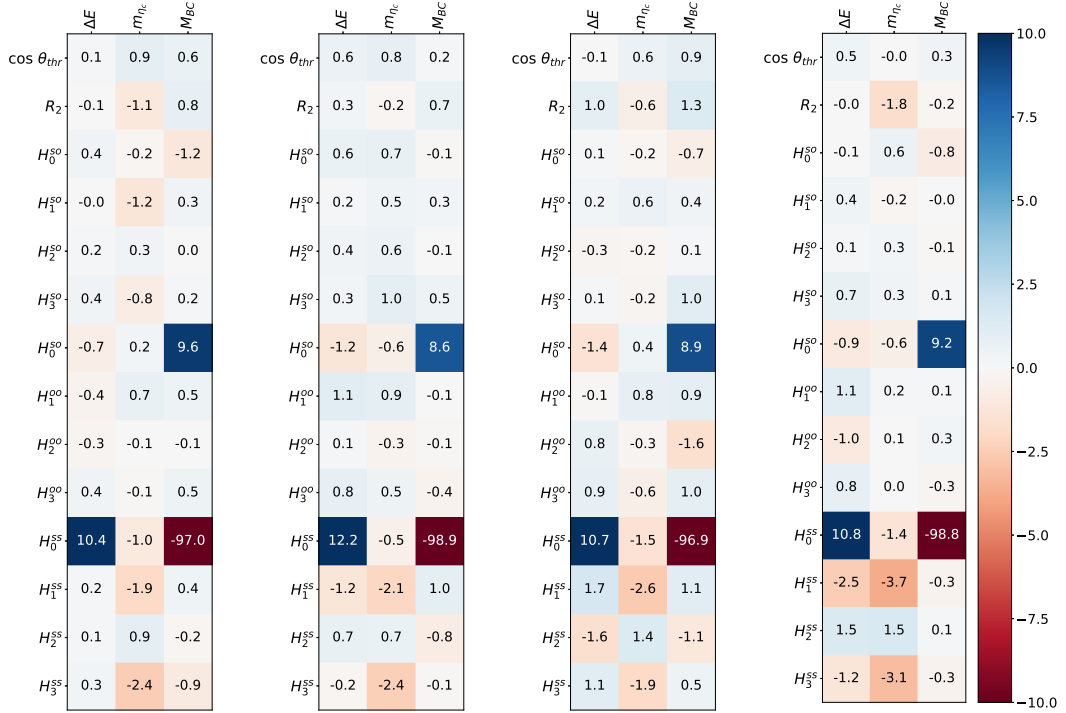
Figure 5.5: Correlations (Pearson correlation coefficient in percent) of event shape variables considered for combination in the Fisher discriminant and the fit variables ΔE, $M(K_S^0 K^+ \pi^-)$ and the M_{bc} variable used for the best-candidate selection. Only variables in the first ten rows are used in the end.

% of continuum background events for SVD1 (SVD2). For the neutral mode it retains 96.3 (96.0) % of signal and rejects 74.9 (70.0) % of continuum background events for SVD1 (SVD2).

Let us briefly explain why such a low cut value is reasonable. If the branching fractions are of interest, one can maximize the statistical significance of the signal over the background using Figure of Merit (FOM), defined as

$$\mathrm{FOM} = \frac{N_{\mathrm{sig}}}{\sqrt{N_{\mathrm{sig}} + N_{\mathrm{bkg}}}}. \tag{5.24}$$

Its dependence on the likelihood ratio cut is shown in Fig. 5.9. The optimal value can be seen to be around 0.9, meaning a very tight cut (only about 50% of the signal remains). The signal and background efficiencies corresponding to different cut values can be read from Fig. 5.10, where also a difference in signal and background efficiencies is shown (black). This is another possible optimization target, giving an optimum cut value of about 0.5.

However, the target of this analysis is the measurement of *CP*-violating parameters, and the result will be statistically limited. All signal events do not contribute with the same 'weight' to the likelihood. The most important events are those with a high probability of correct flavor determination and with large measured lifetimes. As these quantities are generally not correlated with the signal-side kinematics, a more strict continuum suppression cut uniformly re-
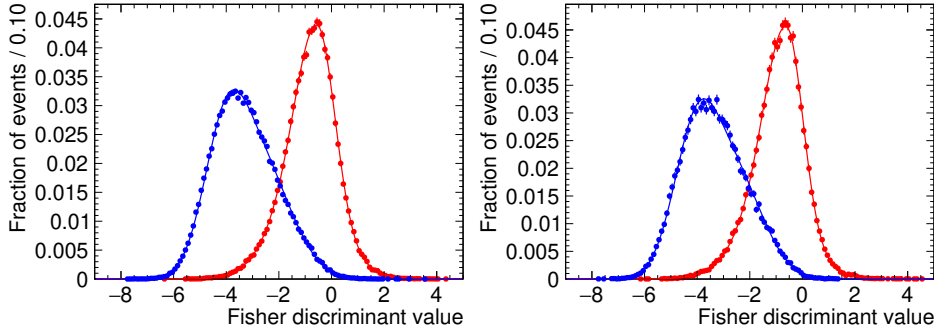
Figure 5.6: Distribution of the Fisher discriminant $\mathcal{F}$ for MC signal (red) and continuum MC (u, d, s, c) background (blue) for the charged control mode (left) and the signal mode (right) using the SVD2 experiment simulation.
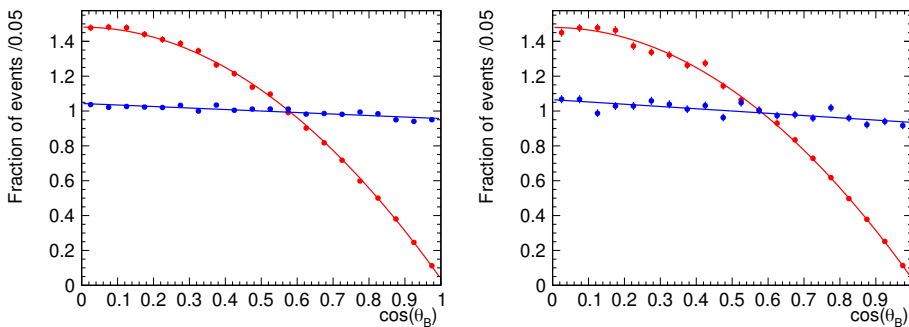
Figure 5.7: Distribution of $\cos\theta_B$ for MC signal (red) and continuum MC (u, d, s, c) background (blue) for the charged control mode (left) and the signal mode (right) using the SVD2 experiment simulation.

moves signal events, even those important ones[4]. Therefore the lowest statistical uncertainty for CP-violation parameters is achieved by maximizing the signal efficiency [40]. The additional continuum background with short lifetimes does not pose a significant problem for the maximum likelihood fit if additional information allows for the separation of signal and background events. For our case, this is achieved by the two fit dimensions in ΔE and M. In principle, one could also model the likelihood ratio distributions for signal and background(s') events and include this information as an additional fit variable. It would, however, make the fit significantly more complicated and slower.

To demonstrate the previous statements, we show here two results from a study using full combined (SVD1 + SVD2) fit in the charged mode. We perform the fit, including the estimation of CP-violating parameters $\mathcal{A}_{CP}$ and $\mathcal{S}_{CP}$ using different values for the $\mathcal{CS}_{\text{cut}}$. The results can be seen in Fig. 5.11.

---

[4]One could consider an r-bin-dependent $\mathcal{CS}_{\text{cut}}$, which would, however, require some additional systematic studies. During the $\mathcal{CS}$ studies for the control mode, we indeed observed that r-bin parametrization of model components is changing. A tighter $\mathcal{CS}_{\text{cut}}$ would result in a limited amount of background available to extract resolution function parameters. Using a different cut value for sideband parameter extraction than for signal to avoid this would present yet another systematic uncertainty, as continuum background suppression variables' correlations with the fit variables are more significant than for signal.
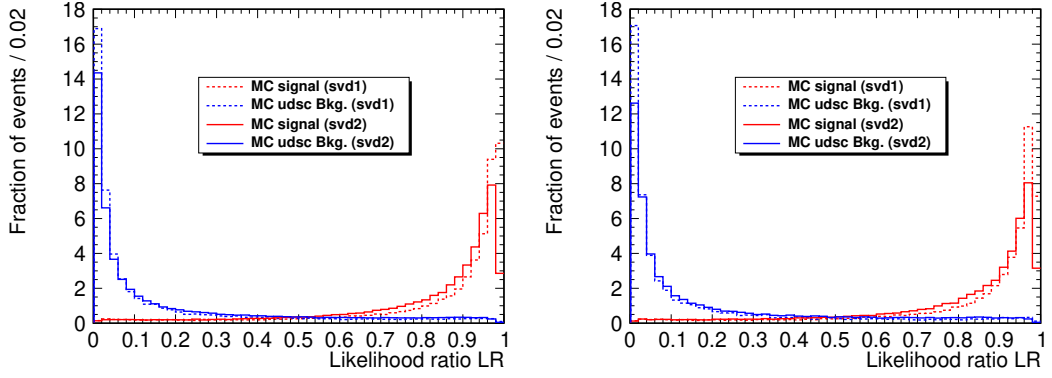
Figure 5.8: Distribution of the likelihood ratio $\mathcal{LR}$ for charged (left) and neutral mode (right) MC signal (red) and continuum MC (u, d, s, c) background (blue). SVD1 experiment is shown using a dashed line, SVD2 using a solid line.
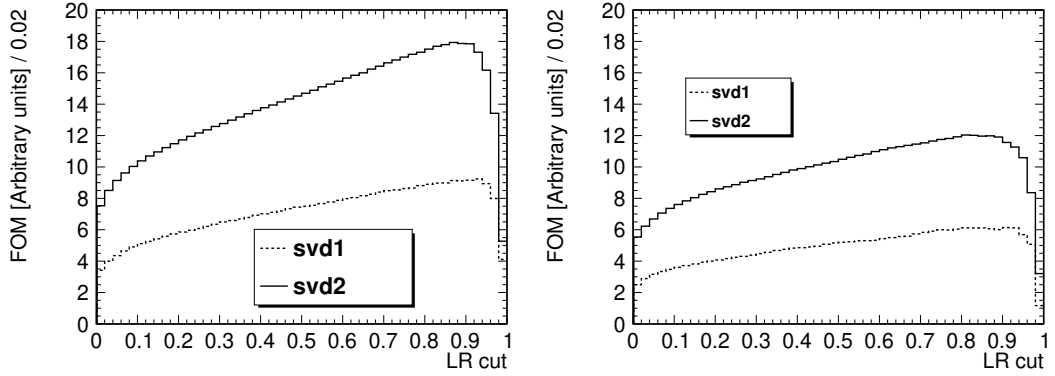
Figure 5.9: Figure of merit (FOM) as function of $\mathcal{CS}_{\mathrm{cut}}$ on the $\mathcal{LR}$ for the charged (left) and neutral mode (right). SVD1 experiment is shown using a dashed line, SVD2 using a solid line.

The estimated values do not show any significant dependence on the cut value. On the other hand, the estimated statistical error (red) on both parameters of interest has a minimum at the cut value of ~ 0.2. This does not agree with any of the previous optimal values, and it is the reason for the low cut value used for this analysis. If an unexpected reason (like systematic errors on CPV parameters generated by the larger background) is found in the analysis, one can consider increasing $\mathcal{CS}_{\mathrm{cut}}$ to about 0.4 without raising the statistical error visibly.

5.3 Peaking Background Vetoes

In this analysis, several backgrounds stemming from generic B decays to charm mesons can be efficiently suppressed by imposing vetoes. These backgrounds have been studied on generic $B\bar{B}$ MC simulation. We form particular combinations of reconstructed final state particles and the respective invariant mass of such combinations. Candidates with combinations corresponding to (production and) decay of D meson in the decay chain are removed. One can see these peaking
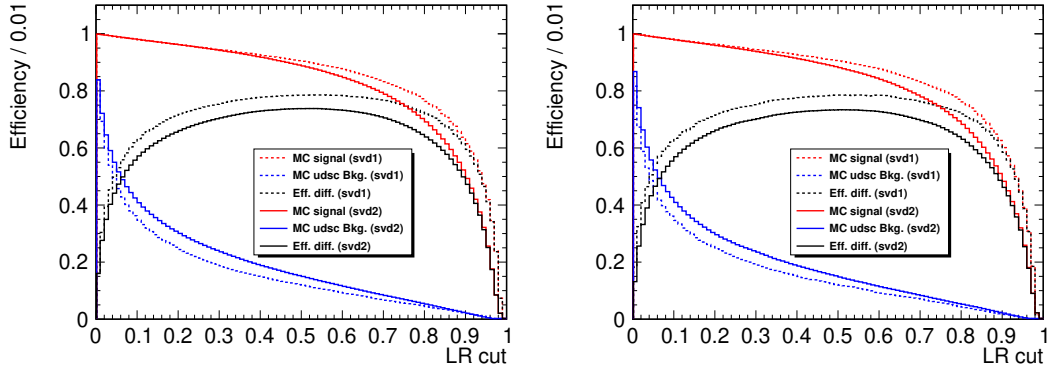
Figure 5.10: Signal (red) and continuum background (blue) efficiency as a function of $\mathcal{CS}_{\text{cut}}$ on the $\mathcal{LR}$ cut for the charged (left) and neutral mode (right). SVD1 experiment is shown using a dashed line, SVD2 using a solid line. The difference between signal and background efficiencies is shown in black.
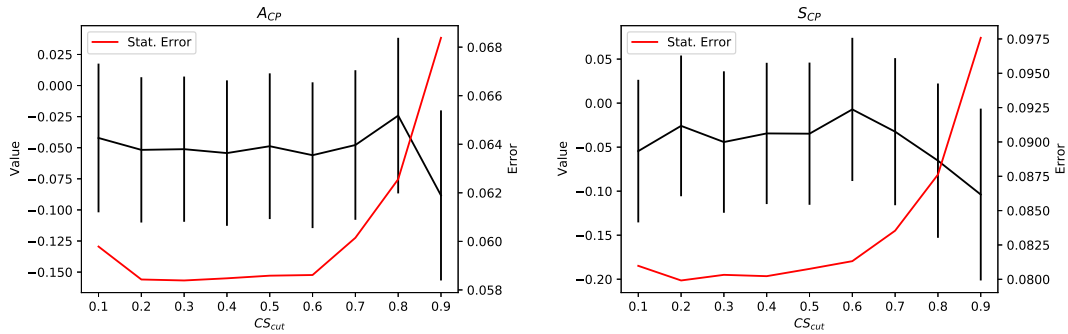
Figure 5.11: Estimated value (black) and statistical error (also shown in red on second axis) for CP-violating parameters $\mathcal{A}_{CP}$ (left) and $\mathcal{S}_{CP}$ (right) in the charged (control) mode from a full combined (SVD1 + SVD2) 4D fit using real Belle data as a function of the likelihood ratio cut.

structures combined in the Dalitz-like plots in Fig. 5.12, where the central position of the peaks at the corresponding decaying particle is denoted. The width of the peaks is estimated using a simple Gaussian fit, which yields a sigma of 6 to 8 MeV/c^2 (depending on the type of D meson), dominated by detector resolution (about 5 MeV/c^2). To exclude these candidates, we reject combinations that fall into a window defined by the central value and width of 15 MeV/c^2, corresponding to about three times the detector resolution. For the neutral mode, we exclude candidates where

- 1.8546 GeV/$c^2 < M(K_S^0 \pi^\pm) < 1.8846$ GeV/c^2, which removes background from $D^\pm \to K_S^0 \pi^\pm$ decays, and

- 1.9535 GeV/$c^2 < M(K_S^0 K^\pm) < 1.9835$ GeV/c^2, for $D_s^\pm \to K_S^0 K^\pm$ decays.

For the charged mode, we veto candidates with

- 1.8499 GeV/$c^2 < M(K^\pm \pi^\mp) < 1.8799$ GeV/c^2, to suppress $D^0 \to K^\pm \pi^\mp$, and
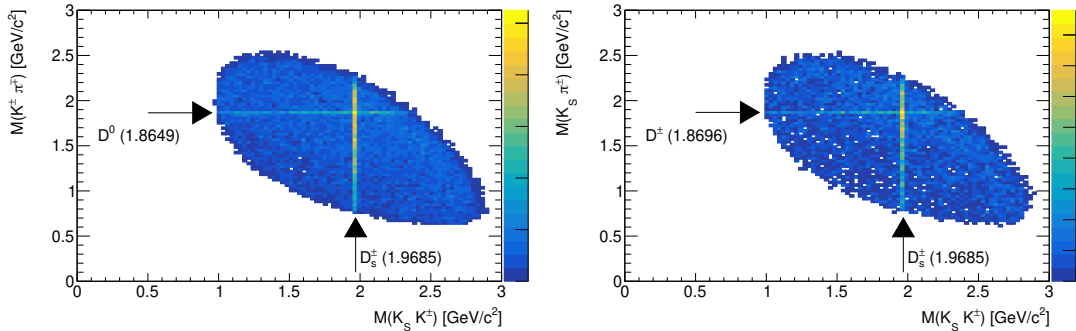
Figure 5.12: Dalitz-like plots (the invariant masses are not squared) for the charged mode (left) and neutral mode (right) with SVD1 and SVD2 experiment combined using generic $B\overline{B}$ MC. The central values of positions of the peaking structures used for vetoes and the corresponding decaying D meson are denoted by arrows. All remaining analysis cuts and continuum suppression is applied for these figures.

- 1.9535 GeV/$c^2 < M(K_S^0 K^\pm) < 1.9835$ GeV/c^2, for $D_s^\pm \to K_S^0 K^\pm$.

Whenever we write $K^\pm$ or K_S^0 above, we check for kaon candidates reconstructed as coming from η_c as well as from the B meson.

These vetoes remove more than 90% of the above peaking backgrounds in ΔE. After the vetoes are applied, the expected signal yield is reduced by 7.4% (6.8%) for neutral (charged) mode. These vetoes reduce the total peaking component in ΔE (dominated now by $b \to u, d, s$ transitions), which is further considered irreducible, to about one-half.

5.4 Remaining Backgrounds

After continuous background suppression and veto cuts, we can estimate the expected background levels by combining the MC data samples. For $q\bar{q}$ continuum background, we studied separately generated six streams of charm and uds MC. The signal is taken directly from ten streams of generic $B\overline{B}$ MC, which is further separated into peaking and non-peaking component in ΔE.

The modes contributing to the peaking component were studied on generic $B\overline{B}$ MC, and their expected yields are listed in Table 5.2 for the charged and in Table 5.3 for the neutral mode, respectively. These backgrounds are composed of $b \to c$ transitions and other charmonia, and thus, some CP-asymmetry can be generated by such backgrounds and systematically bias our measurement. Fortunately, this component is not the dominant contribution to ΔE-peaking backgrounds. The major contribution was investigated on rare-B decays with $b \to u, d, s$. In real data, a direct four-body decay to our final state particles falls under this category. We studied this component on 50 streams of Belle rare-B MC. The individual contributions from MC simulation are combined into histograms in ΔE and $M(K_S^0 K^+ \pi^-)$, which are our 2D fit variables. The peaking components in Fig. 5.13 (charged mode) and Fig. 5.14 (neutral mode) are shown in green colors. For the neutral mode, the level of peaking background is higher than
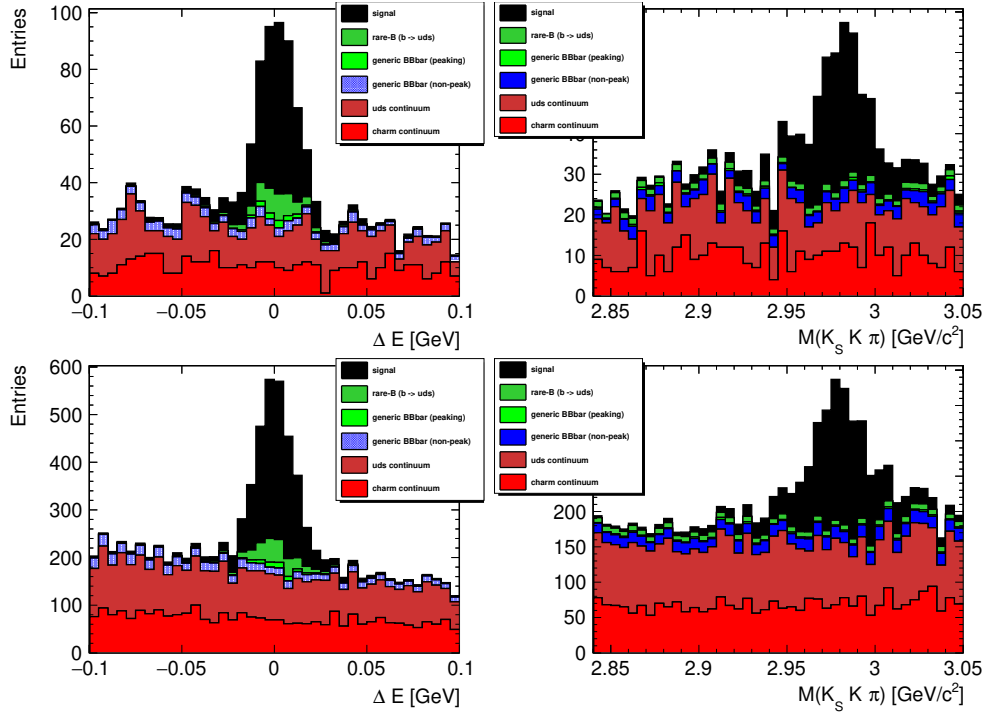
Figure 5.13: Expected signal and background composition from a mix of MC samples (generic, rare-B, and continuum) showing the charged mode for the SVD1 (top) and SVD2 (bottom) experiment.

in the charged mode. All the remaining background from generic $B\overline{B}$ sample non-peaking in ΔE is shown in blue and often considered together with the continuum component in the analysis.

From these rough MC estimates, the fraction of the total peaking background to signal is about 40% for the neutral mode and 21% the the charged mode. The fraction of $b \to c$ background to signal is about 7% for the neutral and 4% for the charged mode. This background is potentially significantly CP-violating in the neutral mode. Its potential effect on the CP-violation measurement will be evaluated in a dedicated systematic study.

While we do not use any of the fractions of background and signal obtained from MC simulations as input for our analysis (except to estimate systematic errors on CPV parameters due to the minor $b \to c$ contribution), we carefully check for data/MC correspondence. As will be shown later, the level of peaking background determined by a direct estimation from the data yields results which are to the first order[5] reproduced by the simulations, taking into account that the branching fractions used for generic and rare-B MC generation can be easily off by 50% and interference effects with non-resonant[6] background are not simulated

---

[5]We will see later that the effective fractions estimated in real data are generally smaller.

[6]By non-resonant background, we mean decays to the same final state particles, but without intermediate formation of the η_c resonance. While many decays have the same final state, their interference with the signal is usually entirely negligible because their phase space is generally different and, in most cases, involves other resonances with mass sufficiently distant from m_{η_c}. Thus such events only fall into the signal window due to mis-reconstruction—by switching kaons from B and from a secondary decay, which just by (a small) chance satisfies all selection
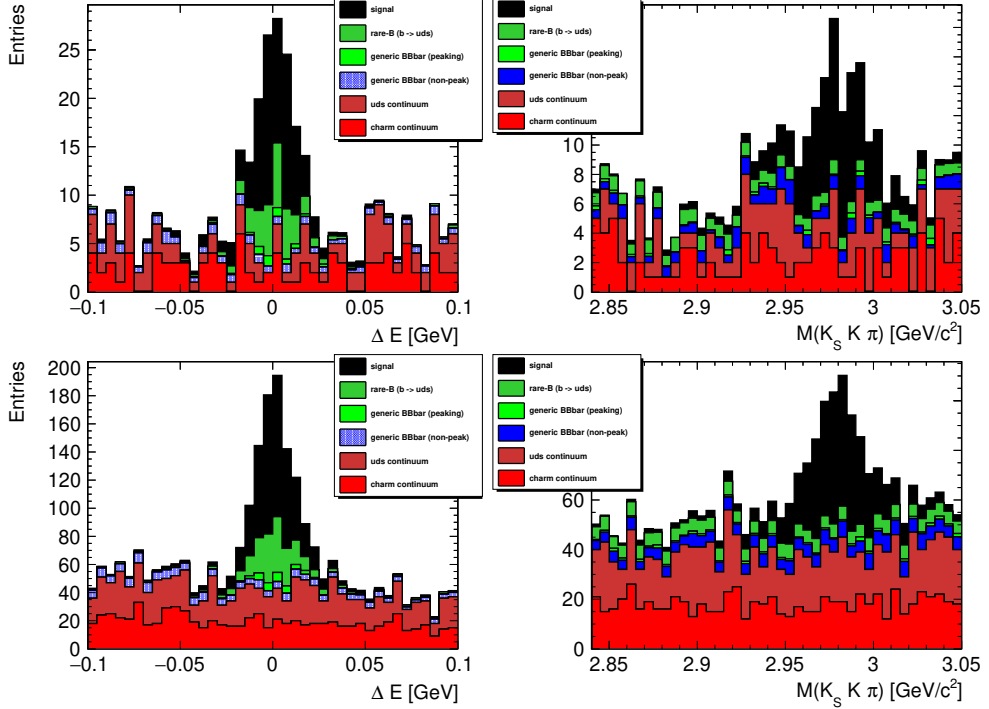
Figure 5.14: Expected signal and background composition from a mix of MC samples (generic, rare-B, and continuum) showing the neutral mode for the SVD1 (top) and SVD2 (bottom) experiment.

at all. This can be for example seen in Tables 5.2 and 5.3, where the correction factors f_{corr} are shown, which update the branching fractions in the decay files used for generic $B\overline{B}$ MC generation to those in more up-to-date PDG from year 2012 [57].

5.5 Summary and Final Detection Efficiency

After all selection criteria summarized in Table 5.4 are applied, the signal efficiency in the signal fit window is computed from the number of reconstructed events N_{rec} and the total number of generated signal MC events N_{gen}. As discussed later, the signal events can be further divided into correctly and incorrectly reconstructed events, where kaons from η_c decay are mismatched with those directly from the B meson. N_{rec} is defined as a sum of those events, which allows us to estimate a proper normalization of the true branching fraction from the total number of observed signal events. The signal efficiency is defined as

$$\epsilon = \frac{N_{\mathrm{rec}}}{N_{\mathrm{gen}}}, \tag{5.25}$$

where $N_{\mathrm{gen}} = 499,995$ for SVD1 experiment and $N_{\mathrm{gen}} = 499,992$ for SVD2 experiment. The difference from exactly half a million events is caused by a rare

criteria. However, a direct four-body decay without peaking structures can have a significant overlap with the η_c decay phase space (one usually assumes a flat non-resonant contribution over $M(K^0_S K^+\pi^-)$).

79

Decay channel	N(SVD1)	N(SVD2)	f_{corr}	N_{corr}(SVD1)	N_{corr}(SVD2)
$B^+ \to \bar{D}^0 K_S^0 \pi^+$	1.1	4.0	1.00	1.1	4.0
$B^+ \to \bar{D}^0 K^{*+}$	1.1	4.0	1.02	1.1	4.1
$B^+ \to \bar{D}^0 K^+$	0.3	2.8	0.41	0.1	1.1
$B^+ \to \bar{D}^0 \pi^+$	0.0	0.6	1.84	0.0	1.1
$B^+ \to J/\psi K^+$	2.6	12.5	1.10	2.9	13.8
$B^+ \to J/\psi K^{*+}$	0.2	0.4	1.26	0.3	0.5
$B^+ \to \eta_c(2S) K^+$	3.3	16.3	0.68	2.2	11.0
$B^+ \to \psi(2S) K^+$	0.4	2.1	0.94	0.4	2.0
$B^+ \to \psi(2S) K^{*+}$	0.0	0.1	0.46	0.0	0.1
$B^+ \to \chi_{c0} K^+$	0.9	4.0	0.45	0.4	1.8
$B^+ \to \chi_{c0} K^{*+}$	2.2	14.8	0.59	1.3	8.7
$B^+ \to \chi_{c1} K^+$	2.6	15.9	0.51	1.3	8.1
$B^+ \to \chi_{c1} K^{*+}$	0.0	0.0	0.0	0.0	0.0
$B^+ \to \chi_{c2} K^{*+}$	0.5	2.1	1.15	0.6	2.4

Table 5.2: Charged mode: Decay modes contributing to generic $B\overline{B}$ MC peaking in ΔE and their expected yields N in SVD1 and SVD2 experiments. A correction factor f_{corr} to adjust branching fractions used for decay files to a more up-to-date PDG values [57] is used to obtain corrected expected yields N_{corr}.

Decay channel	N(SVD1)	N(SVD2)	f_{corr}	N_{corr}(SVD1)	N_{corr}(SVD2)
$B^0 \to D^- K_S^0 \pi^+$	0.4	4.6	1.00	0.4	4.6
$B^0 \to D^- K^*(892)^+$	0.5	2.5	0.96	0.5	2.4
$B^0 \to D^- K^+$	0.4	4.1	1.01	0.4	4.1
$B^0 \to D^- \pi^+$	0.3	0.5	1.00	0.3	0.5
$B^0 \to J/\psi K_S^0$	1.3	8.4	1.10	1.3	8.7
$B^0 \to \eta_c(2S) K_S^0$	2.7	15.2	0.27	0.7	4.1
$B^0 \to \psi(2S) K_S^0$	0.0	0.4	1.00	0.0	0.4
$B^0 \to \chi_{c0} K_S^0$	1.1	5.4	0.28	0.3	1.5
$B^0 \to \chi_{c0} K^*(892)^0$	0.5	2.2	0.78	0.4	1.7
$B^0 \to \chi_{c1} K_S^0$	0.7	5.7	0.72	0.5	4.1
$B^0 \to \chi_{c1} K^*(892)^0$	0.3	2.4	0.54	0.2	1.3
$B^0 \to \chi_{c2} K_S^0$	0.1	0.5	0.20	0.0	0.1
$B^0 \to \chi_{c2} K^*(892)^0$	0.4	1.6	0.48	0.2	0.8

Table 5.3: Neutral mode: Decay modes contributing to generic $B\overline{B}$ MC peaking in ΔE and their expected yields N in SVD1 and SVD2 experiments. A correction factor f_{corr} to adjust branching fractions used for decay files to a more up-to-date PDG values [57] is used to obtain corrected expected yields N_{corr}. All contributions in this table are further considered to be (maximally) CP violating in a dedicated systematic study.

Kaon and pion selection	
Impact parameters	$\lvert dr \rvert < 0.5$ cm & $\lvert dz \rvert < 3$ cm
Kaon PID	$\mathcal{L}_{K/\pi} \geq 0.6$
Pion PID	$\mathcal{L}_{K/\pi} < 0.9$
Electron veto:	$p(e) < 0.95$
K_S^0 selection	
K_S^0 mass window	$482 \text{ MeV}/c^2 < M(\pi^+\pi^-) < 514 \text{ MeV}/c^2$
	`goodKs` cut
Best candidate selection	
M_{bc}-based:	$\min \lvert M_{\mathrm{bc}} - 5.27915 \text{ GeV}/c^2 \rvert$
Continuum suppression cut	
	$\mathcal{CS}_{\mathrm{cut}} > 0.2$
Vetoes	
Neutral mode:	$1.8546 \text{ GeV}/c^2 < M(K_S^0\pi^\pm) < 1.8846 \text{ GeV}/c^2$
	$1.9535 \text{ GeV}/c^2 < M(K_S^0 K^\pm) < 1.9835 \text{ GeV}/c^2$
Charged mode:	$1.8499 \text{ GeV}/c^2 < M(K^\pm\pi^\mp) < 1.8799 \text{ GeV}/c^2$
	$1.9535 \text{ GeV}/c^2 < M(K_S^0 K^\pm) < 1.9835 \text{ GeV}/c^2$
Vertexing & flavour tagging	
No. of SVD hits (≥ 1 track):	≥ 1 in $R - \phi$ plane & ≥ 2 in z-direction
Multi-track vertices:	$h < 50$
	$\sigma_z < 200 \ \mu\text{m}$
Single-track vertices:	$\sigma_z < 500 \ \mu\text{m}$
Signal region	
ΔE fit range:	$-0.1 \text{ GeV} < \Delta E < +0.1 \text{ GeV}$
$M(K_S^0 K^+\pi^-)$ fit range:	$2.84 \text{ GeV}/c^2 < M < 3.05 \text{ GeV}/c^2$
Δt fit range:	$\lvert \Delta t \rvert < 70$ ps
M_{bc} signal region:	$5.271 \text{ GeV}/c^2 < M_{\mathrm{bc}} < 5.290 \text{ GeV}/c^2$
Sideband regions	
M_{bc} sideband region:	$5.200 \text{ GeV}/c^2 < M_{\mathrm{bc}} < 5.265 \text{ GeV}/c^2$
$M(K_S^0 K^+\pi^-)$ sideband region:	$M \in (2.63, 2.84) \text{ GeV}/c^2 \cup (3.15, 3.34) \text{ GeV}/c^2$

Table 5.4: Summary of all selection criteria, background suppression cuts, and definitions of the signal and sideband regions.

	ϵ [%]	Est. N_{sig}
Charged – SVD1	9.60 ± 0.04	341 ± 64
Charged – SVD2	12.93 ± 0.05	1872 ± 350
Neutral – SVD1	5.69 ± 0.03	84 ± 19
Neutral – SVD2	8.66 ± 0.04	522 ± 116

Table 5.5: Signal MC efficiency and estimated signal yield in the signal region. Errors and values for branching fractions are from PDG [34]. Poisson errors are assumed for official Belle $N_{B\overline{B}}$ values for simplicity, as the uncertainties are dominated by the PDG inputs.

case where the simulation fails. The determined signal detection efficiencies are summarized in Table 5.5.

	ϵ [%]	Est. N_{sig}

CHAPTER **6**

Branching Fractions and Background Interference

In this chapter, the 2D data model in the two variables $\Delta E \times M(K_S^0 K^+ \pi^-)$ is built and used to extract signal yields and branching fractions. The last section discusses and estimates the major source of systematic error (interference with non-resonant background).

6.1 2D Data Model

The 2D data model in $\Delta E \times M(K_S^0 K^+ \pi^-)$ allows for some additional independent checks. Namely, a consistency check of signal and background yields and other parameters between the 2D and the full 4D model is an additional confirmation of the correctness of the 4D model. Also, shape parameters for the peaking background will be estimated only using the 2D fit. In this way, possible systematics due to mis-modeling of peaking background in the Δt component will not enter into the determination of these parameters, which will be further fixed in the 4D fit. Finally, a step-by-step approach through the 2D model seems more intuitive and corresponds to the actual analysis flow. To save some space, most fit projections in this chapter are only illustrated for the signal mode and the SVD2 experiment. The differences between SVD1 and SVD2 and control and signal mode are generally negligible, hidden in the statistical noise. Moreover, all relevant projections are shown again when we discuss the final 4D model. Note that except for physics parameters from the PDG, the datasets for the SVD1 and the SVD2 experiments have different parameters, which need to be determined separately in all the individual steps. Only in the final 4D fit the *CP*-violation parameters will be common to both datasets and fitted simultaneously.

6.1.1 Method of (Unbinned) Maximum Likelihood

Given a Probability Density Function (PDF) $\mathcal{P}(x; \alpha)$ for an observable x with a parameter α and a set of N events when we observed $x = x_i$ for event i, the log-likelihood function can be constructed as

$$\ln\mathcal{L}(\alpha) = \sum_{i=1}^{N} \ln \mathcal{P}(x_i; \alpha). \tag{6.1}$$

The goal is to find the best estimates of some parameters of the model (α in this most simple case), which encode the precise form of $\mathcal{P}(x)$, given the observed data, by maximization of the log-likelihood.

	MC value	Data value
	(generator)	(PDG [34])
m_{η_c} [GeV/c^2]	2.9798	2.9839±0.0005
Γ_{η_c} [GeV]	0.0265	0.0320±0.0007

Table 6.1: Parameters for η_c from PDG used for MC generation and fits and for real data.

In addition to the function shape, one could be interested in the best estimation of the total number of observed events [58], i.e., the PDF normalization is now a free parameter. If the number of events follows a Poisson distribution, the extended log-likelihood takes the form

$$\ln\mathcal{L}(\alpha, n) = -n + \sum_{i=1}^{N} \ln n\mathcal{P}(x_i; \alpha), \qquad (6.2)$$

where the additional free parameter n multiplies each 'event probability' $\mathcal{P}(x_i; \alpha)$ and constant factors irrelevant for optimization are omitted.

Just the possibility of evaluating the log-likelihood ($-2\ln\mathcal{L}$ in practice) at a given set of model parameters is enough to find the solution by standard minimization techniques and tools. Such tools evaluate the log-likelihood many times to compute derivatives of this usually multi-dimensional function and search for (ideally the global) minimum. For this purpose, the MINUIT2 package is used.

6.1.2   2D Signal Model and Signal MC

A large simulated data sample of 0.5 million MC signal events (see Sec. 5.5 for details) for each experiment is used to determine most parameters of the signal shape for events that pass the event reconstruction. The MC is generated with somewhat outdated values for mass and width of the η_c resonance. Thus for signal MC fits, the values are fixed to generated ones, while for real data fits, the most up-to-date PDG values are used. The values are summarized in Table 6.1.

In addition, data/MC correction factors discussed later are introduced for generic changes of means and widths of the signal shape. This accounts for additional real detector bias and resolution effects not included in the simulation. Thus these correction factors are only determined from the fit to the control sample and later fixed for the neutral mode.

The signal MC shows a significant correlation of the ΔE and $M(K_S^0 K^+ \pi^-)$ variables. A 2D histogram of the signal MC data sample for SVD2 is shown in Fig. 6.1. From these histograms, the extracted correlation factor ρ is 26 % for charged and 25 % for neutral mode in the SVD1 configuration. For SVD2 configuration, ρ is 24 % for charged and neutral mode. This correlation is naturally expected, as the B-meson momenta (and thus ΔE) is determined using the η_c meson. We attempt to describe this correlation in the signal PDF—the incorporation of the correlation in the correctly reconstructed signal model is discussed in the next Sec. 6.1.3.

Even in the case when the η_c is mis-reconstructed, ΔE might still be determined correctly if all final state particles are correctly assigned to the respective
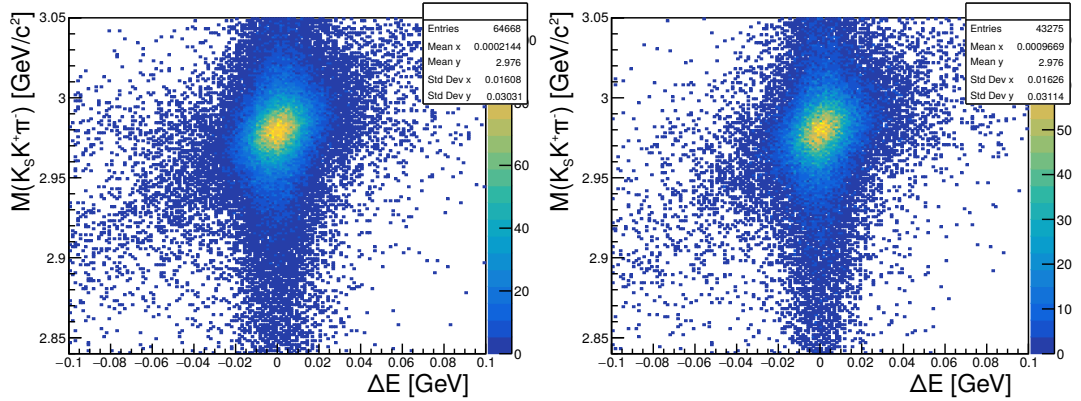
Figure 6.1: 2D signal MC data for SVD2 charged (left) and neutral (right) mode.

B-mesons (on the signal and tag side). The probability of this happening is studied using the MC truth information. Using simulation flags, one can separate the events where both η_c and B-meson are reconstructed correctly. This also includes cases with missed final state radiation photons, misidentified particles (like switched pion and kaon), incorrect assignment of particles after material interaction, and also misidentification of the charge for the case of charged B meson. If under this criteria, only the B-meson is correctly reconstructed, the events are separated, and this component is later called mis-reconstructed signal component. This component and its model are described in detail in Sec. 6.1.4.

6.1.3 Correctly Reconstructed Signal

The PDF for the signal is mainly inspired by the previous work on the $\eta_c \rightarrow p\bar{p}$ decay channel [40], where a single correlation coefficient is introduced to model the correlation of the $\Delta E \times M$ variables for the fully correctly reconstructed signal. Here, this approach seems not sufficient as the correlation factors observed in the MC sample are not well reproduced. Instead, an empirical modification was added, and a second correlation coefficient was introduced. Several statistical checks were performed to select the best way to model the correlation. The final approach was selected mainly for better reproduction of the MC correlation factors and due to a better 2D χ^2 and width of the pull distributions.

In the M variable, the η_c resonance mass spectrum is parametrized as a sum of main and wide components, both in the form of a Voigt function (Breit-Wigner convolved with a Gaussian accounting for detector resolution and bias effects):

$$\tilde{\mathcal{P}}_{\text{sig}}(M) = + (1-f)\mathcal{BW}(M; m_{\eta_c}, \Gamma_{\eta_c}) \otimes G(M; \mu^{\text{Voigt}}, \sigma^{\text{Voigt}}_{\text{main}}) \\ + f\mathcal{BW}(M; m_{\eta_c}, \Gamma_{\eta_c}) \otimes G(M; \mu^{\text{Voigt}}, \sigma^{\text{Voigt}}_{\text{main}} \sigma^{\text{Voigt}}_{\text{tail}}),$$

(6.3)

where $\mathcal{G}(x; \mu, \sigma)$ is a Gaussian with mean μ and width σ, f is the fraction of the wide component, and μ^{Voigt} represents the shift with respect to the central value of the Breit-Wigner, same for the main and wide component. For MC, this shift was found to be compatible with zero, with the only exception being SVD1 charged mode, where the fitter, described later, found a value of $(0.23\pm0.10)\text{MeV}/c^2$.

85

This tiny correction can be neglected. No data/MC correction factor is introduced here, as well as for the detector smearing $\sigma_{\text{main}}^{\text{Voigt}}$ for the main component. The width of the wide component is parameterized with respect to the main component by an additional multiplication factor $\sigma_{\text{tail}}^{\text{Voigt}}$.

To model the correlation, the PDF in ΔE is a conditional probability $\tilde{\mathcal{P}}_{\text{sig}}(\Delta E | M)$, where some of the parameters are themselves functions of $M(K_S^0 K^+ \pi^-)$. The PDF is expressed as a sum of three Gaussians:

$$
\begin{aligned}
\tilde{\mathcal{P}}_{\text{sig}}(\Delta E | M) = &+ (1 - f_1 - f_2) \cdot \mathcal{G}(\Delta E; \tilde{\mu}_{\text{main}} = \mu_{\text{main}} + \mu_{\text{main}}^{\text{CF}} + k_1^{\text{corr}} \mu(M), \\
&\qquad\qquad \tilde{\sigma}_{\text{main}} = \sigma_{\text{main}} \sigma_{\text{main}}^{\text{CF}}) \\
&+ f_1 \cdot \mathcal{G}(\Delta E; \tilde{\mu}_{\text{main}} + \mu_{\text{tail}}^1 + k_2^{\text{corr}} \mu(M), \tilde{\sigma}_{\text{main}} \sigma_{\text{tail}}^1) \\
&+ f_2 \cdot \mathcal{G}(\Delta E; \tilde{\mu}_{\text{main}} + \mu_{\text{tail}}^2, \tilde{\sigma}_{\text{main}} \sigma_{\text{tail}}^2),
\end{aligned}
\tag{6.4}
$$

where f_1, f_2 stand for fractions of the wider components, whose means are expressed as shifts μ_{tail}^1 and μ_{tail}^2 with respect to the total mean of the main Gaussian $\tilde{\mu}_{\text{main}}$, which includes a general shift μ_{main} and data/MC correction factor $\mu_{\text{main}}^{\text{CF}}$, set to zero for MC. Similarly the widths are parametrized by multiplication factors σ_{tail}^1 and σ_{tail}^2, with respect to the total $\tilde{\sigma}_{\text{main}}$. The data/MC correction factor $\sigma_{\text{main}}^{\text{CF}}$ to σ_{main} is again set to one for MC. The correlation is introduced by inserting a linear dependence of the common mean, and the mean of the second Gaussian (as an additional correction) on the M variable:

$$
\mu(M) = M - (m_{\eta_c} + \mu^{\text{Voigt}}),
\tag{6.5}
$$

where the slope is controlled by the parameters k_1^{corr} and k_2^{corr}. The obtained correctly reconstructed signal model parameters are extracted using the following 2D PDF:

$$
\tilde{\mathcal{P}}_{\text{sig}}(\Delta E, M) = \tilde{\mathcal{P}}_{\text{sig}}(\Delta E | M) \tilde{\mathcal{P}}_{\text{sig}}(M).
\tag{6.6}
$$

The fit is repeated for the full signal model discussed below. Fit projections of the 2D model and MC data for the correctly reconstructed signal are illustrated in Fig. 6.2. The systematic pattern in the ΔE projection indicates that the description of the correlation is not perfect. This will be even more evident in Fig. 6.5 showing two-dimensional pulls. Nevertheless, the statistics of real data is almost two orders of magnitude smaller, so the at most $\pm 4\sigma$ pulls in the projection are not relevant for our purposes. Moreover, this deficiency of the model can be evaluated on ensemble tests (with signal events sampled from GEANT3 MC simulation) as a possible fit bias and taken into account. In fact, these signal model deficiencies are fully negligible with respect to other sources, like the signal line shape in $M(K_S^0 K^+ \pi^-)$ modified by interference with the non-resonant background.

The fit window was also extended to the full region, including $M(K_S^0 K^+ \pi^-)$ sidebands to check if the shape parameters changed. The only parameter which changes significantly is the main correlation parameter k_1^{corr}, which gets much smaller in the wider fit window. This is understood to be caused by the empirical description of the signal correlated shape. No further refinements were done to improve this situation as the signal model is good enough at the statistical level for the studied channel. One, however, must be careful when extrapolating signal parameters outside of the signal fit window, as the parameters describing the

Figure 6.2: Projections of the correctly reconstructed signal component 2D data model and MC signal data for the neutral (signal) mode and SVD2 experiment.

correlation might not lead to optimal modeling of the signal shape outside of the signal window. This is later discussed in Sec. 6.1.9.

6.1.4 Mis-reconstructed Signal

The nature of the decay allows swapping of the kaons from B and η_c decay in the reconstruction. If this happens, the signal does not peak in $M(K_S^0 K^+ \pi^-)$ anymore, as the η_c is mis-reconstructed. But for the B meson, the sum of the four-momenta of all the final particles on the signal side will still be correct and result in a peaking component in ΔE, which, however, does not anymore exhibit any correlation with $M(K_S^0 K^+ \pi^-)$, in contrast to the correctly reconstructed signal component. This component of the signal and parameters specific to it are further denoted by a subscript $_{\mathrm{mis}}$.

In the $M(K_S^0 K^+ \pi^-)$ variable, a single-parameter PDF has been found sufficient to describe the mis-reconstructed signal component:

$$\mathcal{P}_{\mathrm{mis}}(M) = \mathcal{N}(1 + c_{\mathrm{mis}}^1 \mathcal{C}_1(M)), \tag{6.7}$$

where $\mathcal{C}_1$ is a first-order Chebyshev polynomial (linear function) and $\mathcal{N}$ is a proper normalization factor. In the ΔE variable, the PDF is almost the same as for the correctly reconstructed signal. The parameters are shared, too, but the correlation is not introduced:

$$
\begin{aligned}
\mathcal{P}_{\mathrm{mis}}(\Delta E) = {} & + (1 - f_1 - f_2)\mathcal{G}(\Delta E; \tilde{\mu}_{\mathrm{main}} = \mu_{\mathrm{main}} + \mu_{\mathrm{main}}^{\mathrm{CF}}, \tilde{\sigma}_{\mathrm{main}} = \sigma_{\mathrm{main}}\sigma_{\mathrm{main}}^{\mathrm{CF}}) \\
& + f_1\mathcal{G}(\Delta E; \tilde{\mu}_{\mathrm{main}} + \mu_{\mathrm{tail}}^1, \tilde{\sigma}_{\mathrm{main}}\sigma_{\mathrm{tail}}^1) \\
& + f_2\mathcal{G}(\Delta E; \tilde{\mu}_{\mathrm{main}} + \mu_{\mathrm{tail}}^2, \tilde{\sigma}_{\mathrm{main}}\sigma_{\mathrm{tail}}^2).
\end{aligned}
\tag{6.8}
$$

Therefore only one parameter (c_{mis}^1) was determined from the fit to the following 2D PDF:

$$\mathcal{P}_{\mathrm{mis}}(\Delta E, M) = \mathcal{P}_{\mathrm{mis}}(\Delta E)\mathcal{P}_{\mathrm{mis}}(M). \tag{6.9}$$

Other model parameters are identical to the correctly reconstructed signal determined in the previous section. The fit and MC data projections are illustrated in Figure 6.3.
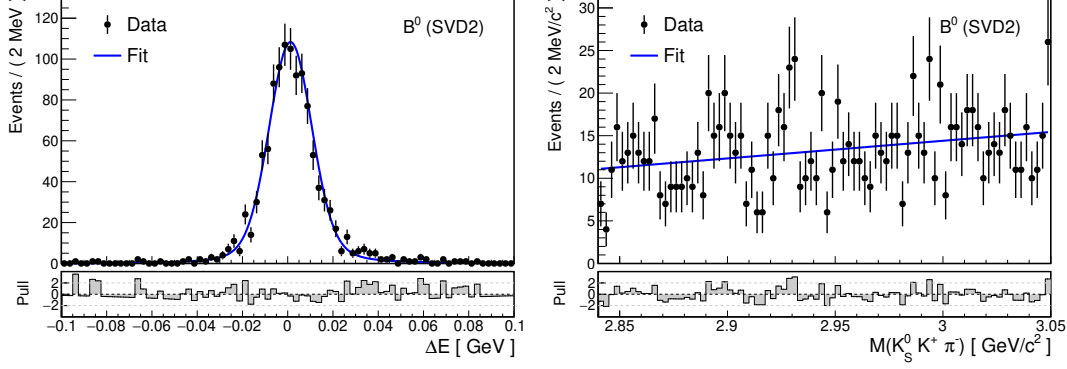
Figure 6.3: Projections of the mis-reconstructed signal component 2D data model and MC signal for the neutral (signal) mode and SVD2 experiment.

6.1.5 Total 2D Signal Model

The total PDF to extract the signal parameter reads

$$\mathcal{P}_{\text{sig}}(\Delta E, M) = (1 - f_{\text{mis}})\tilde{\mathcal{P}}_{\text{sig}}(\Delta E|M)\tilde{\mathcal{P}}_{\text{sig}}(M) + f_{\text{mis}}\mathcal{P}_{\text{mis}}(\Delta E)\mathcal{P}_{\text{mis}}(M), \quad (6.10)$$

where f_{mis} is the fraction of the mis-reconstructed signal. This fraction is higher for the neutral mode due to a larger number of final tracks to be reconstructed (with four charged pion tracks instead of two charged kaon tracks) and thus a larger probability of an incorrect assignment of the daughter to mother particles. This is not a free fit parameter but is determined directly from the MC truth information. The linear slope for the mis-reconstructed signal is fixed to a value determined by the fit to the mis-reconstructed signal in Sec. 6.1.4. This parameter cannot be extracted from the full fit (as well as f_{mis}) because f_{mis} is very small (1–3 %), and the linear shape compensates with the wider Voigt component.

The signal parameters determined from a fit to the total signal are statistically consistent with the results for the correctly reconstructed signal. The largest deviations are in the parameter k_1^{corr}, for the reason of sub-optimal correlation modeling, with the only significant difference for SVD2 in the neutral mode (still within 2σ). All final parameters of the total signal model can be found in the summary Tables A.2 and A.3 of the 2D model.

The fit projections are illustrated in Figure 6.4. In addition, 2D histograms comparing the signal MC and the data model are shown in Figure 6.5. The correlation factors are computed for both signal MC and the fit model and are well reproduced by the full signal model.

6.1.6 Continuum Background

The continuum (without any peaking structures in $\Delta E \times M$) component is mostly composed of $e^+e^- \to q\bar{q}$ processes, where $q = u, d, s, c$ (left-over after continuum suppression), which can be studied using Belle M_{bc} sideband data. In addition, generic $B\overline{B}$ MC can be used to study contributions from $e^+e^- \to b\bar{b}$, where due to mis-reconstruction, usually particles from signal and tag-side B meson are mixed up. To do so, the peaking component (composed of higher $c\bar{c}$ resonances)
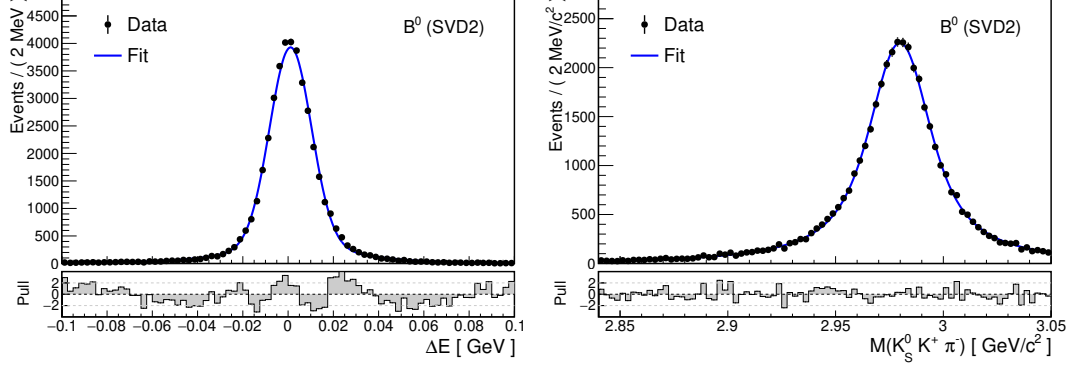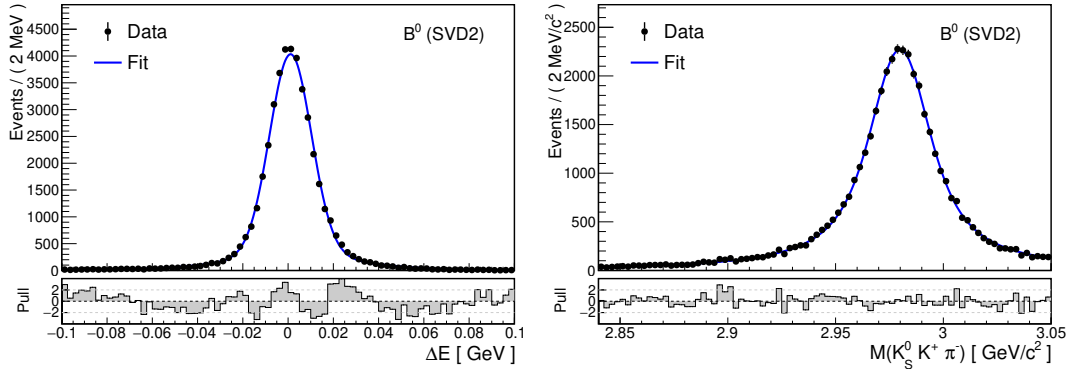
Figure 6.4: Projections of the total signal 2D data model and MC signal data for the neutral (signal) mode and SVD2 experiment.

has to be separated first by removing channels based on MC truth information one by one. The resulting, peaking-free continuum part is fitted, too, but mostly for qualitative evaluation. As will be explained later, the parameters of the continuum background will need to be released again after the fit to the M_{bc} sideband data. Thus the parameters obtained in this section are only relevant for later cross-checks and as starting values for later fits.

The PDF used for the continuum background is simply a Chebyshev polynomial in each variable:

$$
\begin{aligned}
\mathcal{P}_{\mathrm{cb}}(\Delta E) &= \mathcal{N}(1 + c_1^{\Delta E}\mathcal{C}_1(\Delta E)), \\
\mathcal{P}_{\mathrm{bkg}}(M) &= \mathcal{N}(1 + c_1^{M}\mathcal{C}_1(M)),
\end{aligned}
\tag{6.11}
$$

where the subscript $_{\mathrm{bkg}}$ refers to a fundamental assumption of the full data model, which is supported by all available MC studies: the shape of the sum of all backgrounds in M can be described by $\mathcal{P}_{\mathrm{bkg}}(M)$ with some c_1^{M}, even that coming from (ΔE) peaking background. There is no reason to assume that the slopes of individual components are the same (which is not even supported by the extracted fit parameters from samples shown below). Although the slopes of individual continuum (or ΔE-peaking) background components in M are unknown and are not modeled (only their fractions can be extracted), their arbitrary composition can again only have a linear shape, which must be released when the fractions can change (e.g., when opening/changing the fit region to sidebands and vice-versa). As a further dependency of the slope on the ΔE variable can invalidate such an assumption, the (linear) shape in ΔE must also be a floating variable in all fits.

The total 2D PDF for the continuum and non-peaking $B\overline{B}$ background[1] is

$$
\mathcal{P}_{\mathrm{cb}}(\Delta E, M) = \mathcal{P}_{\mathrm{cb}}(\Delta E)\mathcal{P}_{\mathrm{bkg}}(M).
\tag{6.12}
$$

The fit projections and the Belle M_{bc} sideband data are illustrated in Fig. 6.6. The projections for the non-peaking part of the generic $B\overline{B}$ MC are illustrated in Fig. 6.7. Note that all extracted parameters describing the linear background shape will be re-estimated in the total 2D data fit.

---

[1]We often refer to this component as continuum only, with meaning of the continuous shape of the background in the $\Delta E \times M$ variables, without any peaking structures.
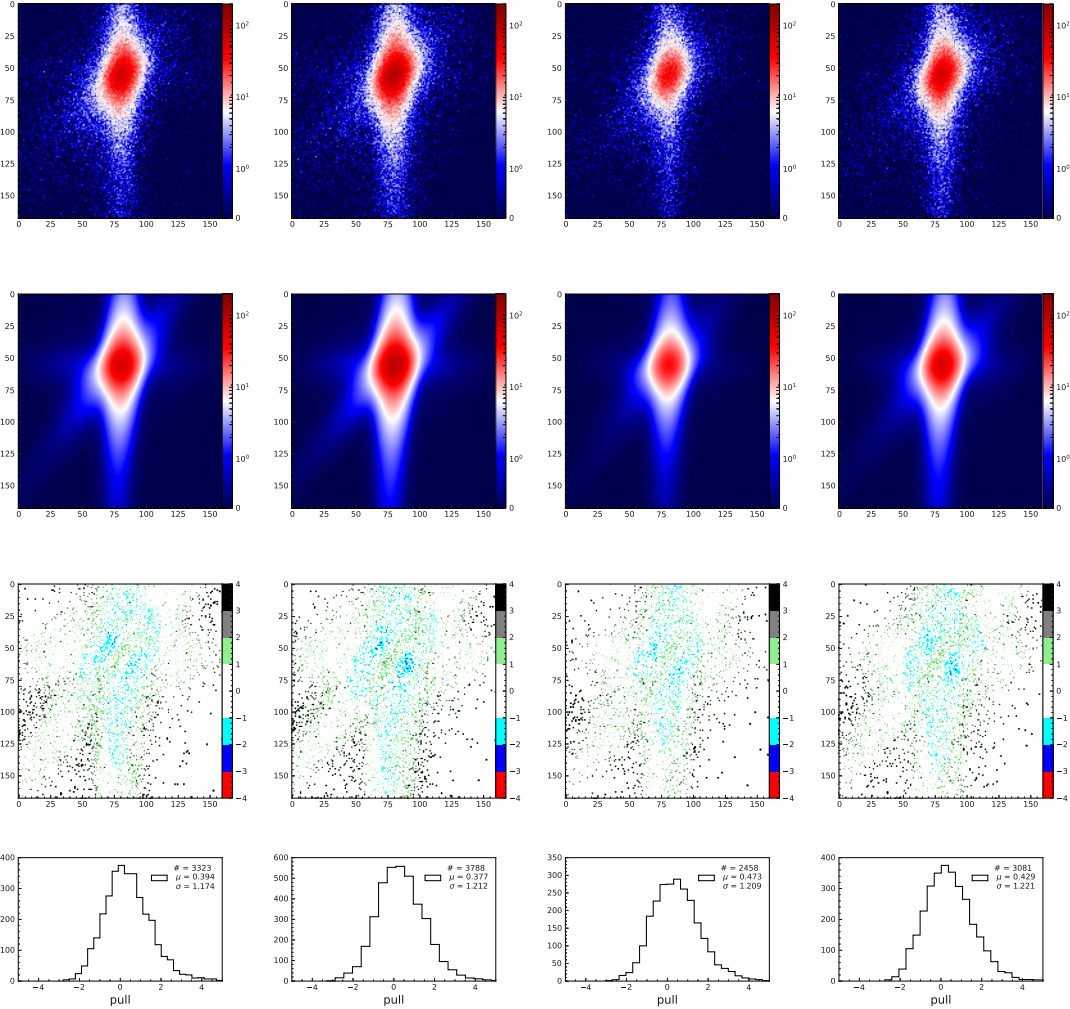
Figure 6.5: Comparison of MC signal data (first row) and fitted signal model (second row). All (logarithmic) scales are set the same (maximum at 200). For more qualitative comparison, the pull plots are shown (third row). Here, a different definition for a pull is used, $\text{pull}_i = (N_i^{\text{data}} - N_i^{\text{fit}})/\sqrt{N_i^{\text{fit}}}$, to fill all bins and avoid division by zero. The pull distributions of bins where $N_i^{\text{data}} > 2$ are shown in 1D histograms (fourth row) with bin count, mean, and standard deviation in the inset. From left to right, distributions for charged (SVD1, SVD2) and neutral (SVD1, SVD2) modes are shown.

The data has been found to be consistent with the simple linear shape in all cases, except for the SVD2 control mode. Here, with the most events, a second-order Chebyshev polynomial would be slightly preferred statistically. However, it has been decided to keep the model identical to the one used for signal extraction. The effect on the control mode is entirely negligible when compared to other systematic uncertainties[2].
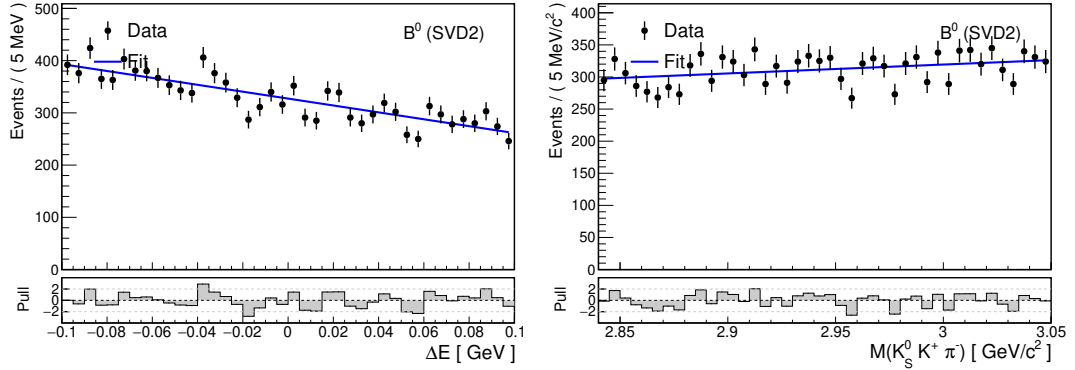
Figure 6.6: Projections of the continuum background 2D data model and Belle M_{bc} sideband data for the neutral (signal) mode and SVD2 experiment.
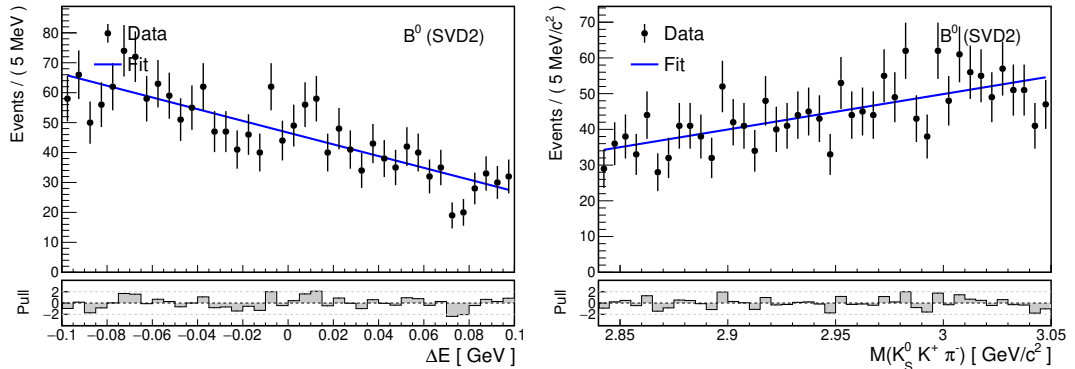
Figure 6.7: Projections of the continuum background 2D data model and generic $B\overline{B}$ MC (non-peaking in ΔE) for the neutral (signal) mode and SVD2 experiment.

6.1.7 Peaking Background

The background component peaking in ΔE is studied on two MC samples: generic $B\overline{B}$ decays ($b \to c$) with ten times larger dataset than experiment luminosity, and rare-B ($b \to uds$) generated with 50 times larger dataset than real Belle data. The rare-B data sample is used only for some qualitative studies. Also, the branching fractions of the individual B decays might not be up-to-date, and some

---

[2]Which we (mostly) do not even evaluate for the control mode as the fit is much slower to perform a full-scale systematic study as done for our signal (with less floating parameters and about a third of the data in the likelihood function).

specific decay channels could be missing or not be modeled correctly in these data samples. Uncertainties in the correctness of the composition of these MC data samples can be overcome if the parameters of the (total) peaking background can be estimated from a fit to the data. According to qualitative studies below, a single Gaussian is sufficient to describe the shape of the peaking background in both MC data samples:

$$\mathcal{P}_{\text{pb}}(\Delta E) = \mathcal{G}(\Delta E; \mu_{\text{pb}}^{\Delta E}, \sigma_{\text{pb}}^{\Delta E}), \tag{6.13}$$

while the shape in $M(K_S^0 K^+ \pi^-)$ is found to be consistent with a linear or even a constant function. Therefore the total PDF used to extract the peaking background shape reads

$$\mathcal{P}_{\text{bkg}}(\Delta E, M) = (1 - f_{\text{pb}})\mathcal{P}_{\text{cb}}(\Delta E)\mathcal{P}_{\text{bkg}}(M) + f_{\text{pb}}\mathcal{P}_{\text{pb}}(\Delta E)\mathcal{P}_{\text{bkg}}(M), \tag{6.14}$$

where f_{pb} is the fraction of the peaking background to the total background. The rest is to accommodate for the possible remaining small non-peaking component. In fact, for further progress, only the parameters $\mu_{\text{pb}}^{\Delta E}$ and $\sigma_{\text{pb}}^{\Delta E}$ are of interest as possible starting values for the data fit or for comparison of the simulated MC background parameters to real data.

The means of the Gaussian shapes are ± 3 MeV around zero. The widths are roughly 10 MeV. The obtained parameters are statistically consistent among the generic $B\overline{B}$ and rare-B samples. The obtained values are summarized later in Sec. 6.1.9 and Table 6.3, where we compare them to the data-driven estimates. The fit projections for generic $B\overline{B}$ MC are illustrated in Fig. 6.8. For the rare-B MC, the projections are illustrated in Fig. 6.9. The discrepancy in the pulls is acceptable, as here, 50 times more events than expected in reality are simulated and a simple shape parametrization is preferred.
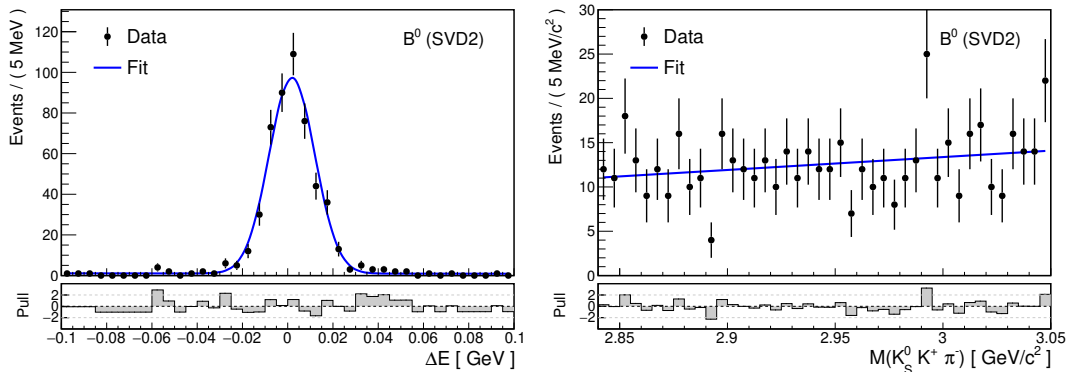
Figure 6.8: Projections of the peaking background 2D data model and generic $B\overline{B}$ MC (peaking in ΔE) for the neutral (signal) mode and SVD2 experiment.

6.1.8 Total 2D PDF and Fit Strategy

The full 2D PDF is a sum of the previously described individual signal (correctly and mis-reconstructed) and background components (peaking and continuum):

$$\mathcal{P}(\Delta E, M) = (1 - f_{\text{bkg}})\mathcal{P}_{\text{sig}}(\Delta E, M) + f_{\text{bkg}}\mathcal{P}_{\text{bkg}}(\Delta E, M), \tag{6.15}$$
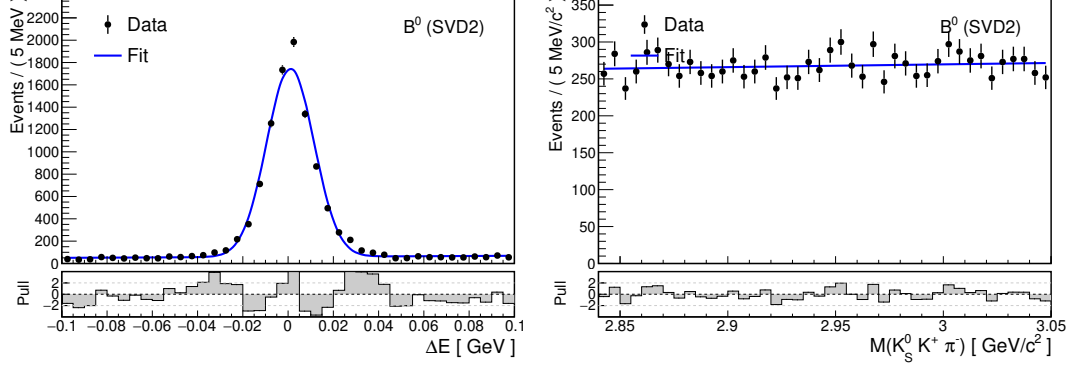
Figure 6.9: Projections of the peaking background 2D data model and rare-B MC for the neutral (signal) mode and SVD2 experiment. The non-peaking component in ΔE is not separated.

where $\mathcal{P}_{\text{sig}}$ is defined in Eq. 6.10, $\mathcal{P}_{\text{bkg}}$ in Eq. 6.14, and f_{bkg} is the total background fraction.

The fit must first be performed in the charged mode to extract data/MC correction factors for the signal shape. This prevents us from performing a fit to the peaking background shape in a simultaneous fit. Instead, the full 2D fit will be first performed with fixed signal shape (with correction factors set to MC default values) and expected yield (from PDG) in the $M(K_S^0 K^+ \pi^-)$ sideband, where the amount of peaking background is large enough to extract mean and width of its Gaussian PDF.

The full 2D fit is performed using the following extended log-likelihood:

$$\ln \mathcal{L} = -(N_{\text{sig}} + N_{\text{bkg}}) + \sum \ln(N_{\text{sig}} + N_{\text{bkg}})\mathcal{P}(\Delta E, M), \qquad (6.16)$$

where the background fraction from Eq. 6.15 is now determined from the additional parameters N_{sig} for the signal and N_{bkg} for the background yield as

$$f_{\text{bkg}} = \frac{N_{\text{bkg}}}{N_{\text{sig}} + N_{\text{bkg}}} \qquad (6.17)$$

and the peaking background yield is now parametrized using a fraction of peaking background to signal $f_{\frac{\text{pb}}{\text{sig}}}$, such that the fraction of peaking background to total background from Eq. 6.14 becomes

$$f_{\text{pb}} = f_{\frac{\text{pb}}{\text{sig}}} \frac{N_{\text{sig}}}{N_{\text{bkg}}}. \qquad (6.18)$$

6.1.9 Extraction of Peaking Background Shape

The mean and width of the Gaussian shape of the peaking background component are estimated by changing the fit window to the $M(K_S^0 K^+ \pi^-)$ sideband. This enlarges the peaking background yield and significantly suppresses the signal. However, the η_c resonance is quite wide, and a non-negligible amount of signal leaks to the sideband. Thus, the signal is first re-fitted in the sideband.

	ϵ^{SB} [%]	Est. N_{sig}(SB)
Charged – SVD1	0.76 ± 0.01	27 ± 5
Charged – SVD2	1.06 ± 0.01	154 ± 30
Neutral – SVD1	0.83 ± 0.01	12 ± 3
Neutral – SVD2	1.25 ± 0.02	76 ± 17

Table 6.2: Signal MC efficiency and estimated signal yield in the $M(K_S^0 K^+\pi^-)$ sideband. Errors and values for branching fractions are from PDG. Poisson uncertainties are assumed for official Belle $N_{B\overline{B}}$ values.

	Charged mode		Neutral mode	
	SVD 1	SVD 2	SVD 1	SVD 2
	Peaking background Gaussian mean $\mu_{\mathrm{PB}}^{\Delta E}$ [MeV]			
generic BB	1.02 ± 0.88	1.30 ± 0.45	2.74 ± 1.33	1.89 ± 0.48
rareB	1.11 ± 0.25	0.56 ± 0.11	1.26 ± 0.31	1.00 ± 0.12
M SB	-2.87 ± 2.11	-0.05 ± 2.04	1.05 ± 3.85	-1.78 ± 1.47
	Peaking background Gaussian width $\sigma_{\mathrm{PB}}^{\Delta E}$ [MeV]			
generic BB	9.86 ± 0.73	11.32 ± 0.42	10.22 ± 1.20	10.03 ± 0.42
rareB	10.43 ± 0.23	10.40 ± 0.10	10.92 ± 0.27	10.40 ± 0.11
M SB	9.99 ± 3.13	11.56 ± 2.57	17.95 ± 3.84	11.35 ± 1.47

Table 6.3: Summary of the extracted peaking background parameters from the generic $B\overline{B}$ MC, rare-B MC, and real data using $M(K_S^0 K^+\pi^-)$ sideband fit.

To extrapolate the signal outside of the signal window, one has to be careful about fractions of components in the model and coefficients of Chebyshev polynomials. These parameters specifically change when the normalization range of the PDFs is changed. However, a direct extrapolation of the signal shape to $M(K_S^0 K^+\pi^-)$ sideband does not yield a good quality fit. This is understood to result from the previously discussed sub-optimal modeling of the signal's correlated nature. Also, the mis-reconstructed signal component is now dominant, especially in the right $M(K_S^0 K^+\pi^-)$ sideband. Therefore the fraction of mis-reconstructed signal is computed again from MC truth, and the full signal fit is repeated with only the Chebyshev coefficient c_{mis}^1 and the main correlation parameter k_1^{corr} floating. Releasing k_1^{corr} was found to be sufficient for a good fit.

The signal yield in the $M(K_S^0 K^+\pi^-)$ sideband can be estimated from the signal efficiency in the sideband ϵ^{SB}, the number of $B\overline{B}$ meson pairs $N_{B\overline{B}}$ and the branching fractions for $B^0 \to \eta_c K_S^0$ (respective $B^+ \to \eta_c K^+$) and $\eta_c \to K_S^0 K^\pm \pi^\mp$ as

$$N_{\mathrm{sig}}(\mathrm{SB}) = \epsilon^{\mathrm{SB}} \times N_{B\overline{B}} \times \mathcal{B}(B^0 \to \eta_c K_S^0 \text{ or } B^+ \to \eta_c K^+) \times \mathcal{B}(\eta_c \to K_S^0 K^\pm \pi^\mp).$$

For this estimation, the Belle efficiency corrections at the 1–2% level are not considered, as the final error, mainly coming from uncertainties of the PDG branching ratios, is around 20%. The values for the sideband signal efficiency and the final estimated signal yields in the sideband are listed in Table 6.2.

The fitted peaking background parameters with N_{sig} fixed to the estimated

values are compared to the previously obtained MC estimates in Table 6.3. The means of the Gaussians are mostly consistent with zero. The widths are mostly similar to those predicted by the generic and rare-B MC samples, about 10 MeV. Only the value of (18 ± 4)MeV for the SVD1 experiment and the neutral mode is an exception, but likely only a statistical fluctuation. Due to the consistency of all the remaining data/MC results, we have decided to use the MC estimate for the width in the SVD1 neutral mode. The data-driven value is included in systematic variations of the final data model parameters.
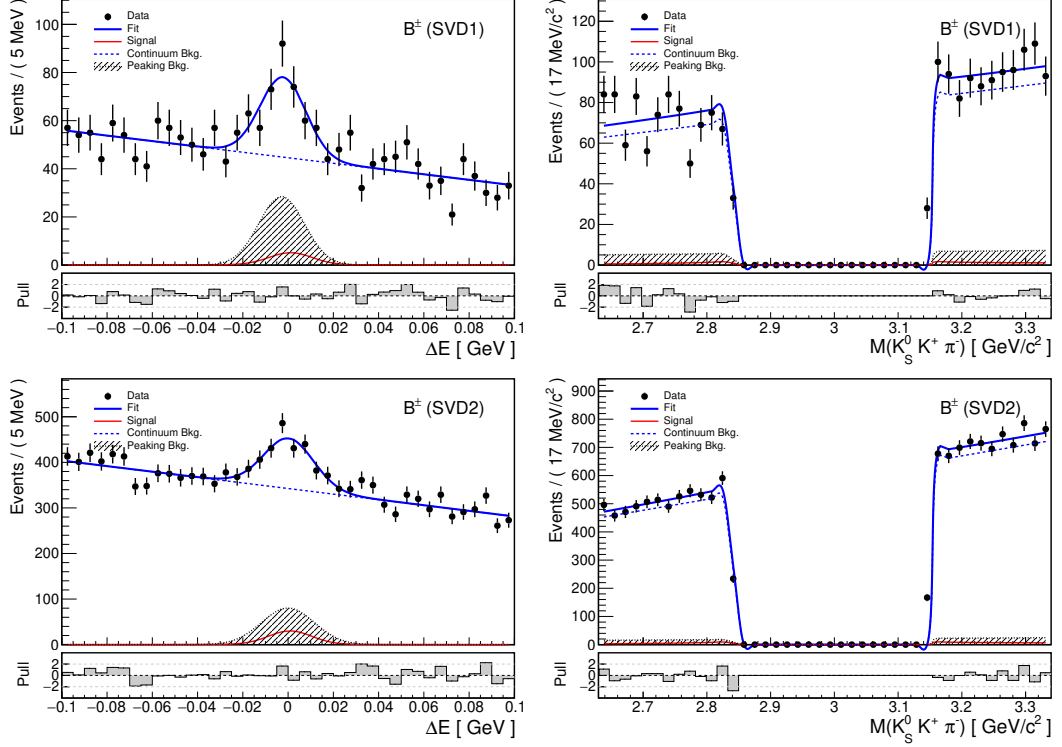
Figure 6.10: Projections of the 2D data model and the Belle data in the $M(K_S^0 K^+ \pi^-)$ sideband for the charged (control) mode and SVD1 (top) and SVD2 (bottom) experiment.

The fit projections with separated individual components are shown in Figures 6.10 and 6.11. Note that the edges of the bins do not match the sideband window, resulting in the drawing artifacts. A very important note is in order here: the depiction of the peaking background shape in the $M(K_S^0 K^+ \pi^-)$ variable is only illustrative. The actual slope could be different; only the total relative fraction to the continuum component is extracted. This will be true from now on in this thesis.

From the extracted fraction of the peaking background to signal $f_{\frac{\text{pb}}{\text{sig}}}$ in the $M(K_S^0 K^+ \pi^-)$ sideband ($f_{\frac{\text{pb}}{\text{sig}}}(SB)$), its linear shape in M, and almost symmetric choice of the sideband regions, one can naively estimate the peaking background fraction to signal in the central (signal) region ($f_{\frac{\text{pb}}{\text{sig}}}(C)$) as:

$$f_{\frac{\text{pb}}{\text{sig}}}(C) \simeq f_{\frac{\text{pb}}{\text{sig}}}(\text{SB}) \frac{\epsilon^{\text{SB}}}{2\epsilon}. \tag{6.19}$$
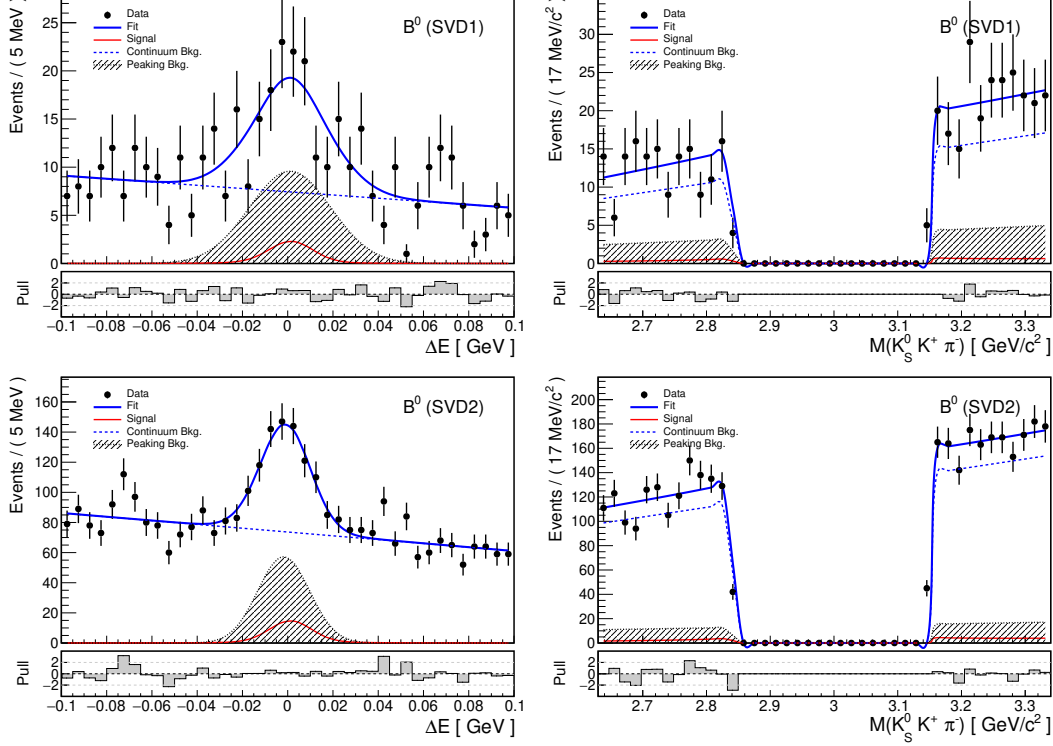
Figure 6.11: Projections of the 2D data model and the Belle data in the $M(K_S^0 K^+ \pi^-)$ sideband for the neutral (signal) mode and SVD1 (top) and SVD2 (bottom) experiment.

This estimate yields:

$$
\begin{aligned}
(21.2 \pm 5.5)\% &\text{ for SVD1, charged,} \\
(12.5 \pm 2.5)\% &\text{ for SVD2, charged,} \\
(51.9 \pm 12.3)\% &\text{ for SVD1, neutral, and} \\
(31.5 \pm 4.6)\% &\text{ for SVD2, neutral.}
\end{aligned}
\tag{6.20}
$$

Note, however, that the signal yield is fixed from the PDG and not a free parameter for this estimate. These numbers are only used as initial values for the fit in the signal region.

6.2 Measurement of Branching Fractions

In this section, the final 2D data model is used to extract the branching fractions from the signal yields—first for the charged control mode, which is used to determine also data/MC correction factors, and then for the neutral signal mode. This is merely for validation of the analysis before extensions to a more complicated fit which includes time difference and flavor dimensions. As the interference effects are only considered as a systematic uncertainty, we do not update the former Belle measurement for the control mode branching fraction [47]. Moreover, the final measurement of the branching fraction for the neutral mode is evaluated in the final fit, although this approach slightly increases systematic uncertainty. In

addition, the cross-checks of the correspondence of the 2D and the final 4D model predictions are important for the validation of the full fit.

6.2.1 Control Sample Measurement

The extended maximum likelihood fit results for the charged mode, including the correction factors for ΔE shape and other model parameters, are summarized in Table A.2. The observed number of events is 360 ± 31 and 1978 ± 75 for the SVD1 and SVD2 experiments, respectively. The fit projections are shown in Fig. 6.12.

The extracted $f_{\frac{\text{pb}}{\text{sig}}}$ can be compared to the estimate 6.20. Within statistical errors (as large as 50%), the values are compatible, but for SVD2, the fitter seems to predict a lower value than expected $(7.4 \pm 3.6)\%$ vs. $(12.5 \pm 2.5)\%$. A suspected bias toward a lower fraction of the peaking background when its yield is small was confirmed by preliminary toy MC studies. Including the fraction as a floating parameter in the final fit is preferred. In this way, the uncertainty is automatically propagated to the statistical uncertainty in the signal yield, including correlations. Moreover, the peaking background fraction and the background slope in the M variable are rather nuisance parameters whose real purpose is to partially absorb interference effects in the real data. With a non-resonant interfering background, separation to individual components requires an amplitude model for a proper description as a PDF[3].
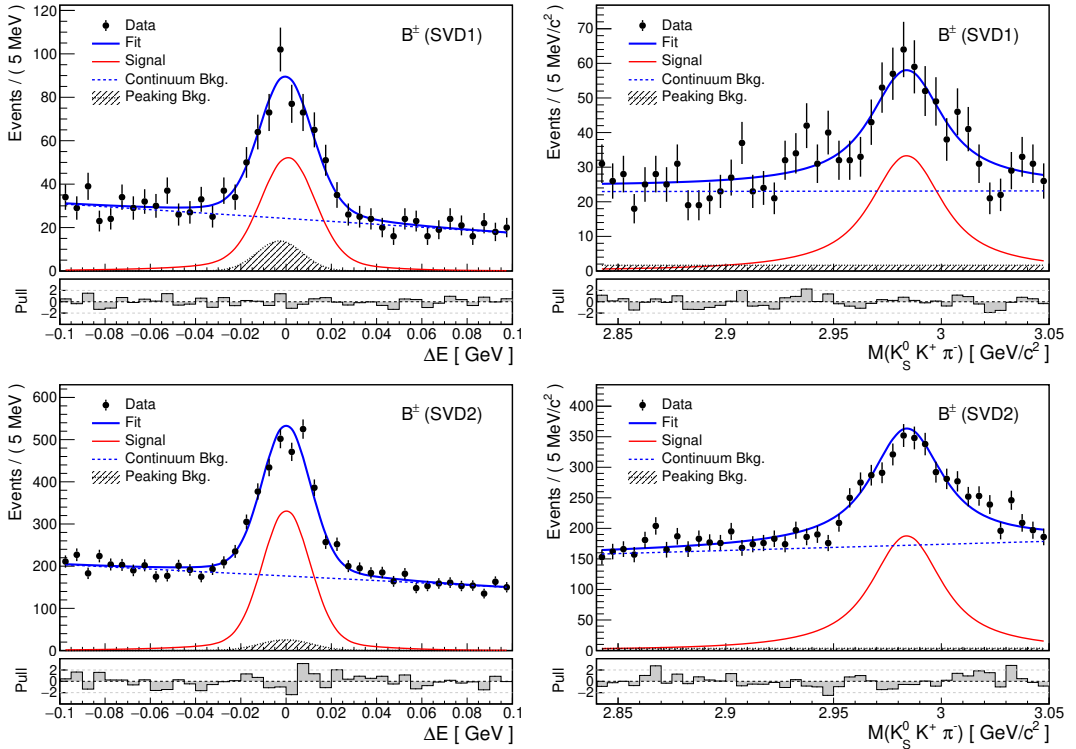
Figure 6.12: Projections of the final 2D data model and Belle data for the control mode and SVD1 (top) and SVD2 (bottom) experiment.

---

[3]With interference, individual components can have effectively negative weights due to destructive interference. PDFs and probabilities cannot be negative.

The extracted signal yield is compatible with the expectations from PDG; see Table 5.5. From the fitted signal yield, signal efficiency, and number of B-mesons for the given experiment (SVD1 or SVD2), the product of branching ratios can be extracted as

$$\mathcal{B} = \frac{N_{\text{sig}}}{\epsilon N_{B\overline{B}}}, \tag{6.21}$$

which yields

$\mathcal{B} = (2.47 \pm 0.23) \times 10^{-5}$ for SVD1,
$\mathcal{B} = (2.47 \pm 0.10) \times 10^{-5}$ for SVD2, and $\qquad$ (6.22)
$\mathcal{B} = (2.47 \pm 0.09) \times 10^{-5}$ as weighted average for the charged mode.

We will compare these results to the expectations in Sec. 6.2.3.

6.2.2 Neutral Mode Measurement

The fit results and other model parameters for the measurement of the signal yield in the neutral mode are shown in Table A.3. The observed number of events is 64 ± 15 and 548 ± 39 for the SVD1 and SVD2 experiments, respectively. The fit projections are shown in Fig. 6.13. The two data/MC calibration factors are fixed from the charged mode. As for the charged mode, the extracted peaking background fraction is compatible with the sideband extrapolation 6.20. The extracted signal yields are compatible with PDG expectations in Table 5.5 as well. Using Eq. 6.21, the product of branching ratios can be extracted:

$\mathcal{B} = (0.75 \pm 0.18) \times 10^{-5}$ for SVD1,
$\mathcal{B} = (1.02 \pm 0.08) \times 10^{-5}$ for SVD2, and $\qquad$ (6.23)
$\mathcal{B} = (0.98 \pm 0.07) \times 10^{-5}$ as weighted average for the neutral mode.

6.2.3 Summary

Concerning statistical errors only, the measured products of branching fractions for charged and neutral mode

$$\begin{aligned}
\mathcal{B}(B^+ \to \eta_c K^+) \times \mathcal{B}(\eta_c \to K_S^0 K^\pm \pi^\mp) &= (2.47 \pm 0.09) \times 10^{-5}, \\
\mathcal{B}(B^0 \to \eta_c K_S) \times \mathcal{B}(\eta_c \to K_S^0 K^\pm \pi^\mp) &= (0.98 \pm 0.07) \times 10^{-5}
\end{aligned} \tag{6.24}$$

are compatible with values from PDG[4]:

$$\begin{aligned}
\mathcal{B}(B^+ \to \eta_c K^+) \times \mathcal{B}(\eta_c \to K\bar{K}\pi) \times 1/3 &= (2.34 \pm 0.43) \times 10^{-5}, \\
\text{or explicitly } \mathcal{B}(B^+ \to \eta_c K^+, \eta_c \to K_S^0 K^\pm \pi^\mp) &= (2.7 \pm 0.6) \times 10^{-5},
\end{aligned} \tag{6.25}$$

$$\mathcal{B}(B^0 \to \eta_c K^0) \times \mathcal{B}(\eta_c \to K\bar{K}\pi) \times 1/2 \times 1/3 = (0.97 \pm 0.21) \times 10^{-5}.$$

---

[4]The explicit value obtained in Ref. [47] takes interference with non-resonant background into account in both the central value and the uncertainty, dominated systematically by the interference effect. As the neutral mode was not studied, we have partially reproduced this study in the next section to estimate at least the systematic uncertainty.
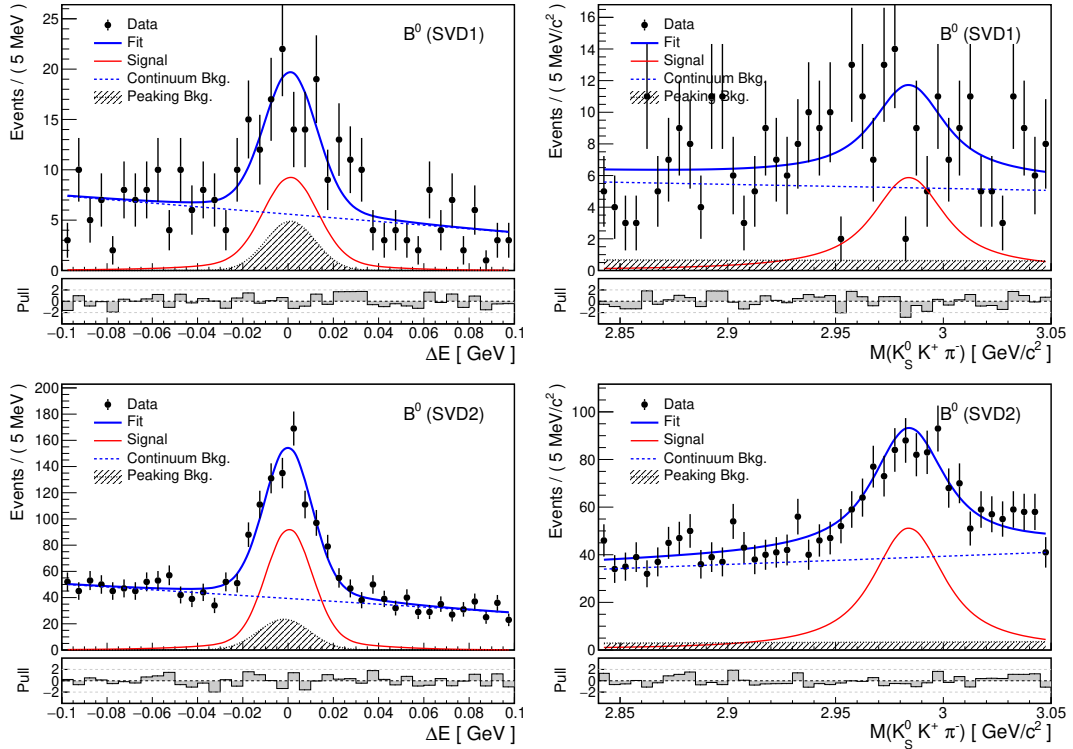
Figure 6.13: Projections of the final 2D data model and Belle data for the signal mode and SVD1 (top) and SVD2 (bottom) experiment.

The largest systematic error of the branching fraction measurement stems from the interference of the signal decay with a non-resonant (direct four-body) decay to the same final state. This effect was evaluated for the charged mode in Refs. [59, 47]. Taking the interference into account leads to a significant modification of the signal yield and additional large systematic error stemming from the model uncertainty. The next section is devoted to the estimation of this uncertainty also for the neutral (signal) mode. As it is very large, we do not estimate other systematic uncertainties here, but only in the final 4D fit and only for the signal mode.

6.3 Signal and Background Interference

So far, we have silently assumed that the signal and peaking background components can be described as independent PDFs. This would be, for example, the case of backgrounds with final state particle content different from the signal or when intermediate resonances which are far[5] apart in invariant mass are involved. In the case of $B \to \eta_c(K_S^0 K^+ \pi^-) K$ decays, multiple processes with the same final state constitute a significant contribution. Quantum-mechanical amplitudes of decays with the same final state can interfere with the signal amplitude and modify the signal line shape in the η_c invariant mass spectrum and

---

[5]Such that their mass peaks do not overlap. This is thus an issue for wide resonances, like the η_c meson.

apparent branching fraction. In addition, signal and background interference can lead to parasitic CP asymmetries, which can pollute our measurement of $\sin(2\phi_1)$ via $\mathcal{S}_{CP}$ ($\mathcal{A}_{CP}$) by introduction of additional strong and weak phases to the $B \to (K_S^0 K^+ \pi^-)K$ and $\overline{B} \to (K_S^0 K^+ \pi^-)K$ amplitudes[6].

Our full 4D fit model is already too slow and complicated to include more effects arising from the possible signal and background interference. Also, a full amplitude analysis with multiple components of various weak and strong phases is completely out of the scope of this thesis and beyond the statistical power of the used dataset. Instead, we devote this chapter to a set of independent studies with a simple χ^2-based fit, which will allow us to get more insight into the possible scale of the interference effects and put some constraints on their impact. The sole purpose of this chapter is thus only an estimation of the systematic uncertainty on the branching fraction measurement and derivation of a modified line shape which can be used to probe related fit biases of the CPV parameters using ensemble studies.

The effect of non-resonant background interference on the CPV parameters might lead to CP asymmetries, which are not constant as a function of the η_c resonance mass. Wide resonances are yet another interesting source of a strong[7] phase difference. We will see in this section explicitly that the phase of the Breit-Wigner amplitude changes rapidly around the mass peak, see 6.17 (right). If the non-resonant background strong phase is different (usually assumed to be constant in the simplest case), its relative weak phase can be revealed. As the non-resonant contribution should be mostly induced by $b \to s$ penguin decays, there is small pollution with such a weak phase difference. Any such effect should likely be small, but a direct confirmation from real data is, of course, always preferable. We will try to check for such a possibility in the final cross-checks assuming the simplest possible model proposed after a discussion with the review committee of the analysis.
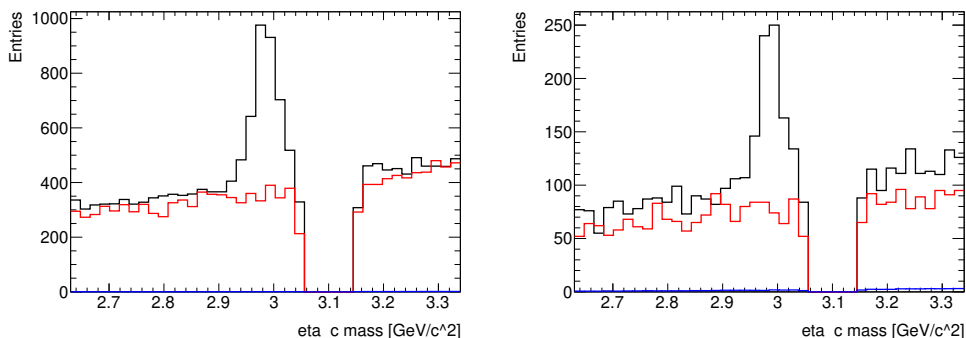
Figure 6.14: Histograms of the reconstructed η_c invariant mass for the real Belle data of the charged (left) and the neutral mode (right) from the signal (black) and sideband (red) regions in ΔE. The scaled distribution of the mis-reconstructed MC signal is shown in blue.

---

[6]In practice one of course measures (time-dependent) decay rates, not the (complex) amplitudes.

[7]In our context, strong phases are any phases that are CP invariant. In the time-dependent CP violation, this (time-dependent) phase comes from the unitary time evolution.

6.3.1 Background-subtracted Distributions

For all plots in this chapter, real data from the SVD1 and SVD2 experiments are merged to increase statistics. The continuum background is subtracted using a simple histogram method. The ΔE window is separated to signal ($|\Delta E| < 0.05$ GeV) and sideband regions (0.05 GeV $< |\Delta E| < 0.1$ GeV). The distribution of the reconstructed η_c invariant mass from the sideband region is then subtracted from the signal region distribution. In addition, we subtract a scaled histogram of the mis-reconstructed signal as obtained from the signal MC. The effect of neglecting the mis-reconstructed signal is small and most relevant for the neutral mode (where it is more abundant). The distributions before these subtractions are shown together in Fig. 6.14.

The distributions after background subtractions are shown in Fig. 6.15, where the arrows denote the boundaries of the additional separation into the η_c invariant mass signal and sideband regions. The presence of (ΔE) peaking background, flat in reconstructed invariant η_c candidate mass, is clearly visible and about 2–3 times larger for the neutral mode than the charged mode, in agreement with the more advanced extraction using the 2D likelihood fit.
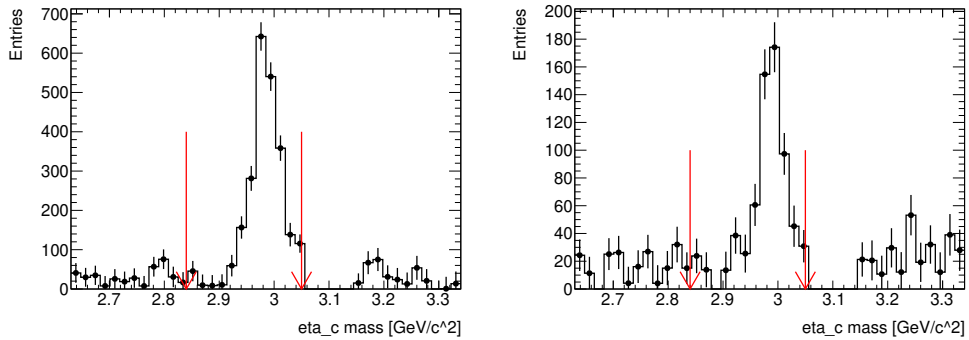
Figure 6.15: Background-subtracted distributions of the reconstructed η_c invariant mass with red arrows denoting boundaries of the η_c mass signal region for the charged (left) and neutral mode (right). The number of entries for some bins is negative as a result of the background subtraction.

The same approach is used to obtain background-subtracted Dalitz plots in the η_c mass signal region and distribution of $\cos(\theta)$, where θ is the angle between the momentum of K_S^0 from η_c and K_S^0 ($K^\pm$ for charged mode) from the B meson in the rest frame of the η_c candidate. These distributions are shown in Fig. 6.16. Several structures can be observed in the Dalitz plots, which are consistent with other measurements [60, 46] and suggests intermediate resonant structures in the η_c decay, with most prominent bands around 2 (GeV$^2/c^4$) with possible contributions from $K_0^*(1430)$, $K_0^*(1950)$ or $K_2^*(1430)$; the diagonal bands indicate the presence of a_0 and a_2 resonances [60].

6.3.2 Signal Line Shape with and without Interference

Let us consider a simplified signal model (compared to our full likelihood fit), where the broad η_c resonance with nominal mass m_{η_c} and width Γ is parametrized
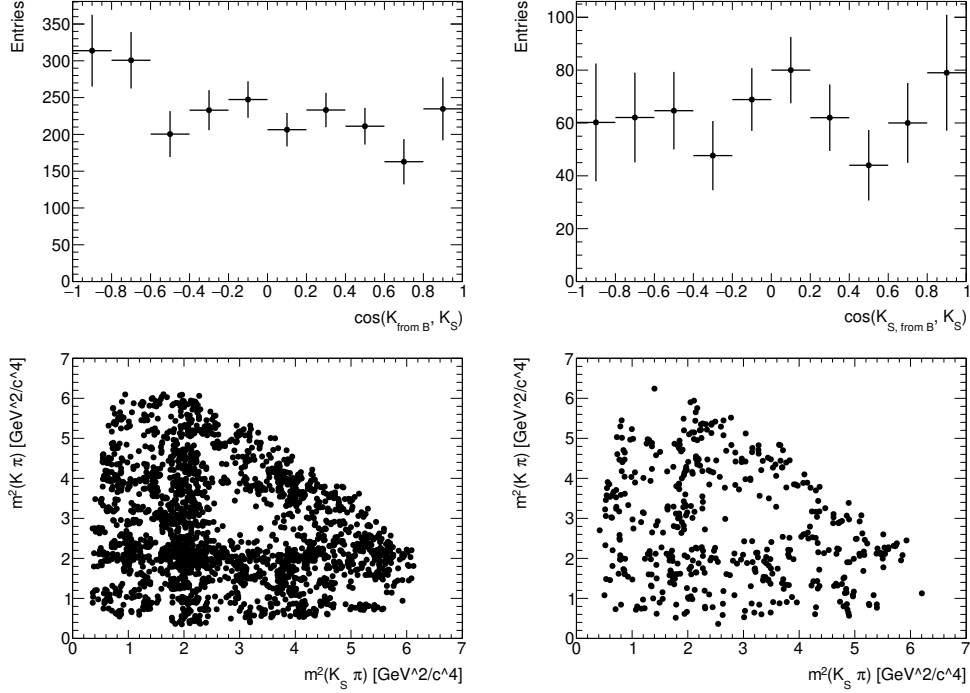
Figure 6.16: Background-subtracted distribution of the helicity cosinus between kaon from B and K_S^0 from η_c (top) and Dalitz plots (bottom). Figures for the charged (left) and neutral mode (right) are shown for SVD1 and SVD2 experiments combined.

by squared Breit-Wigner amplitude convolved with detector resolution, represented by a single Gaussian with width σ. Let us further assume the nominal case, where a constant peaking background does not interfere with the signal. The distribution of the reconstructed invariant mass $M \equiv M(K_S^0 K^+ \pi^-)$ can be written as

$$F(M) = N \left(\left| \frac{1}{M^2 - m_{\eta_c}^2 + im_{\eta_c}\Gamma} \right|^2 + \alpha^2 \right) \otimes G(M; 0, \sigma), \qquad (6.26)$$

where we omitted normalization constants, N is signal yield, and α^2 parametrizes the size of the peaking background contribution. For the resolution σ, we use the value determined in our full likelihood MC signal fit, 9 MeV/c^2. Now assume instead all the peaking background in ΔE is, in fact, a constant (in M) non-resonant contribution to the same final state and interferes with our signal. For the constant term, a complex phase ϕ arises, and the contribution is added coherently to the signal as

$$F(M) = N \left| \frac{1}{M^2 - m_{\eta_c}^2 + im_{\eta_c}\Gamma} + \alpha e^{i\phi} \right|^2 \otimes G(M; 0, \sigma). \qquad (6.27)$$

The square of the absolute value yields four terms:

$$\frac{1}{(M^2 - m_{\eta_c}^2)^2 + m_{\eta_c}^2\Gamma^2} + \frac{2\alpha\cos(\phi)(M^2 - m_{\eta_c}^2)}{(M^2 - m_{\eta_c}^2)^2 + m_{\eta_c}^2\Gamma^2} + \frac{-2\alpha\sin(\phi)m_{\eta_c}\Gamma}{(M^2 - m_{\eta_c}^2)^2 + m_{\eta_c}^2\Gamma^2} + \alpha^2.$$
$$(6.28)$$

102

The first one corresponds to the standard Breit-Wigner distribution. The second term is proportional to the real part of the Breit-Wigner amplitude, and the third one to its imaginary part. The distinct individual contributions (after convolution with detector resolution) are shown in Fig. 6.17. Depending on the size and phase of the interfering background, the measured signal yield can be significantly different. A fit to the data with assumption of no and full signal interference is compared in Fig. 6.18 and 6.19 for the charged and neutral modes, respectively. An asymmetric signal shape yields a better fit in both cases, and the difference in the measured signal yields suggests significant interference effects.

For the charged mode, the interference effect is less apparent in the signal shape asymmetry. When the interference is taken into account, the signal yield is reduced by 11% for the charged mode. For the neutral mode, the signal yield is reduced by 9% in the case of full interference. Note also the corresponding increase in the estimated statistical error of the signal yield (shown in the legend) in the case of interference. Furthermore, the parameters N, α, and ϕ are strongly correlated (from 58% to 85%).

The qualitative results for the charged mode correspond to previous results of a similar simplified study [59]. However, if we use the above-determined changes to the signal yield estimates as a systematic uncertainty, we get less than half of the number obtained in the full interference study [47] for the charged mode. Thus we use a more advanced interference model to estimate the systematics of the branching fraction in the next section. The results of this simple study will be used to estimate systematic uncertainty on the CPV parameters in ensemble studies with a modified signal line shape.
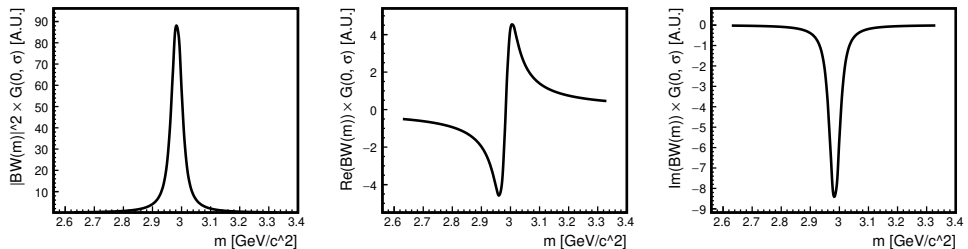
Figure 6.17: The square of Breit-Wigner amplitude (left), the real part (middle), and the imaginary part (right) convolved with detector resolution.

6.3.3 Angular Analysis of Four-Body Decay

To study the non-resonant background contributions with decay to the same final state, but without formation of the η_c resonance, we need to resort to an analysis of the full four-body decays $B^\pm \to K^\pm (K_S^0 K^\pm \pi^\mp)$ and $B^0 \to K_S^0 (K_S^0 K^\pm \pi^\mp)$. Narrow resonances have been removed by peaking background vetoes; see Sec. 5.3. Contributions from $b \to c$ decays (in particular with other charmonia) are small and can be neglected here. Events with wrongly identified final state particles constitute only a small (about 5% on rare-B MC) fraction of the peaking background. What remains are mostly true decays into the same final state as our signal. These can be further separated into direct four-body decays and quasi-three-body, and quasi-two-body decays with the formation of intermediate
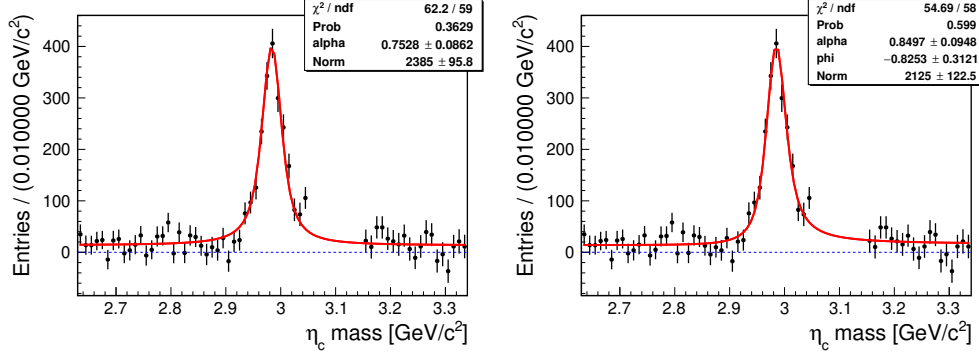
Figure 6.18: The fit result with the assumption of no interference (left) and with a simple interference model (right) for the charged control mode.
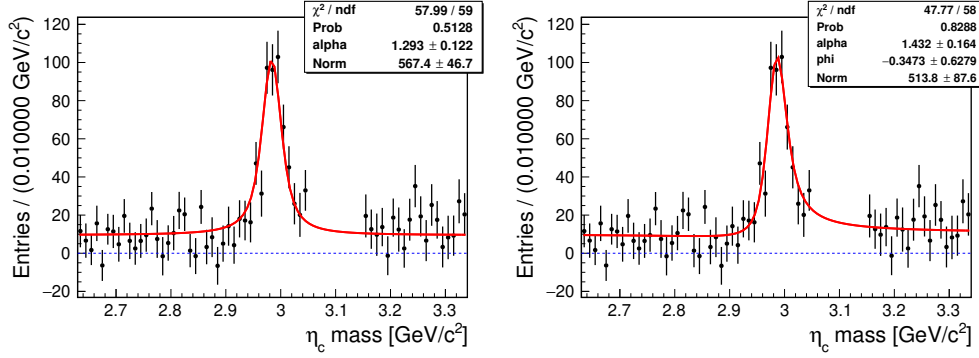
Figure 6.19: The fit result with the assumption of no interference (left) and with a simple interference model (right) for the neutral mode.

hadronic resonances. For example, the neutral rare-B MC sample contains about $1/4$ of $B^0 \to K_S^0 K_S^0 K^{*0}(892)$ decays which pass the event selection.

A general decay configuration of a (pseudo-) scalar particle (B meson) to four scalars is described by 3×4 momentum components of the final state particles. Three degrees of freedom correspond to the arbitrary orientation of the decaying B meson frame. With the additional four components of the four-momentum conservation law, this leaves five degrees of freedom to fully describe the final state configuration and thus the amplitude of the process. We select conveniently the same as in Ref. [47]:

- the invariant mass of the $K_S^0 K^{\pm} \pi^{\mp}$ combination with K_S^0 from the η_c candidate (usually denoted as M or $M(K_S^0 K^+ \pi^-)$ in this work), shown in Fig. 6.15,

- $\cos\theta$, where θ is the angle between the momentum of K_S^0 from η_c and K_S^0 ($K^{\pm}$ for charged mode) from the B meson in the rest frame of the η_c, see Fig. 6.16,

- two Dalitz variables $q_1^2 = M(K^{\pm}\pi^{\mp})$ and $q_2^2 = M(K_S^0 \pi^{\mp})$. Their distributions in the signal regions are shown in Fig. 6.16,

- angle ϕ between the planes defined by cross products of the vector of kaon (from the B meson) and pion, respectively charged kaon (from η_c) in the rest frame of the η_c candidate.

104

We set up a simple χ^2 fit of a 2D histogram in $s = M(K_S^0 K^+ \pi^-)$ and $\cos\theta$. Such a fit permits the separation of different (S, P, D) wave amplitude contributions because the pseudo-scalar η_c decay should exhibit a flat distribution in $\cos\theta$. Separation of other than S wave might help to constrain the interference effect and estimate possible systematic error on the signal yield when the interference is neglected in our full likelihood fit.

Similar to the previous section, the observed distribution can be obtained as a square of the sum of signal and S, P, and D background amplitudes, constant in the full mass window. For a full detailed formalism, we refer to Ref. [47]. The used fit function reads

$$
\begin{aligned}
F(s, x) =& (1 + \epsilon_1 x + \epsilon_2 x^2 + \epsilon_3 x^3) \times \\
& [S^2(x)(I_{\eta\eta}(s) + \alpha^2 + 2\sqrt{N}\alpha I_{\eta\eta}(s)) + P^2(x)\beta^2 + D^2(x)\gamma^2 \\
& + 2S(x)P(x)\beta(\sqrt{N}I_{\eta P}(s) + \alpha\Pi_{SP}) \\
& + 2S(x)D(x)\gamma(\sqrt{N}I_{\eta D}(s) + \alpha\Pi_{SD}) \\
& + 2S(x)P(x)\beta\gamma\Pi_{PD}],
\end{aligned} \tag{6.29}
$$

where $x = \cos\theta$, $s = M(K_S^0 K\pi)$, $S(x) = \frac{1}{\sqrt{2}}$, $P(x) = \sqrt{\frac{3}{2}}x$, $D(x) = \frac{3}{2}\sqrt{\frac{5}{2}}(x^2 - \frac{1}{3})$. Parameters ϵ_1, ϵ_2, ϵ_3 parametrize the detection efficiency as a function of x. The yields of the individual, S, P, and D background amplitudes are characterized by α, β, and γ, respectively. The parameters Π_{SP}, Π_{SD} and Π_{PD} represent the overlap integrals of the complex background amplitudes over the Dalitz variables and ϕ:

$$
\Pi_{ij} = \int\int\int \Re(A_i(q_1^2, q_2^2)A_j^*(q_1^2, q_2^2))q_1^2 q_2^2 d\phi, \tag{6.30}
$$

where $i, j = S, P, D$ and $i \neq j$. Similarly, the overlap integrals of η_c and background amplitudes are written as

$$
\xi_i(\cos\theta_i + i\sin\theta_i) = \int\int\int A_\eta(q_1^2, q_2^2)A_i^*(q_1^2, q_2^2)q_1^2 q_2^2 d\phi \tag{6.31}
$$

with six fit parameters $\xi_S = \xi$, $\theta_S = \theta$, ξ_P, θ_P, ξ_D, θ_D, where $0 \leq \xi_i \leq 1$ and $0 \leq \theta_i < 2\pi$, as the amplitudes are normalized over the Dalitz variables and ϕ. With this parametrization, the s-dependent terms can be written as

$$
\begin{aligned}
I_{\eta i}(s) = \xi_i \bigg[&\cos\theta_i \Re\left(\frac{1}{s^2 - m_{\eta_c}^2 + im_{\eta_c}\Gamma}\right) \otimes G(s; 0, \sigma) \\
&+ \sin\theta_i \Im\left(\frac{1}{s^2 - m_{\eta_c}^2 + im_{\eta_c}\Gamma}\right) \otimes G(s; 0, \sigma) \bigg],
\end{aligned} \tag{6.32}
$$

where the real and imaginary part of the Breit-Wigner amplitude convolved with the resolution function Gaussian, shown in Fig. 6.17, are precomputed to make the fit faster. For the resolution σ, we use the value determined in our full likelihood MC signal fit, 9 MeV/c^2.

The efficiency correction factors are determined from a fit to our full MC signal samples, with removed mis-reconstructed signal contributions. No resonant substructures are simulated in our MC, so the signal distribution over $\cos\theta$ should be flat. The deviation from flatness is attributed to the detection efficiency and parametrized by a third-order polynomial, fixed for later data studies. In contrast

to Ref. [47], we consider η_c mass and width as fixed externals parameters. This leaves 12 parameters to be determined by the 2D fit.

We believe (based on actual experience) that all these parameters cannot be reliably determined with the available statistics. In addition, amplitude fits suffer from issues with multiple minima (due to ambiguities from phases), and this is also our case. Therefore, we sample the starting values for the fit parameters uniformly in their range (reasonable values are used for α, β, and γ, which are not limited by physics bounds) and perform the fit. In many cases, the fit fails to estimate the statistical errors properly and can result in suboptimal final χ^2. We generate 1000 random sets of starting values for the parameters and select the solution with the lowest χ^2. However, the minimum forms a plateau for most parameters with almost identical χ^2 and different values of the fit parameters.

In Fig. 6.20 and Fig. 6.21 the best-fit projections are shown along with histograms of all reached fitted values of the parameters. Most of these fits are very close to the minimum. Based on our limited confidence in the fit, we prefer to estimate the uncertainties of the parameters by the standard deviation of these histograms, including all outliers. While this approach might seem too conservative, its impact on the estimated signal yield is of a similar scale as estimated in Ref. [47] for the charged mode in more advanced analysis. For the neutral mode, this effect is naturally larger due to the higher background level. These estimates yield a systematic error for the signal yield of 16% and 23% for the charged and neutral modes, respectively. These values are simply obtained as `(Std Dev)/Mean` from the histograms for parameter N in Figs. 6.20 and 6.21 and used to calculate the branching fraction uncertainty in the final result.
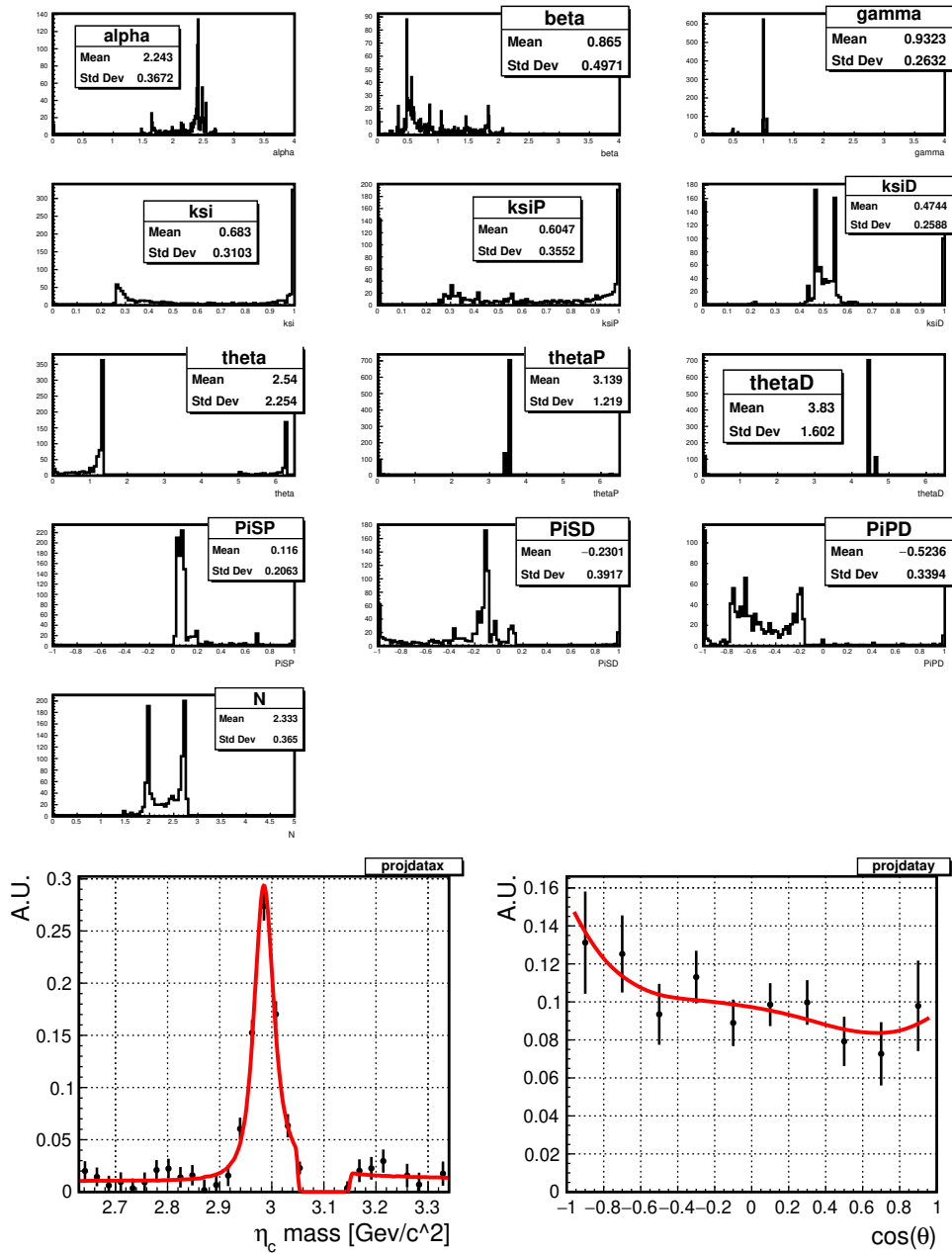
Figure 6.20: Fit projections (bottom) with data overlaid of the best fit for the charged mode. The upper grid of plots shows histograms of achieved optimal values for all floating parameters when the fit is initialized with random values sampled uniformly 1000 times.
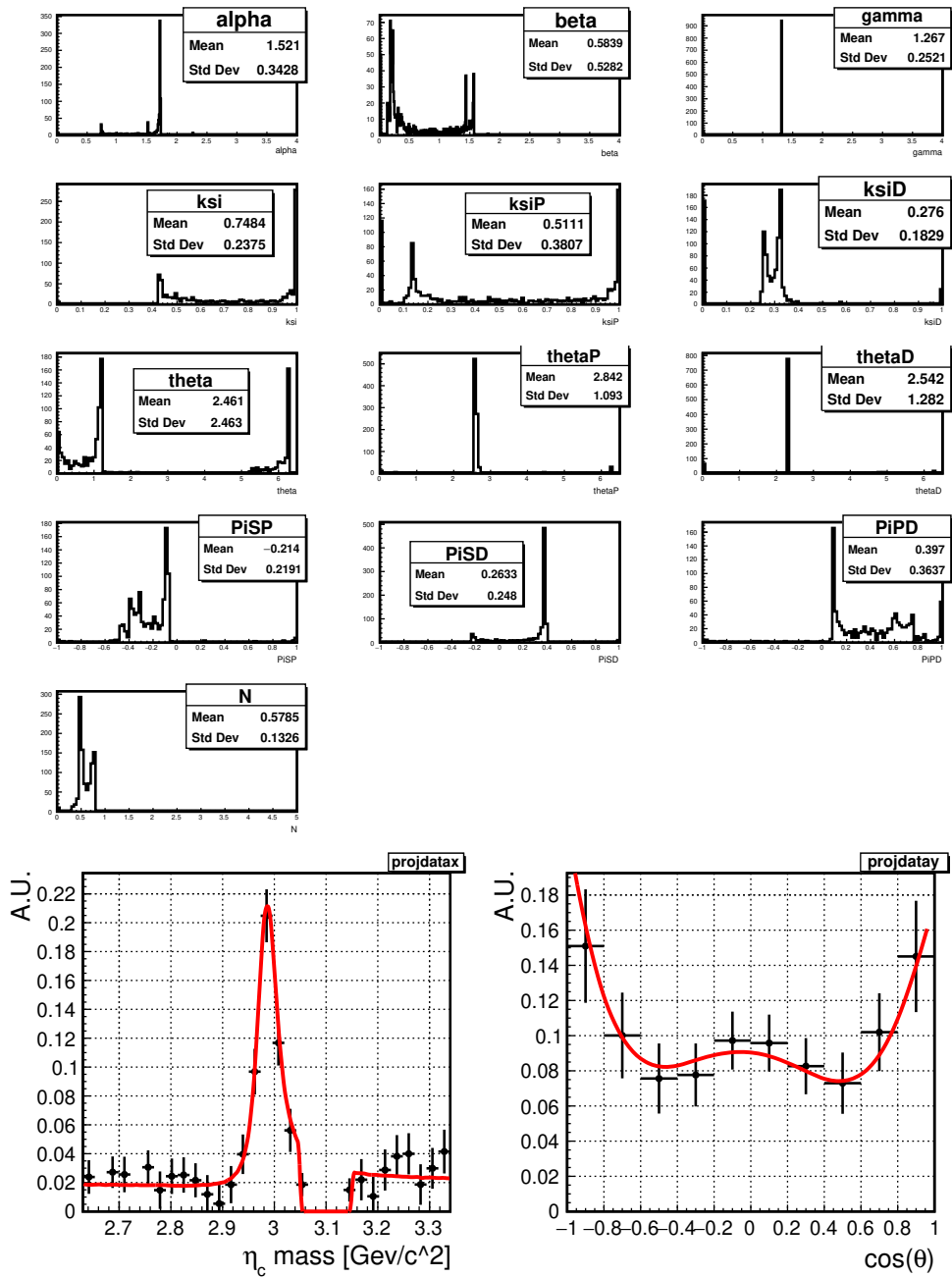
Figure 6.21: Fit projections (bottom) with data overlaid of the best fit for neutral mode. The upper grid of plots shows histograms of achieved optimal values for all floating parameters when the fit is initialized with random values sampled uniformly 1000 times.

Time-Dependent Measurements and Consistency Checks

7.1 4D Data Model

In this section, we extend the data model to the full fit in $\Delta E \times M \times \Delta t \times q \times l$, where Δt is the reconstructed proper time difference between B_{tag} and B_{sig}, q is the determined flavor q of B_{tag} and l is the index of the r-bin category. We call this a 4D model to distinguish it from a potential flavor-blind variant (not implemented in this work explicitly) which could be used to extract the B meson lifetime.

The discrete variable q is determined by the flavor tagging algorithm along with the probability of a correctly determined flavor characterized by categorizing each event into one of seven r-bins indexed by l. See also Sec. 4.6.3 and 5.1.

7.1.1 Flavour Tagging Quality Parametrization

The fractions of signal or background events in each of the seven r-bins are denoted as f^l, $l \in \{0, 1, 2, 3, 4, 5, 6\}$. For better stability of the extended maximum likelihood fit, the seven fractions f^l are transformed into a set of six actual fit parameters $\tilde{f}^k$, $k = 0...5$ as follows:

$$
\begin{aligned}
f^0 &= (1 - \tilde{f}^5)(1 - \tilde{f}^4)(1 - \tilde{f}^3)(1 - \tilde{f}^2)(1 - \tilde{f}^1)(1 - \tilde{f}^0) \\
f^1 &= (1 - \tilde{f}^5)(1 - \tilde{f}^4)(1 - \tilde{f}^3)(1 - \tilde{f}^2)(1 - \tilde{f}^1)\tilde{f}^0 \\
f^2 &= (1 - \tilde{f}^5)(1 - \tilde{f}^4)(1 - \tilde{f}^3)(1 - \tilde{f}^2)\tilde{f}^1 \\
f^3 &= (1 - \tilde{f}^5)(1 - \tilde{f}^4)(1 - \tilde{f}^3)\tilde{f}^2 \\
f^4 &= (1 - \tilde{f}^5)(1 - \tilde{f}^4)\tilde{f}^3 \\
f^5 &= (1 - \tilde{f}^5)\tilde{f}^4 \\
f^6 &= \tilde{f}^5 .
\end{aligned}
\tag{7.1}
$$

This satisfies by construction the required normalization constraint

$$
\sum_{l=0}^{6} f^l = 1.
\tag{7.2}
$$

The parameters $\tilde{f}^k$ are limited to the interval $[0, 1]$ in the fitter. For the signal (and peaking background), we introduce data/MC correction factors as additional

parameters. These parameters are set to one for MC, determined from the fit to the control mode, and fixed for the final fit in the neutral mode. Thus for the signal, the transformation in Eq. 7.1 is obtained by the replacement

$$\tilde{f}^k \to \eta_k^{\mathrm{CF}} \tilde{f}^k \tag{7.3}$$

where η_k^{CF} is a data/MC correction factor to the parameter $\tilde{f}^k$ and is set to vary in a range of $[0, 2]$ in the final fit to the control mode.

From a study with MC and real M_{bc} sideband data, it was concluded that two qualitatively different distributions of r-bin fractions could be observed. Samples that do not peak in ΔE exhibit a different r-bin distribution from the peaking backgrounds and signal. Therefore two sets of parameters are introduced, f_{sig}^l (and corresponding $\tilde{f}_{\mathrm{sig}}^k$ and η_k^{CF}, same for signal and peaking background), and f_{bkg}^l (with corresponding fit parameters $\tilde{f}_{\mathrm{bkg}}^k$) for continuum backgrounds; $l \in \{0, 1, 2, 3, 4, 5, 6\}$, $k \in \{0, 1, 2, 3, 4, 5\}$. The shape of continuum backgrounds in Δt does not depend on the r-bin, but we need to parametrize the distribution to be able to extract the data/MC correction factors for the signal from the final fit to the control sample.

7.1.2 Proper-Time Resolution Function for B Mesons

A crucial part of extracting *CP*-violation parameters from the observables is the decoupling of detector resolution effects, inclusive nature of the tag-side vertex reconstruction, and kinematic approximation, as briefly mentioned in Sec. 4.6.2. This is achieved by convolution of the physics distribution in Eq. 4.49 with the proper-time resolution function for B mesons $\mathcal{R}_{B\overline{B}}(\Delta t)$.

The $\mathcal{R}_{B\overline{B}}(\Delta t)$ resolution function and its parameters are provided by the `tatami` software package [61], which allows for fast analytical convolutions with the physics distribution. The functional form of $\mathcal{R}_{B\overline{B}}(\Delta t)$ is studied on MC simulations and parameters are extracted from high-statistics data samples and prepared by the I*CP*V (Indirect *CP* violation) working group [62]. We are using the latest available parameter set, denoted as `2010mdlh`, same as for the final Belle $\sin 2\phi_1$ analysis [55]. A different set of parameters is used for MC simulation and real data, as, for example, the vertex resolution is worse for data than for MC. The parameters are also different for the SVD1 and SVD2 experiment configurations (and some parameters are different for charged and neutral B mesons).

We do not determine any of the parameters, as those are mode-independent. The next section will verify that this common resolution function describes our MC signal sample well. Moreover, we could just copy-paste the codes used for $\eta_c \to p\bar{p}$ analysis [40], already verified by the previous work. For brevity, we are not going to explain all details and list all the parameters. Such a detailed summary can be found in Ref. [40]. The `tatami` package also provides estimated uncertainties for the parameters, which can be used for convenient systematic studies.

Let us briefly explain what the $\mathcal{R}_{B\overline{B}}(\Delta t)$ function actually looks like. The total resolution function is obtained (for a fixed set of its conditional variables) as a convolution of four components:

$$\mathcal{R}_{B\overline{B}} = \mathcal{R}_{\mathrm{det}_{\mathrm{rec}}} \otimes \mathcal{R}_{\mathrm{det}_{\mathrm{tag}}} \otimes \mathcal{R}_{\mathrm{np}} \otimes \mathcal{R}_{\mathrm{k}}. \tag{7.4}$$

The components follow from the three main effects affecting the proper-time reconstruction:

- Detector resolution for the B vertex determination for the fully reconstructed signal side $\mathcal{R}_{\mathrm{det_{rec}}}(\delta z)$ and the tag side[1] $\mathcal{R}_{\mathrm{det_{tag}}}(\delta z)$, where δz is the difference between the true and reconstructed signal or tag-side vertex z-position. This is a Gaussian for multi-track events and a sum of two Gaussians for single-track events. The width is given by event-by-event uncertainty on $\sigma_z^{\mathrm{rec,tag}}$ estimated from the vertex fit, scaled by parameters determined from fits to data. For multi-track vertices, an additional linear dependence of the width scaling is introduced, which depends on the reduced vertex fit quality h (for a vertex fit without the IP constraint), taking into account correlations observed on MC samples.

- Non-primary tracks (mostly from the charm and K_S^0 decays) affecting the tag-side vertex determination are taken into account in $\mathcal{R}_{\mathrm{np}}$. $\mathcal{R}_{\mathrm{np}}(\delta z_{\mathrm{tag}})$ is a sum of a delta function for prompt decays and an asymmetric double-sided exponential distribution with an event-by-event lifetime. This lifetime is a linear function of $\sigma_z^{\mathrm{rec,tag}}$ and the vertex quality h. Further differentiation of parameters is made among events with single and multi-track vertices as well as for events where a high momentum primary lepton is used at the tag side.

- Kinematic approximation due to neglected CM frame movement of the B mesons, which is reflected in $\mathcal{R}_{\mathrm{k}}(\Delta t - \Delta t_{\mathrm{true}})$. This is again in the form of exponential-like distributions, which are further functions of $\cos\theta_B^{\mathrm{CM}}$ and the B meson (charged or neutral) mass.

A detailed explanation of the individual components and how their parameters are obtained from MC and data can be found in Ref. [62].

Finally, let us note that some small fraction of outliers with a large Δt is not well described by a convolution with the resolution function and is instead parametrized by a Gaussian distribution with width σ_{ol} ($\sim$ 40ps) and a fraction f_{ol} ($\sim$ 3% for multi-track and $< 10^{-3}$ for single-track vertices). These outliers are added to the total PDFs in Δt for each component (signal and backgrounds) without convolution with the physics distribution.

7.1.3 Signal PDF

The two-dimensional[2] signal PDF in $\Delta t \times q$ follows from a convolution of the Δt resolution function for B mesons $\mathcal{R}_{B\overline{B}}$ and the true physics distribution, which has to be modified to take into account the effect of imperfect flavor determination. We model the distribution for each r-bin l separately (and fit simultaneously) to

---

[1]Those have the same definition, but different parameters

[2]The distribution is normalized over both variables Δt and discrete q

	Charged mode		Neutral mode	
	SVD 1	SVD 2	SVD 1	SVD 2
$\mathcal{S}_{CP}$				
Fit result	0.007 ± 0.016	0.009 ± 0.013	0.709 ± 0.020	0.703 ± 0.015
Generated	0.0		0.6889	
$\mathcal{A}_{CP}$				
Fit result	-0.009 ± 0.011	0.001 ± 0.009	-0.006 ± 0.014	0.022 ± 0.011
Generated	0.0		0.0	

Table 7.1: 4D Fit results for CPV parameters using signal-only MC data. Generated and fitted values are shown.

reach the best sensitivity. For a given r-bin l, we define

$$
\mathcal{P}_{\text{sig}}(\Delta t, q|l) = (1 - f_{\text{ol}}) \frac{e^{-|\Delta t|/\tau_B}}{4\tau_B} \Big\{ (1 - q\Delta w_l) +
$$
$$
q(1 - 2w_l) \times [\mathcal{S}_{CP} \sin(\Delta m_d \Delta t) + \mathcal{A}_{CP} \cos(\Delta m_d \Delta t)] \Big\} \otimes \quad (7.5)
$$
$$
\mathcal{R}_{B\overline{B}}(\Delta t) + f_{\text{ol}} \frac{1}{2} \mathcal{G}(\Delta t; 0, \sigma_{\text{ol}}),
$$

where w_l is mistag probability and Δw_l is the difference between wrong tag probabilities in r-bin l, see also Table 5.1. In addition, a fraction of outliers f_{ol}, not included in the resolution function, is added. The full 4D signal PDF reads

$$
\mathcal{P}_{\text{sig}}(\Delta E, M, \Delta t, q, l) = \mathcal{P}_{\text{sig}}(\Delta E|M) \mathcal{P}_{\text{sig}}(M) \mathcal{P}_{\text{sig}}(\Delta t, q|l) \mathcal{P}_{\text{sig}}(l)
$$

and its parameters are extracted using the MC signal samples.

The signal r-bin distributions $\mathcal{P}_{\text{sig}}(l) = f^l(\tilde{f}_{\text{sig}}^k, \eta_k^{\text{CF}})$ with $k = 0..5$ (see Eq. 7.1 and Eq. 7.3) for the signal (and peaking background, see below) component are further modified by data/MC correction factors η_k^{CF} (fixed to one for the MC fits) extracted from the control mode. The fit is also repeated for most signal shape parameters determined in the 2D fit to confirm the consistency of the results. The fitted parameters can be found in the summary Tables A.4 and A.5. The projections of data and PDFs for the SVD2 experiment are shown in Fig. 7.1. For the SVD1 experiment, the projections are shown in Fig. B.1.

This high-statistics signal-only fit also allows us to validate that correct CPV parameters can be extracted. In Table 7.1, the fitted CPV parameter values are compared to the values used for MC production. All results are statistically consistent (within 2σ) with the generated values. In a similar way, we checked fits to the B meson (charged and neutral) lifetimes with fixed CPV parameters and observed no significant deviations from the true (generated) lifetimes.

7.1.4 Peaking Background PDF

The PDF for the peaking background in $\Delta t \times q \times l$ is exactly the same as for the signal, but for the nominal fit result, the parameters $\mathcal{S}_{CP}$ and $\mathcal{A}_{CP}$ are set to
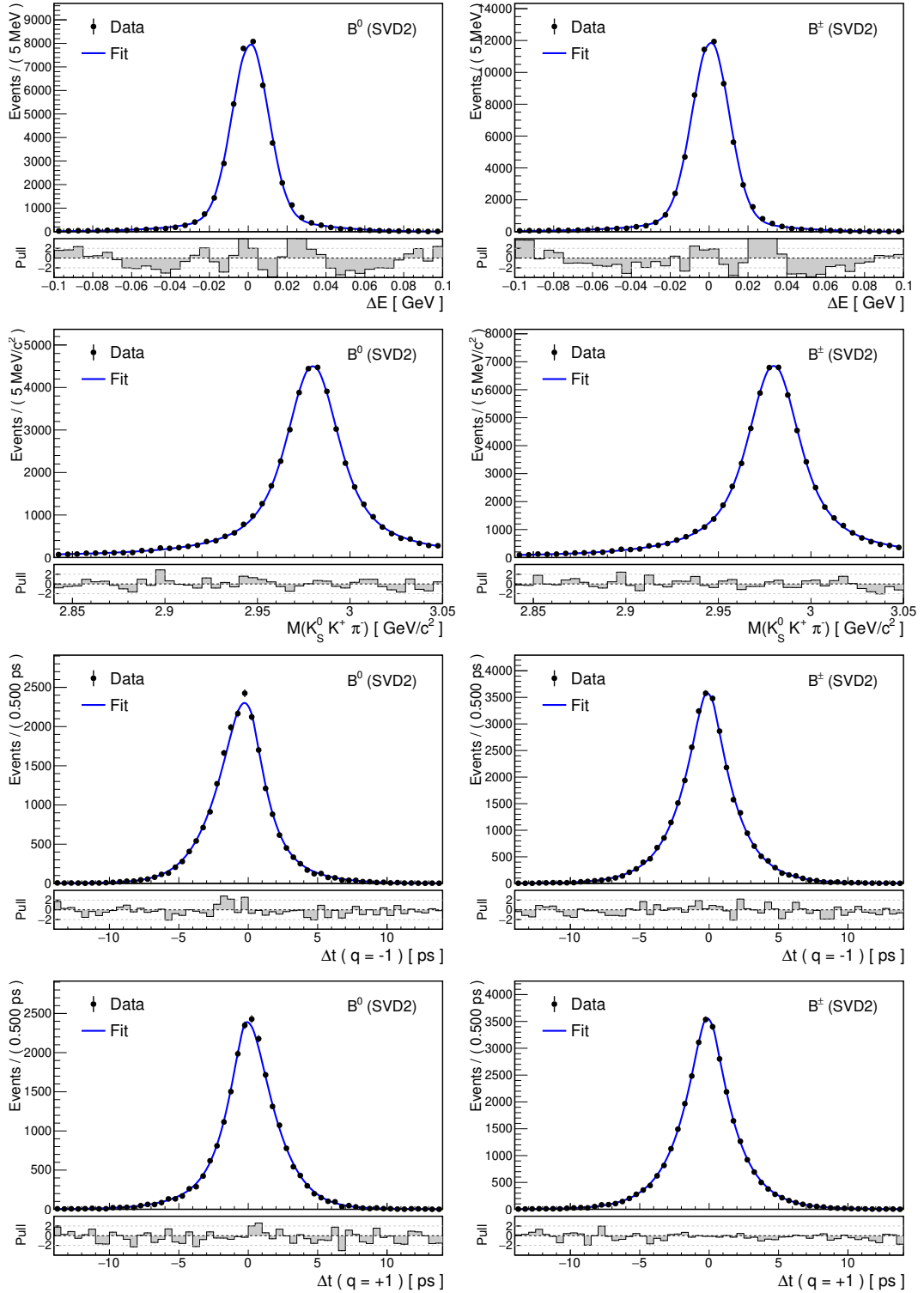
Figure 7.1: Projections of the SVD2 experiment MC signal and fitted signal PDF for the neutral (left) and control mode (right).

zeros (no *CP*-violation), i.e.:

$$\mathcal{P}_{\text{pb}}(\Delta t, q|l) = (1 - f_{\text{ol}})\frac{e^{-|\Delta t|/\tau_B}}{4\tau_B}\left[1 - q\Delta w_l\right] \otimes \mathcal{R}_{B\overline{B}}(\Delta t) + f_{\text{ol}}\frac{1}{2}\mathcal{G}(\Delta t; 0, \sigma_{\text{ol}}).$$

(7.6)

In some specific studies, we will determine effective $\mathcal{S}_f$, $\mathcal{A}_f$ for the peaking background. To accommodate for this case in the model, we use the full PDF and only set the *CP*V parameters to zero for the nominal fit. The full 4D PDF for the peaking background

$$\mathcal{P}_{\text{pb}}(\Delta E, M, \Delta t, q, l) = \mathcal{P}_{\text{pb}}(\Delta E)\mathcal{P}_{\text{bkg}}(M)\mathcal{P}_{\text{pb}}^l(\Delta t, q, l)\mathcal{P}_{\text{sig}}(l).$$

is using identical *r*-bin distribution as the signal component (including data/MC corrections from the control mode).

We have performed lifetime fits for the simulated MC signal, (peaking) generic $B\overline{B}$ and rare-*B* samples. The extracted lifetimes and *r*-bin distributions are consistent in all cases. This also means that the remaining small contamination from long-lived *D* mesons (see Table 5.3 and 5.2) can be neglected.

7.1.5 Continuum Background PDF

For the purpose of the 4D fit, the total continuum (i.e. non-peaking) background in $\Delta E \times M$ has to be further separated into a part stemming from non-*B* decays produced via $e^+e^- \to q\overline{q}$, $q = u, d, s, c$ denoted as $\mathcal{P}_{q\overline{q}}(\Delta t, q)$ and mis-reconstructed decays from $b\overline{b}$, denoted as $\mathcal{P}_{B\overline{B}}(\Delta t, q)$. Note that the shapes of these PDFs do not depend on q, and thus it is only present as a normalization factor $1/2 = \mathcal{P}_{\text{bkg}}(q)$. These decays occur typically when particles are mismatched between the tag and signal side of the reconstructed event. As a result, the effective lifetime for such decays is smaller than for the fully correctly reconstructed *B* decays. We model this contribution using the standard resolution function for *B* mesons $\mathcal{R}_{B\overline{B}}$ (as in the case of signal) but with a modified, effective lifetime τ_{eff}:

$$\mathcal{P}_{B\overline{B}}(\Delta t, q) = (1 - f_{\text{ol}})\frac{e^{-|\Delta t|/\tau_{\text{eff}}}}{4\tau_{\text{eff}}} \otimes \mathcal{R}_{B\overline{B}}(\Delta t) + f_{\text{ol}}\frac{1}{2}\mathcal{G}(\Delta t; 0, \sigma_{ol}),$$

(7.7)

taking again into account the effect of outliers as for the signal.

We extract the only free parameter, τ_{eff}, from a simultaneous fit to the SVD1 and SVD2 generic MC $B\overline{B}$ samples (with signal and peaking contributions removed). From the fit, we extracted the values

$$\text{Charged: } \tau_{\text{eff}} = 1.152 \pm 0.017 \text{ ps}$$
$$\text{Neutral: } \tau_{\text{eff}} = 1.111 \pm 0.033 \text{ ps}.$$

(7.8)

The fit projections are shown in Fig. 7.2, where we combined for the first time the SVD1 and SVD2 experiments into a single figure. We will continue to use this compact representation for all fits simultaneous in SVD1 and SVD2 experiments.

The model of the dominant part of the continuum background $\mathcal{P}_{q\overline{q}}(\Delta t, q)$ is composed of a distribution that describes a prompt component for events with
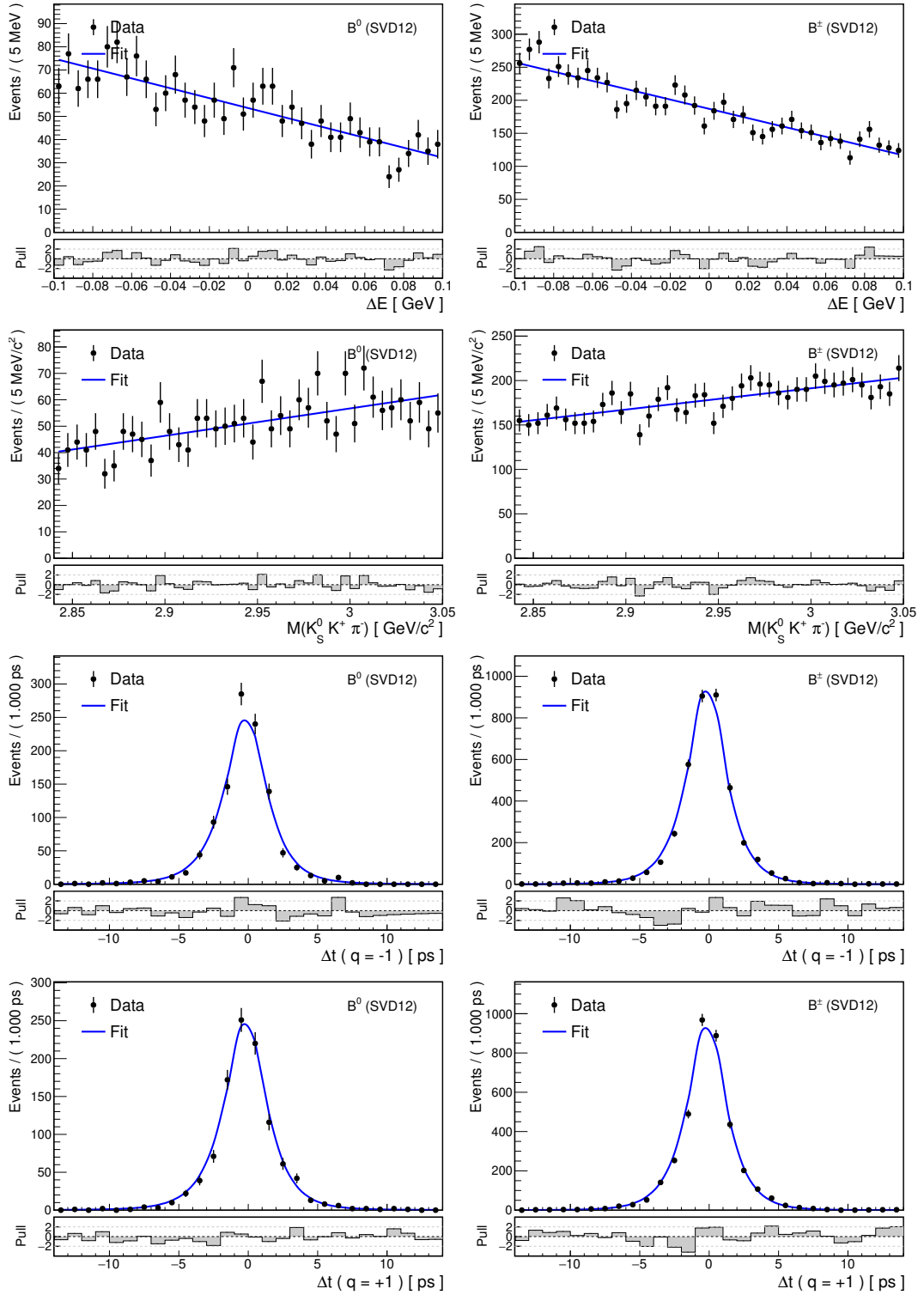
Figure 7.2: Projections of the combined SVD1 and SVD2 experiment generic $B\overline{B}$ MC (with peaking background contributions removed) and fitted PDFs for the neutral (left) and control mode (right).

a track coming directly from the interaction point (modeled as a delta function) with a fraction f_δ, and a contribution of non-prompt decays (due to charm mesons) with non-negligible effective lifetime $\tau_{\rm bkg}$. These distributions are convolved with a background resolution function, and a fraction for outliers is added:

$$
\begin{aligned}
\mathcal{P}_{q\bar{q}}(\Delta t, q) = \quad & \frac{1}{2}(1 - f_{\rm ol})\left\{ f_\delta \delta(\Delta t - \mu_\delta) + (1 - f_\delta)\frac{e^{-|\Delta t|/\tau_{\rm bkg}}}{2\tau_{\rm bkg}} \right\} \otimes \mathcal{R}_{q\bar{q}}(\Delta t) \\
& + \frac{1}{2}f_{\rm ol}\mathcal{G}(\Delta t; 0, \sigma_{\rm ol}),
\end{aligned}
\tag{7.9}
$$

where $\mathcal{R}_{q\bar{q}}$ is the background resolution function. In contrast to $\mathcal{R}_{B\overline{B}}$, the background resolution function is specific to this analysis and has to be determined from a fit to a relatively high-statistics $M_{\rm bc}$ sideband data sample. The resolution function has different parameters for events with single and multi-track vertices used for the B vertex reconstruction. In both cases, the resolution function is parametrized as a sum of two Gaussians:

$$
\mathcal{R}_{q\bar{q}}(\Delta t) = (1 - f_{\rm bkg}^{\rm tail})\mathcal{G}(\Delta t; \mu_{\rm bkg}, s_{\rm bkg}^{\rm main}\sigma_{\rm vtx}) + f_{\rm bkg}^{\rm tail}\mathcal{G}(\Delta t; \mu_{\rm bkg}, s_{\rm bkg}^{\rm tail}s_{\rm bkg}^{\rm main}\sigma_{\rm vtx}),
\tag{7.10}
$$

where $f_{\rm bkg}^{\rm tail}$ is the fraction of the wide tail component and $s_{\rm bkg}^{\rm main}$ and $s_{\rm bkg}^{\rm tail}$ are scale factors which multiply the event-dependent Δt resolution error

$$
\sigma_{\rm vtx} = \frac{1}{(\beta\gamma)_{\Upsilon(4S)}c}\sqrt{(\sigma_z^{\rm sig})^2 + (\sigma_z^{\rm tag})^2},
\tag{7.11}
$$

where $\sigma_z^{\rm sig}$ and $\sigma_z^{\rm tag}$ are the vertex position uncertainties (in z) of the reconstructed signal and tag B meson vertex, respectively. The parameters $f_{\rm bkg}^{\rm tail}$, $\mu_{\rm bkg}$, $s_{\rm bkg}^{\rm main}$ and $s_{\rm bkg}^{\rm main}$ are determined separately for events with only single charged track (with enough SVD hits) used for vertex reconstruction on either tag or signal side (single-track vertices) and the rest (multi-track vertices).

The $M_{\rm bc}$ sideband data sample has been found sufficient to determine all parameters of the continuum background. We extract the $\mathcal{P}_{q\bar{q}}(\Delta t, q)$ shape parameters in the 4D fit together with the distribution of r-bin fractions for the continuum background. As the same r-bin distribution is also used for $\mathcal{P}_{B\overline{B}}$, we have checked that generic $B\overline{B}$ MC reproduces the data distribution well. The fitted parameters of $\mathcal{P}_{q\bar{q}}$, fixed for the final fits, can be found in the summary Tables A.4 and A.5. The fit projections for the SVD2 experiment are shown in Fig. 7.3. For the SVD1 experiment, the projections are shown in Fig. B.2.

While the Δt shape parameters of $\mathcal{P}_{q\bar{q}}$ are determined fully from the $M_{\rm bc}$ sideband (almost free of B decays) and $\mathcal{P}_{B\overline{B}}$ from generic $B\overline{B}$ MC (with signal and peaking background removed), the fraction $(1 - f_{qq}^{\Delta t})$ of $\mathcal{P}_{B\overline{B}}$ still needs to be estimated in the signal region, where both components are present in a significant amount. The initial strategy was to estimate the fraction in the final fit. However, some concerns arise when estimating fractions only from the Δt component while using event-dependent resolution. Therefore for the measured neutral mode, we will estimate this fraction only from a fit to the $M(K_S^0 K^+\pi^-)$ sideband data (with subtracted leaking signal) in Sec. 7.2.1. The control mode measurement is performed using the former approach and we cross-check that both results are

consistent. The full continuum background PDF is a sum of the two components

$$\mathcal{P}_{\mathrm{cb}}(\Delta E, M, \Delta t, q, l) =$$
$$\mathcal{P}_{\mathrm{cb}}(\Delta E)\mathcal{P}_{\mathrm{bkg}}(M)\Big(f_{qq}^{\Delta t}\mathcal{P}_{q\bar{q}}(\Delta t, q) + (1 - f_{qq}^{\Delta t})\mathcal{P}_{B\overline{B}}(\Delta t, q)\Big)\mathcal{P}_{\mathrm{bkg}}(l),$$
(7.12)

which share their shape in $\Delta E \times M \times q \times l$ and only differ in the Δt dimension. The background model is completely flavor-symmetric, i.e., the shape does not depend on q. The r-bin distribution $\mathcal{P}_{\mathrm{bkg}}(l)$ is distinct from signal and depends only on the parameters $\tilde{f}_{\mathrm{bkg}}^{k}$, $k = 0..5$. Thus no correction factors are introduced, and the shape is extracted directly from the fit to M_{bc} sideband data for the neutral mode and from the final fit in the control mode.

7.1.6 Full Log-Likelihood Function

The *CP*-violation parameters are extracted along with the signal and background yields and additional nuisance parameters (the exact configuration of fixed and floating parameters differs for the control charged and the neutral mode) using the following extended log-likelihood

$$\ln \mathcal{L} = -(N_{\mathrm{sig}} + N_{\mathrm{bkg}}) + \sum_{i} \ln \Big(N_{\mathrm{sig}}\mathcal{P}_{\mathrm{sig}}(\Delta E^i, M^i, \Delta t^i, q^i, l^i)$$
$$+ f_{\frac{\mathrm{pb}}{\mathrm{sig}}} N_{\mathrm{sig}}\mathcal{P}_{\mathrm{pb}}(\Delta E^i, M^i, \Delta t^i, q^i, l^i)$$
$$+ (N_{\mathrm{bkg}} - f_{\frac{\mathrm{pb}}{\mathrm{sig}}} N_{\mathrm{sig}})\mathcal{P}_{\mathrm{cb}}(\Delta E^i, M^i, \Delta t^i, q^i, l^i)\Big)$$
(7.13)

either for each data set separately or, as in the final fit, simultaneously for the SVD1 and SVD2 experiments with the only common parameters $\mathcal{S}_{CP}$, $\mathcal{A}_{CP}$, and all input physics parameters from the PDG ($\tau_{B^\pm}/\tau_{B^0}$ is released only for validation tests):

$$\ln \mathcal{L} = \ln \mathcal{L}(\mathcal{S}_{CP}, \mathcal{A}_{CP}, ...; (\Delta E^i, M^i, \Delta t^i, q^i, l^i) \in \mathrm{SVD1})$$
$$+ \ln \mathcal{L}(\mathcal{S}_{CP}, \mathcal{A}_{CP}, ...; (\Delta E^i, M^i, \Delta t^i, q^i, l^i) \in \mathrm{SVD2}).$$
(7.14)

Every event is classified by the flavor tagger into one of seven r-bins l (which influence signal Δt shape through w_l and Δw_l). The fact that the r-bin distribution is uneven and different for signal and (continuum and non-peaking) background is reflected in the r-bin fraction parameters and the corresponding PDFs[3]. These represent the probabilities of observing a particular r-bin in a signal/background event and play a similar role as Punzi constraint terms on the likelihood [63].

The PDF is, in contrast to the 2D case, a function of several (conditional) event-dependent variables besides the explicitly stated five ($\Delta E, M, \Delta t, q, l$). The shape Δt resolution functions depend on event-dependent candidate properties. For $\mathcal{R}_{B\overline{B}}$, the actual parametrization depends on event categorization based on the following flags:

- whether only a single charged track or multiple tracks were used for B_{tag} or B_{sig} vertex reconstruction (single or multi-track flag),

---

[3]Which can be represented as histograms with seven bins.
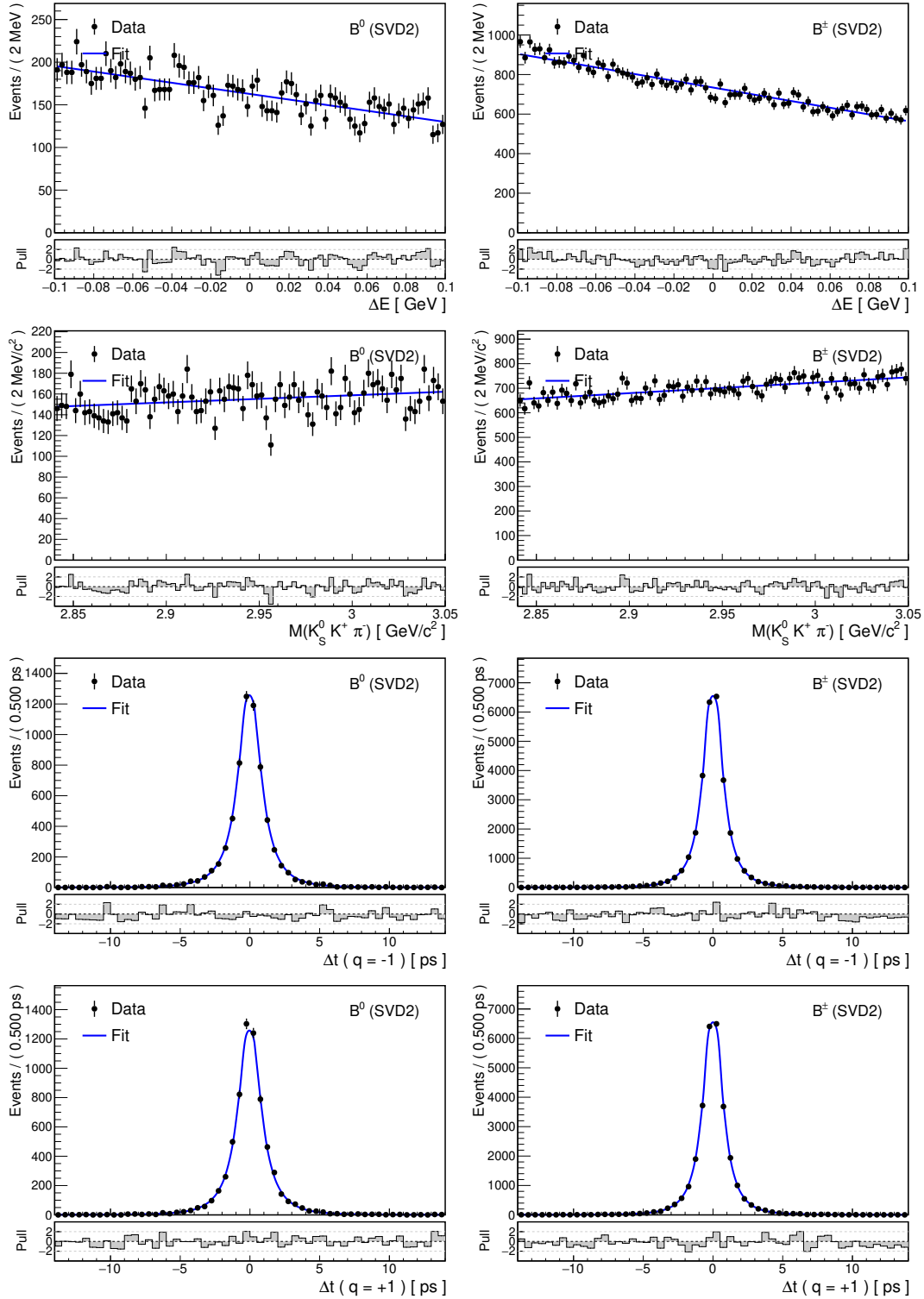
Figure 7.3: Projections of the SVD2 experiment M_{bc} sideband data and fitted continuum PDF for the signal (left) and control mode (right).

- whether a lepton was used for the B_{tag} vertex reconstruction (lepton flag),

and on multiple properties which are the output of the reconstruction procedure, namely:

- The vertex reconstruction error in z for signal and tag B meson, σ_z^{tag} and σ_z^{sig}

- Reconstructed vertex number goodness-of-fit h for signal and tag side.

- Kinematic properties of the B meson $\cos\theta_B^{\text{CM}}$, p_B^{CM}, and E_B^{CM} needed for parametrization of the kinematic approximation part of the resolution function. All these inputs are also event-dependent, depending on the properties of the reconstructed B_{sig} candidate and the calibrated beam energy.

For $\mathcal{R}_{q\bar{q}}$ only σ_z^{tag} and σ_z^{sig} and single/multi-track flag are needed.

The distributions of these variables are not part of the data model. This has some potential drawbacks related to a so-called Punzi effect [63] if the distributions of these variables differ for signal and background events, see also Sec 7.3.5.

7.2 Control Mode Measurement

The control sample is used not only to check the consistency of the fit on a very similar decay but also to determine data/MC correction factors to the signal r-bin fractions and to the mean and width of the ΔE signal shape, later fixed in the final fit to neutral mode. Many parameters of the full 4D fit have to be estimated from MC studies or fits to special data samples. The parameters fixed in the final control mode fit are:

- Signal shape parameters in $\Delta E \times M$ determined from a fit to the MC signal sample. The mass and width of η_c are fixed to the most up-to-date PDG values, as well as parameters for B meson mixing and B lifetimes (only released for validity check).

- Effective lifetime τ_{eff} for the non-peaking $B\overline{B}$ background (included in our continuum component) is extracted from the fit of the generic $B\overline{B}$ MC sample (non-peaking, with removed signal).

- Peaking background mean and width extracted from the $M(K_S^0 K^+\pi^-)$ sideband.

- Continuum background Δt shape parameters determined from the M_{bc} sideband.

The parameters determined in the final control mode fit, different for the SVD1 and SVD2 experiments are:

- Signal and background yields, and the fraction of peaking background.

- Data/MC correction factors for signal r-bin fractions and ΔE shape.

- Continuum background $\Delta E \times M$ shape parameters.

- Fraction $f_{qq}^{\Delta t}$.

Finally, the only floating parameters common to SVD1 and SVD2 are the physics parameters $\mathcal{S}_{CP}$ and $\mathcal{A}_{CP}$. The results of separate fits have also been compared to the simultaneous fit results. All determined parameters are entirely consistent for both fit configurations. In addition, the signal and background yields and peaking background fractions are fully[4] consistent with the result of the 2D fit. This is an important cross-check validating the model in the Δt and flavor dimensions. For the control sample, the fit results are shown together with other model parameters in the summary Table A.4. The observed number of events is 358 ± 30 and 1990 ± 70 for the SVD1 and SVD2 experiments, respectively. The physics parameters of interest

$$
\begin{aligned}
\mathcal{S}_{CP} &= -0.03 \pm 0.08 \\
\mathcal{A}_{CP} &= -0.05 \pm 0.06 \\
\mathcal{B} &= (2.49 \pm 0.09) \times 10^{-5}
\end{aligned}
\tag{7.15}
$$

are consistent with the SM predictions and the PDG-based estimate for $\mathcal{B}$ ($\mathcal{B} = (2.34 \pm 0.43) \times 10^{-5}$, respective $\mathcal{B} = (2.7 \pm 0.6) \times 10^{-5}$, see Sec. 6.2.3) within statistical errors.

The fit and data projections are shown in Fig. 7.4, where SVD1 and SVD2 experiment data and models are combined. The distribution of r-bins is shown in Fig. 7.5. We project only a limited, signal-enhanced region, defined by the following selections

$$
\begin{aligned}
&-40 \text{ MeV} < \Delta E < 40 \text{ MeV} \\
&2935 \text{ MeV}/c^2 < M(K_S^0 K^+ \pi^-) < 3035 \text{ MeV}/c^2 \\
&|\Delta t| < 7\text{ps}.
\end{aligned}
\tag{7.16}
$$

For projections in the full signal region, see Fig. C.1. After projecting the signal-enhanced region, the estimated background PDF is subtracted from the data histograms for each flavor (charge) of the tagged B_{tag}. The upper plot in Fig. 7.6 compares these background-subtracted distributions to the signal PDF projection. The bottom plot shows the raw asymmetry directly, without any background subtractions, formed as $(N^+ - N^-)/(N^+ + N^-)$, where the $N^{+,-}$ is the number of observed events in each flavor category, with Poisson errors. The same asymmetry operation is applied to the histograms accumulated from the total model PDFs.

7.2.1 Extracting $f_{qq}^{\Delta t}$ from $M(K_S^0 K^+ \pi^-)$ Sideband

In the nominal control sample measurement, we have determined the fraction of the (mis-reconstructed, non-peaking) $B\overline{B}$ background to the total continuum background, defined as $1 - f_{qq}^{\Delta t}$. This fraction is determined separately for the SVD1 and SVD2 experiment in the final fit (see Table A.4) and using only the

---

[4]Signal and background yields are almost identical. Some variations can be observed in the estimated peaking background fractions (to signal), but only in the SVD1 experiment with much less data. Nevertheless ,the values agree well within 1σ.
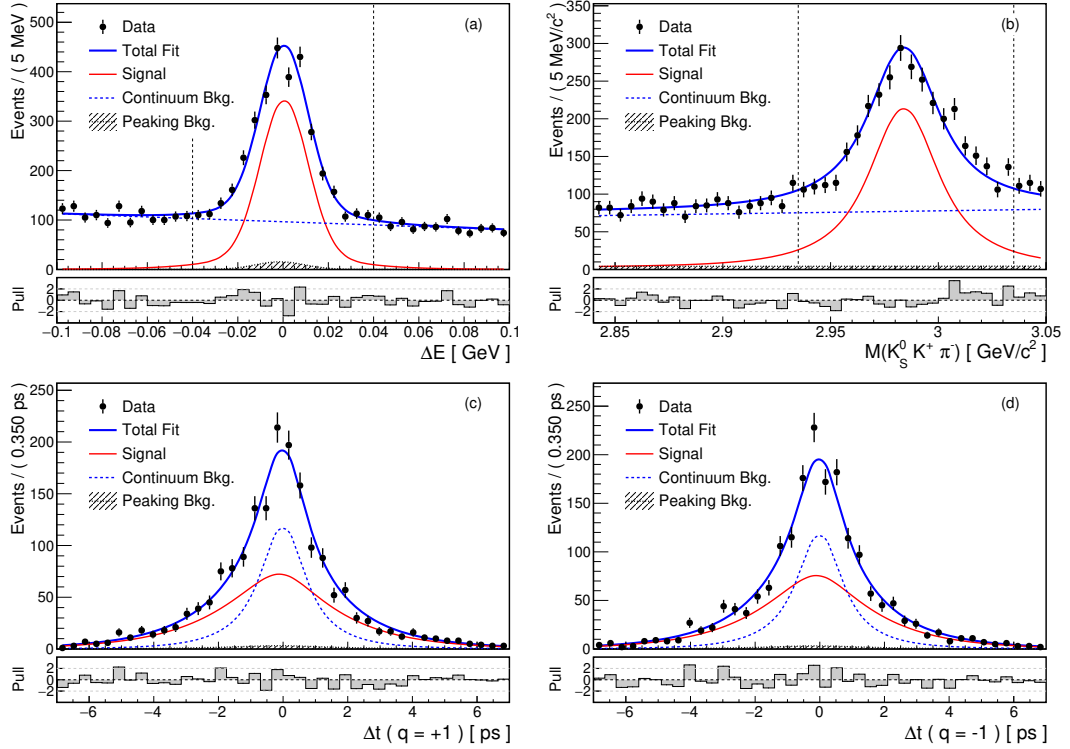
Figure 7.4: Control mode projections in ΔE (a), $M(K_S^0 K^+ \pi^-)$ (b) and proper time difference Δt for B^+ (c) and B^--tagged (d) events (q is the flavor of B_{tag}) in the signal-enhanced region (the distributions in Δt are for events in the dotted window in ΔE and M). Curves show the fit model and its components, and points represent the data. SVD1 and SVD2 experiment data and model are combined. Continuum background component includes non-peaking $B\overline{B}$ background.
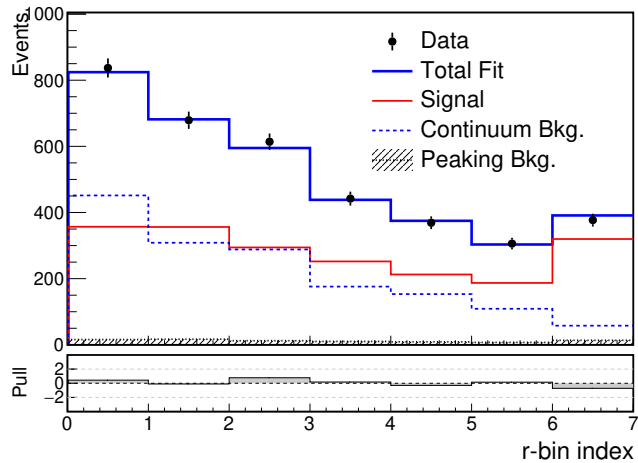
Figure 7.5: Control mode projections of r-bin distributions in the signal-enhanced region. Lines show the fit model and its components, and points represent the data. SVD1 and SVD2 experiment data and model are combined. Continuum background component includes non-peaking $B\overline{B}$ background. Note that the pulls are small because the r-bin fraction parameters (and data/MC correction factors) are floating.
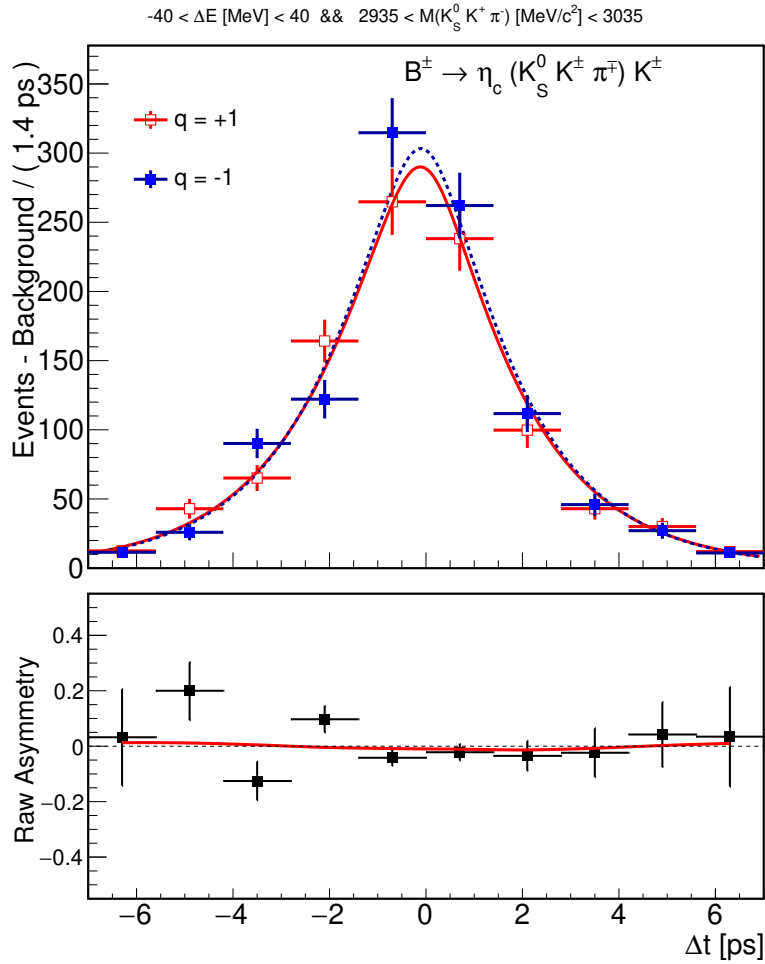
Figure 7.6: Control mode projections of background-subtracted distributions of the estimated number of signal B^+ and B^--tagged events (q is the flavor of B_{tag}) (top) and raw asymmetry $(N^+ - N^-)/(N^+ + N^-)$ of the number of total events (bottom). The distribution is integrated over all r-bins. Only the signal-enhanced region is projected. Curves show the fit model; points represent the data. SVD1 and SVD2 experiment data and model are combined.

information in the Δt dimension. The extraction of this parameter is possible due to the effective lifetime τ_{eff} of the $B\overline{B}$ component (about 1.11 ps), which is larger than the lifetime of the remaining continuum background from u, d, s, c decays (determined from M_{bc} sideband with lifetime dominated by charm contributions). However, the resolution functions in the Δt dimension depend on event-dependent variables, which are not included in the data model. This could cause biases due to the Punzi effect [63]. Therefore for the final neutral mode measurement, we will fix this fraction.

To determine this fraction in the neutral mode prior to un-blinding, we can extract $f_{qq}^{\Delta t}$ using the $M(K_S^0 K^+ \pi^-)$ sideband, this time in the full 4D fit. We need to fix signal CPV parameters to some values to do so. We check the determined fraction with several assumed (extremal, zero, or expected) values for the signal CPV. The floating parameters are the fraction of peaking background to signal, background yield, continuum shape parameters, and $f_{qq}^{\Delta t}$. All other parameters are set to their final values prepared for the unblinded fit, including data/MC correction factors determined from the nominal fit to the control sample. When varying the assumed signal CP violation, we observe only small, sub-percent variations of the estimates for $f_{qq}^{\Delta t}$. Thus we will use the values determined when $\mathcal{S}_{CP} = 0$ and $\mathcal{A}_{CP} = 0$, which read:

$$f_{qq}^{\Delta t}(\text{SVD1 neutral}) = 0.834 \pm 0.114$$
$$f_{qq}^{\Delta t}(\text{SVD2 neutral}) = 0.803 \pm 0.030$$

(7.17)

To confirm the method of $f_{qq}^{\Delta t}$ extraction is reasonable, we use the control mode results and fix all parameters except for a fraction of peaking background to signal, background yield, continuum shape parameters, and $f_{qq}^{\Delta t}$. We then perform the full 4D fit (with nominal CPV parameters at zero) in the $M(K_S^0 K^+ \pi^-)$ sideband and obtain

$$f_{qq}^{\Delta t}(\text{SVD1 charged}) = 0.843 \pm 0.039$$
$$f_{qq}^{\Delta t}(\text{SVD2 charged}) = 0.820 \pm 0.012$$

(7.18)

which are statistically consistent with the results in the signal window in Table A.4.

7.3 Validity Studies and Consistency Checks

After confirmation from the control channel, we continued with cross-checks of the analysis and data model before moving to the signal channel. Next to ensemble tests and specialized fits, dedicated studies are performed to evaluate specific sources of systematic uncertainties.

7.3.1 Control Mode Lifetime Fit

The lifetime fit can serve as a validation of the resolution function. We perform the final fit in the charged mode with $\mathcal{S}_{CP}$, $\mathcal{A}_{CP}$ fixed to zero, and $\tau_{B^\pm}$ (common to SVD1 and SVD2 experiment) floating. The resulting

$$\tau_{B^\pm} = (1.696 \pm 0.045) \text{ ps}$$

(7.19)

123

is consistent with the world average (1.641±0.008) ps within statistical error (1.2σ away). Performing the same fit separately for the SVD1 and SVD2 experiments yields

$$\text{SVD1}: \quad \tau_{B^\pm} = (1.737 \pm 0.115) \text{ ps}$$
$$\text{SVD2}: \quad \tau_{B^\pm} = (1.688 \pm 0.049) \text{ ps.} \tag{7.20}$$

7.3.2 Systematic Study of $\mathcal{CS}_{\text{cut}}$ for Control Mode

For the control mode, we vary the continuum suppression cut on the likelihood ratio. The estimations for the *CP*-violating parameters as a function of this cut were already shown in Fig. 5.11. In fact, only after this study the $\mathcal{CS}_{\text{cut}}$ was finally fixed. The estimated values for the *CP*V parameters are stable within a wide range of $\mathcal{CS}_{\text{cut}}$. At tight cut values, larger oscillations are caused by significant variations of the statistics. The statistical errors, shown in red, exhibit a shallow minimum around the loosest reasonable value of the cut of ~ 0.2. This shows that maximizing signal efficiency is the primary goal of reducing statistical error on the parameters of interest in our fit model. Moreover, the background parameter extraction uses the same cut (not to generate additional systematics). The loose cut value results in a higher precision of the background parameters. Finally, the loose cut means that signal yield is almost unaffected by small changes in the cut value, whereas a tight cut may generate additional systematic effects (due to rapidly changing signal yield for a slightly varied cut).

7.3.3 *CP* Violation in $M(K_S^0 K^+ \pi^-)$ Sideband

As discussed in section 5.4, a fraction of the total peaking background, stemming from $b \to c$ decays, is expected to be *CP*-violating and thus can bias our measurement. From the MC studies, the estimated fraction of $b \to c$ backgrounds to the total peaking background ($b \to c + b \to u, d, s$) is about 14% for the neutral mode. To account for uncertainties in MC background composition, we should increase this fraction by a 50% safety factor to 21%. Even with such a safety factor, it is desirable to validate our assumptions on this limited amount of *CP*-violation observed in the peaking background on real data, considering possible effects from both $b \to c$ and $b \to u, d, s$, without assumptions that $b \to u, d, s$ background is fully *CP*-conserving. To do so, we have slightly modified the fit model and made the *CP*V parameters $\mathcal{S}_f$ and $\mathcal{A}_f$ of a fraction f_{CP} of the peaking background floating. We then performed the full fit in the $M(K_S^0 K^+ \pi^-)$ sideband. For the fit to be stable, one must fix the *CP*V parameters of the signal, which is leaking to the sideband. We fix $\mathcal{A}_{CP}$ to zero and perform the fit in steps of fixed values for assumed $\mathcal{S}_{CP}$ for the signal and the *CP*-violating peaking background fraction f_{CP}.

The result of this study can be seen in Fig. 7.7. In the bottom two and a half rows corresponding to $f_{CP} \leq 0.2$, the fitter is no anymore stable, but we can do some estimates based on its consistent behavior in the remaining region. The second row shows how the estimated $\mathcal{S}_f$ of the peaking background is compensating the assumed signal $\mathcal{S}_{CP}$ as expected. Assuming all the peaking background is *CP*-violating with some effective $\mathcal{S}_f$ and $\mathcal{A}_f$, the *CP*-violation is consistent with zero. With a lower f_{CP}, the estimated deviation from *CP*-
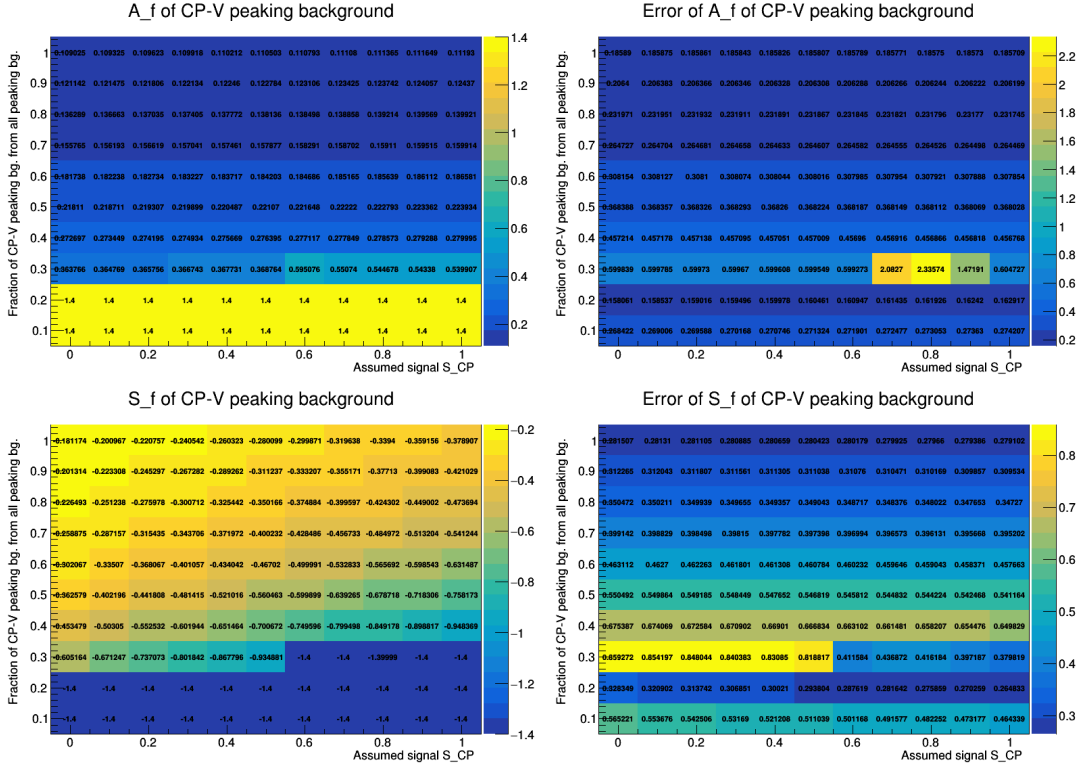
Figure 7.7: Study of CP violation in the $M(K_S^0 K^+ \pi^-)$ sideband for the neutral mode. The effective CP-violation parameters $\mathcal{A}_f$ (top row) and $\mathcal{S}_f$ (bottom row) are determined by the full 4D fit in the $M(K_S^0 K^+ \pi^-)$ sideband as a function of assumed signal $\mathcal{S}_{CP}$ and fraction of the CP-violating peaking background to the total peaking background. The determined values are shown in the left column, with their statistical error estimates in the right column.

conserving values gets larger, but it nowhere reaches the extremal values and does not get statistically significant.

Therefore, also the data suggest that even if significant CP-violation contributions are present in the peaking background, these should be sufficiently small. Fixing $\mathcal{S}_f$ and $\mathcal{A}_f$ to -1 and +1 for about 21% of the peaking background, as suggested by the MC study, then would approximately correspond (at most) to the fitter estimates extrapolated to the region of $f_{CP} = 0.2$.

If the SM expectation ($\mathcal{S}_{CP} = 0.7$) is assumed for the signal, we can give a more concrete estimate. The largest predicted CP-violation occurs for $f_{CP} = 0.4$. For lower values of f_{CP}, the fit is no anymore able to reliably predict CPV parameters of the peaking background. For this value, one can extract the effective CP-violation of the fraction f_{CP} of the peaking background as

$$
\begin{aligned}
\mathcal{S}_{f,f_{CP}=0.4} &= -0.80 \pm 0.66 \\
\mathcal{A}_{f,f_{CP}=0.4} &= +0.28 \pm 0.46
\end{aligned}
\tag{7.21}
$$

If we assume all peaking background is CP-violating, the effective values are

$$
\begin{aligned}
\mathcal{S}_{f,f_{CP}=1} &= -0.32 \pm 0.28 \\
\mathcal{A}_{f,f_{CP}=1} &= +0.11 \pm 0.19
\end{aligned}
\tag{7.22}
$$

We can see that a negative value, slightly larger than the estimated statistical error, is extracted for $\mathcal{S}_f$ in both cases. The estimated $\mathcal{A}_f$ is consistent with zero, and the statistical error is almost twice larger than the estimated parameter value. We take the estimated value or the error, whatever is larger in absolute value, to be conservative. We will then use these estimates to sample the *CP*-violating peaking background for a toy MC study to estimate the expected systematic effect on the measured *CP*V parameters of the signal.

7.3.4   Ensemble Tests with Toy MC

An important family of validations of the fitter is based on the generation of testing datasets and performing the final fit to check its statistical properties (fit biases, error estimation, and stability of the fit). These datasets can be generated directly using the data model and PDFs for backgrounds and the signal or sampled from the results of the full GSIM simulation. An important remark is that our data model does not contain all information to generate the data samples. It is missing the prescription on how to sample event-dependent variables needed to parametrize the Δt resolution functions for signal and background. These variables need to be sampled from distributions observed in simulation or real data. The statistics generated for each pseudo-experiment follow a Poisson distribution with a mean equal to the expected signal and background yields from the 2D data fit. Also all other fixed parameters are set to their expected values.

PDF-based Ensemble Tests

The PDF-based ensemble tests serve as a consistency check of the fit procedure, and thus no significant systematics is expected if the fitter works correctly. One of the main goals is to confirm that the fit can correctly predict *CP*-violation parameters in the presence of other floating nuisance parameters, from signal and background yields to continuum background shape. In particular, the fraction of the peaking background to signal is expected to cause statistical issues, especially for the SVD1 experiment, where the peaking background yield is very small (about 20 events).

We have generated 1000 pseudo-experiments from the full 4D data model. The distributions of the fitted parameter values, their pulls, and estimated statistical errors are shown in Fig. 7.8. The pulls are defined as the difference between the true value used for generation and the determined parameter value divided by the estimated error. As can be seen from the pull distributions, no statistically significant biases are observed, and the widths are consistent with unity. Also, the estimated errors agree very well with the variance of the determined parameter values. The only exception is the fraction of peaking background to signal with about -0.14 σ bias. For the SVD1 experiment, this parameter is consistent with being both zero and one, given its large error. The fitter is hitting the preset parameter limits in some cases, and the distributions are non-Gaussian. These effects arise due to the small peaking background yield in the SVD1 experiment. But as can be seen, they do not cause any issues in the determination of the parameters of interest (*CP*-violation parameters and signal yields) and thus are acceptable. In addition, with these parameters floating, the correlations to the parameters of interest are kept in the solution and should properly inflate their

Figure 7.8: Ensemble test with PDF-based toys. The distributions of the estimated parameter values, their pulls, and estimated errors are shown for all floating parameters. First row: $\mathcal{S}_{CP}$ and $\mathcal{A}_{CP}$. The following rows show parameters different for SVD1 (left three columns) and SVD2 experiment (right three columns)—signal yields, background yields, the fraction of peaking background to signal, and linear slopes of the continuum background in ΔE and M. Each distribution is fitted with a single Gaussian in the displayed range, and its mean and sigma are reported. For estimated parameter values, the true generated value is shown.

statistical errors. We will consider biases from this study as **fit bias**. Where the bias is smaller than its statistical error, this error will be taken as fit bias instead. The corresponding values are 0.008 and 0.004 fo $\mathcal{S}_{CP}$ and $\mathcal{A}_{CP}$, respectively. As the biases in the GSIM-based toys (see below) are larger and should already contain these simple fit biases, only the larger estimates will be used for the final result to avoid double-counting of the systematic uncertainties.

Moreover, as the control mode is used to extract the data/MC calibration factors for the signal mode, we have also checked the fitter behavior in PDF-based ensemble tests. However, as there is about three times more data and many more floating parameters, the control mode fit is much slower, and thus we generated only 100 pseudo-experiments. We have observed no statistically significant biases or non-unit pulls for any of the floating parameters.

Fit Linearity

Fit linearity shows whether the model can reliably estimate the parameters of interest in their full possible range. PDF-based toy MC is used for this study as the generation of new signal MC data samples with different CPV parameters would be too time-consuming. The method used to evaluate the nominal fit bias in the previous section is used to estimate the bias for each set of 1000 pseudo-experiments. From the mean value of the Gaussian fit, the true (generated) values are subtracted. As an error, the estimated fit error of the mean (from 1000 pseudo-experiments) is used. This procedure is repeated for a range of different generated values for $\mathcal{S}_{CP}$ and $\mathcal{A}_{CP}$. Significant deviations from a straight line with an intercept at zero indicate deviations from fit linearity.

Fig. 7.9 shows the fit linearity for varied $\mathcal{S}_{CP}$. A slight trend can be observed in the left plot (but not statistically significant), while the right plot for $\mathcal{A}_{CP}$ (generated in all cases using zero value) shows no visible slope or deviation from an intercept at zero. The small fit bias observed in the nominal fit bias study is clearly present in the figures for $\mathcal{S}_{CP}$, but no additional systematics is derived from this study as the nominal fit bias (0.008) already includes this effect.
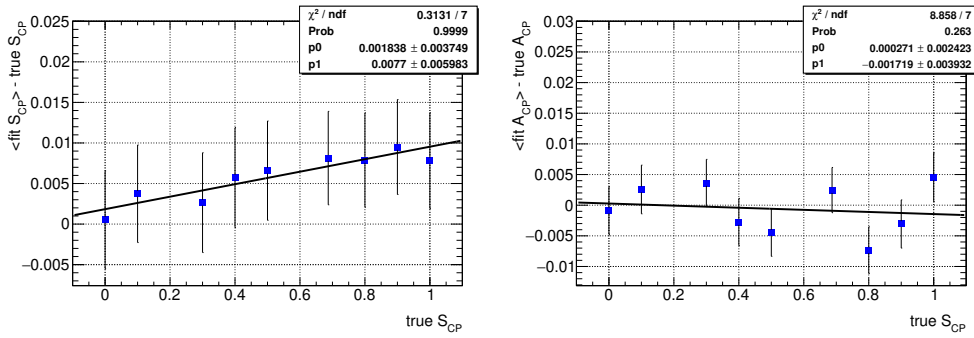
Figure 7.9: Linearity test for CPV parameters when $\mathcal{S}_{CP}$ is varied.

Fig. 7.10 shows the fit linearity for varied $\mathcal{A}_{CP}$. The generated $\mathcal{S}_{CP}$ was set to 0.687. No statistically significant deviations from fit linearity are observed, and the slight fit biases are consistent with the nominal (PDF-based) fit bias study (0.004). In conclusion, the observed fit linearity is good, and possible corresponding systematic errors are already taken into account by the nominal fit bias study.
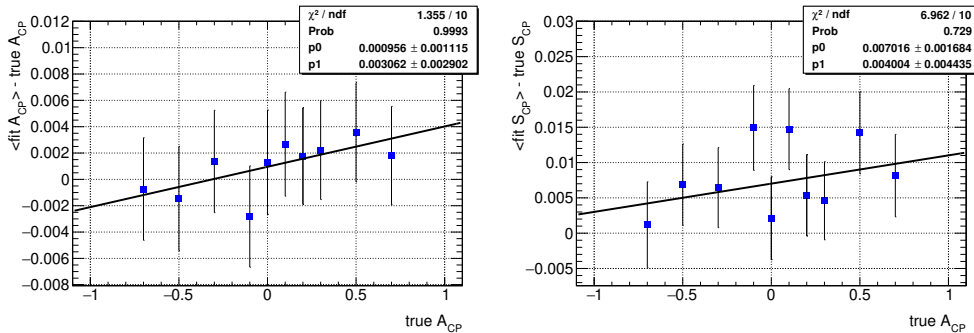
Figure 7.10: Linearity test for CPV parameters when $\mathcal{A}_{CP}$ is varied.

GSIM-based Ensemble Tests

The GSIM-based ensemble tests serve as a realistic test of the fitter. While the backgrounds are sampled from their PDFs as in the previous PDF-based test, the signal events are sampled directly from the large simulated signal MC sample[5]. The event-based variables for the peaking background are sampled from the MC signal (as the signal itself), which is the most realistic option.

Therefore GSIM-based tests check for systematic effects stemming from the approximation of the true (at least in MC) signal shape in ΔE, M and the true Δt resolution function. In $\Delta E \times M$, we have, for example, only an empirical model for the signal correlation, which is not perfect and could result in overestimation or underestimation of the true signal yield. The Δt resolution function, being a mix of empirical and analytical components and corrections, is based on a number of approximations. The impact of these approximations can be studied by fitting large-statistics MC samples (see Sec. 7.1.3 for signal-only fits) or with GSIM-based ensemble tests, which also validate the effects of the backgrounds on the full fit.

We have again generated 1000 pseudo-experiments and performed the same study as for PDF-based ensemble tests. The results are shown in Fig. 7.11. As expected, the observed biases are larger and statistically significant for CPV parameters and signal yields with. We attribute these biases to approximations in the signal and resolution function model (and partially to event-dependent variables, discussed in more detail in Sec. 7.3.5). The most significant bias is observed for $\mathcal{S}_{CP}$, about $+0.25 \; \sigma$. Also, for $\mathcal{A}_{CP}$ a smaller $+0.14 \; \sigma$ bias is observed. For further discussion, the absolute values of the biases, $+0.0341$ for $\mathcal{S}_{CP}$ and $+0.016$ for $\mathcal{A}_{CP}$ are more relevant as the predicted errors from the GSIM simulation are smaller than expected for real data due to the resolution function being too optimistic in MC.

These CPV parameter biases might be consistent with those observed in the fit to the large signal simulated signal sample, see Table 7.1. The statistical errors are still large for a conclusive result, but a general trend for a slight bias towards positive values of the CPV parameters seems to be observed consistently. These biases will be taken as systematic errors (we do not make any corrections), but we will first try to decouple the Punzi effect in the next section.

7.3.5 Systematics due to Punzi Effect

We believe the Punzi effect [63] is one of the main arguments against the simultaneous determination of signal and background yields together with the CP-violation parameters in the presence of event-dependent resolution function(s). If the distribution of event-dependent variables entering the likelihood is significantly different for signal and background, biases might be introduced. This would be truly dangerous if only the Δt dimension would be used to extract both yields and CP-violation parameters. However, in our case, the yields are constrained by the additional dimensions. This should, to some degree, reduce the possible bias due to the Punzi effect. For this hypothesis, the comparison of the

---

[5]While the same event is allowed to be taken multiple times, this is very unlikely due to the high statistics of the signal MC sample.
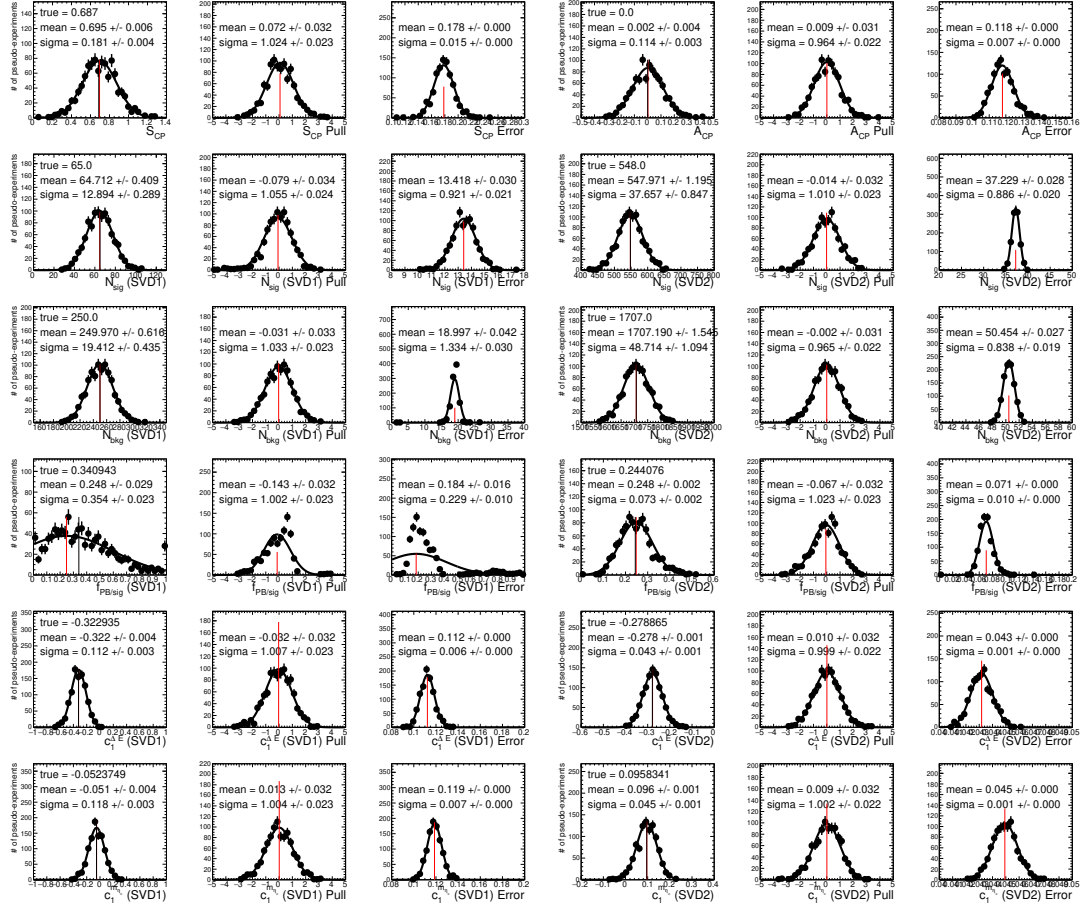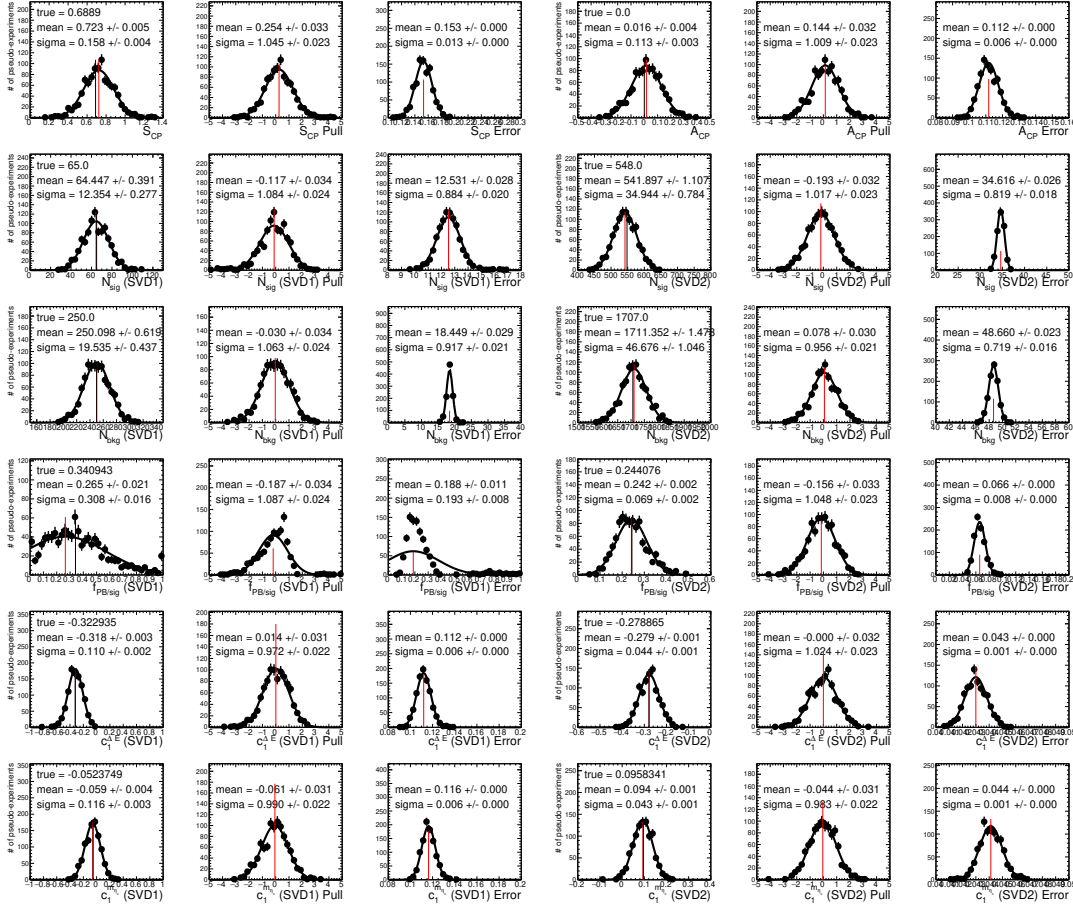
Figure 7.11: Ensemble test with GSIM-based toys. The distributions of the estimated parameter values, their pulls, and estimated errors are shown for all floating parameters. First row: $\mathcal{S}_{CP}$ and $\mathcal{A}_{CP}$. The following rows show parameters different for SVD1 (left three columns) and SVD2 experiment (right three columns)—signal yields, background yields, the fraction of peaking background to signal, and linear slopes of the continuum background in ΔE and M. Each distribution is fitted with a single Gaussian in the displayed range, and its mean and sigma are reported. For estimated parameter values, the true generated value is shown.

control mode fit in 2D versus 4D served as a cross-check. The differences in the yields and peaking background fractions when the time dimension is included are much smaller than their statistical errors.

The proper way to treat the Punzi effect is to separately model the event-dependent variables' distributions for the signal and background components. The likelihood function is then extended, treating original signal and background PDFs as conditional PDFs depending on the event variables, multiplied by the proper PDFs of these variables. However, there is a considerable number of such variables for the Belle resolution function. Even if included in the model, there will be uncertainties in the true distribution of the event variables in the data. Therefore we would anyway need to evaluate systematics stemming from uncertainties in the proper model of the Punzi constraint terms. In addition, our full fit is already quite time-consuming, and the addition of so many additional

PDFs to be evaluated for each event would make it too slow.

Fortunately, ensemble tests allow quantifying the effect of neglecting the Punzi terms in the likelihood. These terms will factorize from the likelihood if their PDFs are the same for all signal and background components. We can generate an artificial toy dataset, where we sample all event variables from the same source. The technically easiest option, which also seems the most realistic, is to use GSIM-based toy MC, where event variables for all backgrounds are sampled from the MC signal. We can then compare the bias in the *CP*-violation parameters to the nominal case to estimate the systematic error due to the Punzi effect.
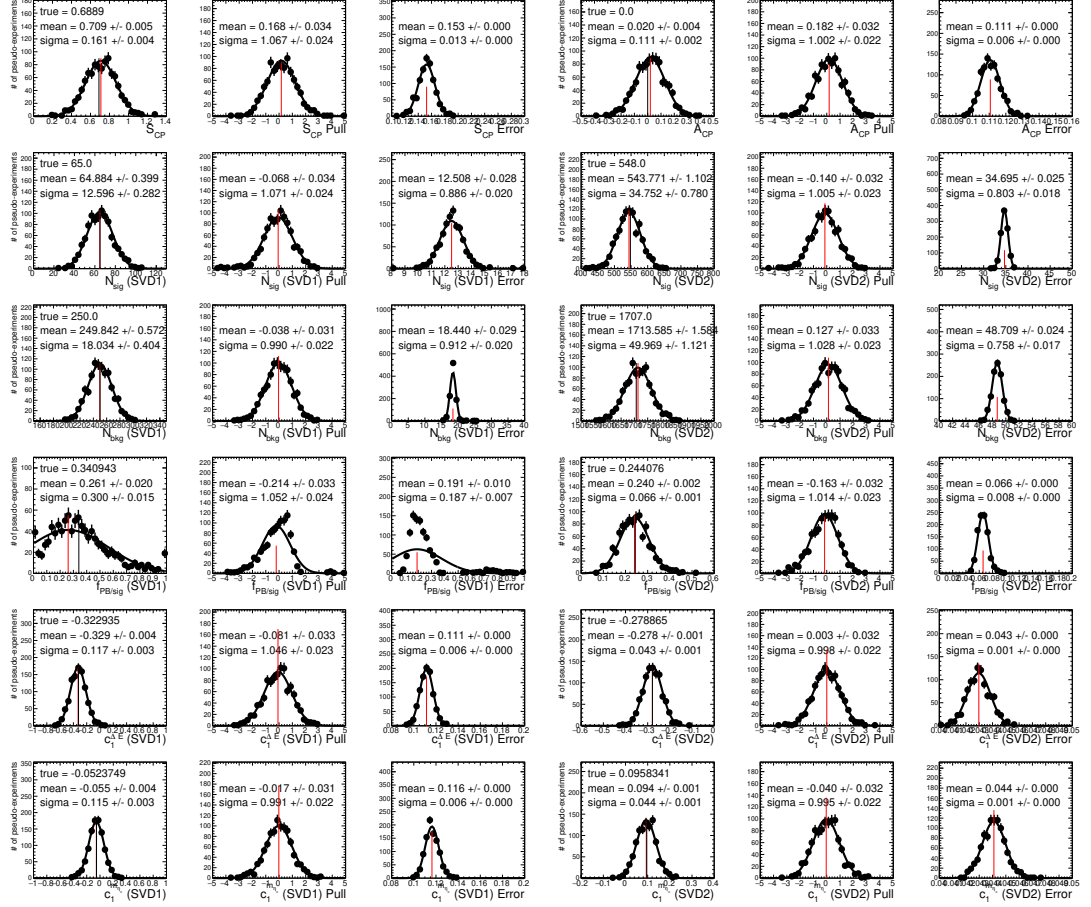
Figure 7.12: Ensemble test with artificial GSIM-based toys with all event-based variables sampled from signal MC to evaluate the Punzi effect. The distributions of the estimated parameter values, their pulls, and estimated errors are shown for all floating parameters. First row: $\mathcal{S}_{CP}$ and $\mathcal{A}_{CP}$. The following rows show parameters different for SVD1 (left three columns) and SVD2 experiment (right three columns)—signal yields, background yields, the fraction of peaking background to signal, and linear slopes of the continuum background in ΔE and M. Each distribution is fitted with a single Gaussian in the displayed range, and its mean and sigma are reported. For estimated parameter values, the true generated value is shown.

The results of the ensemble test with this artificial dataset are shown in Fig. 7.12. As can be seen, the $\mathcal{S}_{CP}$ bias of $+0.020\pm0.005$ is significantly smaller than $+0.034\pm0.005$ in our nominal fit case. The bias on $\mathcal{A}_{CP}$ gets slightly larger

by $+0.004\pm0.004$, but not significantly. This, together with a set of smaller tests, suggests that the observed biases are to a large degree related to the event-dependent resolution function and neglected distributions of the event-dependent variables in our data model (with significant backgrounds). The candidates for the most problematic ones are likely σ_z or $\cos\theta_B$. We initially suspected these biases could be reduced by limiting the amount of the (dominant) continuum background. This was not confirmed on toy simulations—the bias is almost unchanged when the toy MC for the neutral mode is repeated with updated signal and background yields corresponding to $\mathcal{CS}_{\mathrm{cut}}$ equal to 0.5 or 0.8 (determined from a special 2D fit of the neutral mode). In addition, at tighter cut values, the estimated statistical error grows considerably (to more than 0.2 for $\mathcal{S}_{CP}$ for $\mathcal{CS}_{\mathrm{cut}} = 0.8$) due to the reduced signal efficiency. Therefore we will consider the difference (in quadrature) of the bias from the nominal GSIM result and this special dataset as another source of systematic error due to the Punzi effect. The quadratic sum of the bias due to the Punzi effect and the 'pure' fit bias equals the maximal observed bias in our studies (in the nominal GSIM-toys).

7.3.6 Systematics due to *CP*-Violating Peaking Background

In Sections 5.4 and 7.3.3, we have estimated the possible amount of *CP*-violating contributions to the peaking background. This estimate can be done in different ways, yielding slightly different results. Our small study on data suggests setting at most 30% of the peaking background as potentially *CP*-violating (with extremal values) or to use the *CP*-violation parameters (or their errors) and the fraction extracted from the scan of the $M(K_S^0 K^+\pi^-)$ sideband.

Similar to study performed in Section 7.3.3, we modify the PDF for the peaking background such that a fraction f_{CP} of it is *CP*-violating with parameters $\mathcal{S}_f, \mathcal{A}_f$. We test the extremal values $\{-1,+1\}$. We also test two values for f_{CP}: 21% (MC estimate with safety limit) and 30% (maximum reasonable value, which seems rejected by data). In addition, the two sets of values extracted from the sideband data in Eqs. 7.21 and 7.22 are included in our study. We then generate 1000 pseudo-experiments using GSIM-based toys and fit them with the nominal fit model, which assumes $f_{CP} = 0$ (all peaking background *CP*-conserving).

The results are shown in Fig. 7.13. The impact of *CP*-violation in the peaking background is more threatening for $\mathcal{S}_{CP}$, where the maximal bias of 0.046 is reached for the extremal MC estimates. The largest deviation of 0.050 is observed for the configuration $f_{CP} = 0.4, \mathcal{A}_f = 0.28, \mathcal{S}_f = -0.8$, extracted from the sideband data. For $\mathcal{A}_{CP}$, the largest observed difference to the nominal fit model (GSIM) bias is 0.019 (from MC estimates). We will use the largest of the observed biases as the corresponding systematics. These values are the expected dominant systematic errors of our measurement for $\mathcal{S}_{CP}$.

7.3.7 Validation with Randomized Flavor

Before performing the final *CP* fit in the neutral mode, an additional cross-check can be done by randomization of the determined neutral B_{tag} meson flavor. This
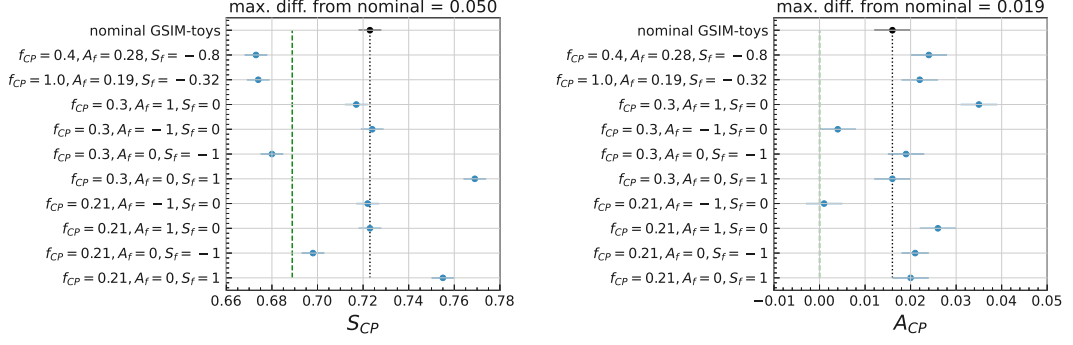
Figure 7.13: Systematics due to *CP*-violating background. The nominal result of the GSIM-based toys is shown in black. Various values for a fraction of the *CP*-violating background f_{CP} and its effective *CP*-violation parameters $\mathcal{S}_f$ and $\mathcal{A}_f$ are used to generate special toy datasets, fitted with the nominal fit model.

allows us to extract B^0 lifetime τ_{B^0} with fixed $\mathcal{A}_{CP} = 0$, $\mathcal{S}_{CP} = 0$. We get

$$\tau_{B^0} = (1.577 \pm 0.079) \text{ ps,} \tag{7.23}$$

which is in agreement with the world average. In addition, with a fixed $\tau_{B^0} = 1.519$ ps, we get the values

$$\begin{aligned} \mathcal{S}_{CP} &= -0.277 \pm 0.171 \\ \mathcal{A}_{CP} &= -0.050 \pm 0.117, \end{aligned} \tag{7.24}$$

which are statistically consistent with zero, as expected for the fit with randomized flavor information.

We have also checked the yields and other parameters are entirely consistent (within $\ll 1\sigma$) with the predictions of the 2D fit. The only slight exception is the peaking background yield for the SVD1 experiment, which is, however, still compatible within 1σ. These results (except for *CP*V parameters) differ only negligibly from the final estimates after un-blinding, discussed in the next chapter.

133

CP-Violation Measurement

In this chapter, we present the measurement of the *CP*-violation parameters in $B^0 \to \eta_c K_S^0$ decays followed by $\eta_c \to K_S^0 K^+ \pi^-$ using the full Belle dataset of 772×10^6 of $B\overline{B}$ pairs and the roughly 600 expected signal events.

We have followed the standard Belle procedures for blind analysis. All cross-checks, validations, consistency checks, and control mode fits have been performed before the *CP* fit in the signal mode. The last small modification to the analysis was done after validation with randomized flavor information, where we replaced the width of the peaking background Gaussian in ΔE by its MC estimate. In addition, all the systematic uncertainties have also been evaluated before the final fit, using partial blinding of the final values. In this approach, we added unknown offsets in the range of the expected statistical error to $\mathcal{A}_{CP}$ and $\mathcal{S}_{CP}$ to be able to compute differences due to various systematic sources without knowing the exact result. This strategy was chosen to avoid any surprises. No modifications have been done after un-blinding[1].

After all the above checks passed, the review committee agreed to un-blind the analysis. Based on ensemble and random flavor studies for the signal mode, the expected statistical significance should be comparable to BaBar, although very slightly larger for $\mathcal{S}_{CP}$. We will return to the reasoning behind the larger uncertainty in Chapter 9.

In this chapter, we will discuss the result, the sources of systematic uncertainties, and summarize the outcome in terms of its statistical significance.

8.1 Fit Results

In contrast to the control mode, the limited statistics for our signal mode force us to fix additional parameters, in particular

- the fraction $f_{q\bar{q}}^{\Delta t}$, extracted from the $M(K_S^0 K^+ \pi^-)$ sideband data, see Sec. 7.2.1,

- the background r-bin fractions $\tilde{f}_{\text{bkg}}^i$, $i = 0 \ldots 5$, extracted from the M_{bc} sideband data, see Sec. 7.1.5, and

---

[1]There is only one small exception: we have found an error in the peaking background width used for the control mode in the SVD1 experiment and a small bug in the parametrization of the mis-reconstructed signal in the 4D model. As the control mode is used to set calibration factors for the neutral mode, all relevant fits have been repeated. The SVD1 experiment contains a minor part of the entire dataset, and the effect of correcting the error resulted in a change in the SVD1 signal yield by just one single event. Effects on *CP*V parameters of interest were only on the fourth significant digit.

- the calibration factors for η_c signal shape in ΔE as well as signal r-bin calibration factors $\tilde{\eta}_i^{\mathrm{CF}}$, $i = 0 \ldots 5$, extracted from the final fit to the control mode, see Sec. 7.2.

This leaves twelve floating parameters for the final fit: five (times two) specific to each experiment (SVD1 and SVD2) configuration (N_{sig}, N_{bkg}, $f_{\frac{\mathrm{pb}}{\mathrm{sig}}}$ and background slopes $c_1^{\Delta E}$ and c_1^M) and two common physics parameters $\mathcal{S}_{CP}$ and $\mathcal{A}_{CP}$. The estimated values of the floating fit parameters in the final fit are shown in the summary Table A.5. The observed number of events is 68 ± 14 and 548 ± 37 for the SVD1 and SVD2 experiments, respectively. We also evaluate asymmetric `MINOS` errors for the CPV parameters, which yields $\mathcal{S}_{CP} = 0.588^{+0.163}_{-0.168}$ and $\mathcal{A}_{CP} = 0.161^{+0.116}_{-0.117}$. The standard practice at Belle is to report asymmetric errors if they differ by more than 10%. If this is not the case, only the larger uncertainty is reported as symmetric. Following this practice, the result with statistical errors only is

$$\mathcal{S}_{CP} = 0.588 \pm 0.168$$
$$\mathcal{A}_{CP} = 0.161 \pm 0.117. \tag{8.1}$$

We also determine the branching fraction as

$$\mathcal{B} = \frac{N_{\mathrm{sig}}^1/\epsilon_1 + N_{\mathrm{sig}}^2/\epsilon_2}{N_{B\bar{B}}}, \tag{8.2}$$

where the MC signal efficiencies ϵ_i are corrected by factors accounting for data/MC difference in the K_S^0 reconstruction and detection efficiency, pion ID efficiency, and kaon ID efficiency (see Sec. 8.3). The result is

$$\mathcal{B} = (9.79 \pm 0.64) \times 10^{-6} \tag{8.3}$$

concerning statistical uncertainties only.

We prefer to determine the branching fraction in this final simultaneous fit, albeit this means additional systematic sources will appear. In fact, the simultaneous fit predicts slightly smaller statistical errors as additional information is provided by the time and flavor dimensions. Moreover, the estimated signal yields from the 2D and the 4D fit are almost identical. However, from the study of interference with the non-resonant peaking background, we already know that the total uncertainty will be entirely dominated by the interference (23% for signal yield in neutral mode as estimated in Sec. 6.3.3), which we will quote in our final result separately.

The fit and data projections are shown in Fig. 8.1, where SVD1 and SVD2 experiment data and models are combined. The distributions of r-bins are shown in Fig. 8.2. We project only a limited, signal-enhanced region, defined already in Eq. 7.16. For projections in the full signal region, see Fig. C.2.

After this projection, the estimated background PDF is subtracted from the data histograms for each flavor (q) of the tagged B_{tag}. The upper plot in Fig. 8.3 compares these background-subtracted distributions to the signal PDF projection. The bottom plot shows the raw asymmetry directly, without any background subtractions, formed as $(N^+ - N^-)/(N^+ + N^-)$, where the $N^{+,-}$ is the number of observed events in each flavor category, with Poisson errors. The same asymmetry operation is applied to the histograms accumulated from the total model PDFs.
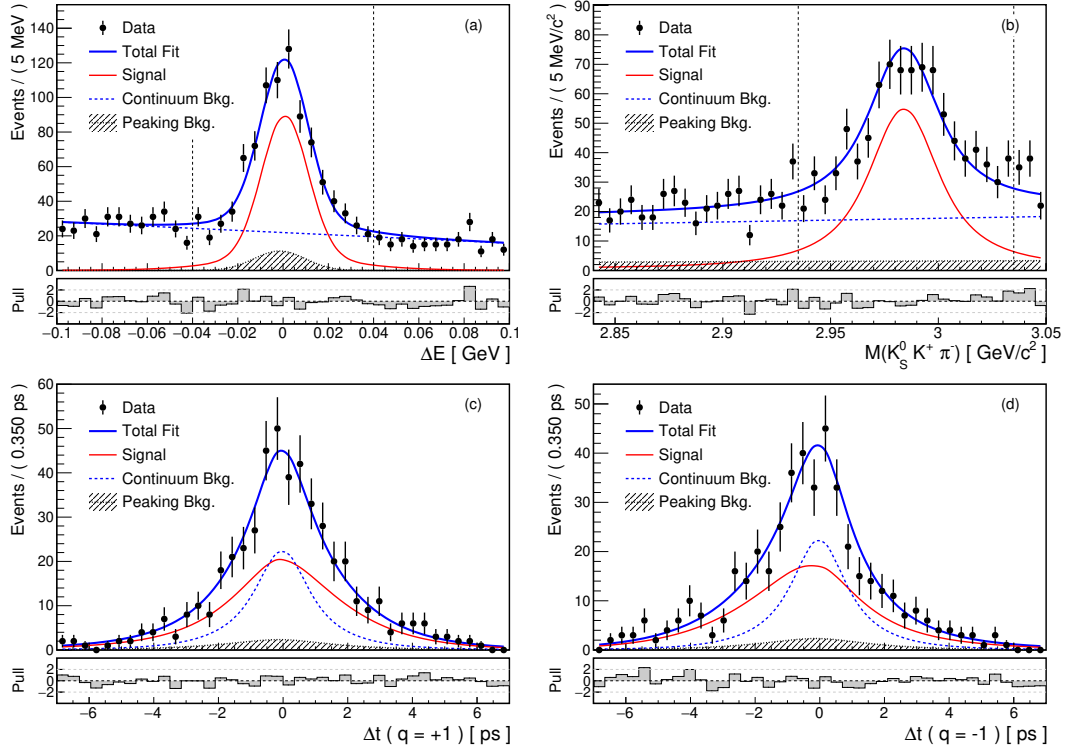
Figure 8.1: Projections in ΔE (a), $M(K_S^0 K^+ \pi^-)$ (b) and proper time difference Δt for B^0 (c) and $\overline{B}^0$-tagged (d) events in the signal-enhanced region (the distributions in Δt are for events in the dotted window in ΔE and M). Curves show the fit model and its components, and points represent the data. SVD1 and SVD2 experiment data and model are combined. Continuum background component includes non-peaking $B\overline{B}$ background.
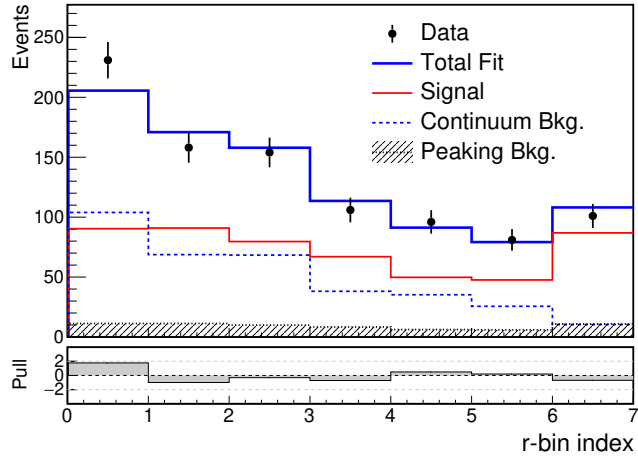
Figure 8.2: Projections of r-bin distributions in the signal-enhanced region. Lines show the fit model and its components, and points represent the data. SVD1 and SVD2 experiment data and model are combined. Continuum background component includes non-peaking $B\overline{B}$ background.

8.2 Consistency Checks

The review committee of the analysis proposed two additional specific validations. First, we check that background suppression performance, optimized on contin-
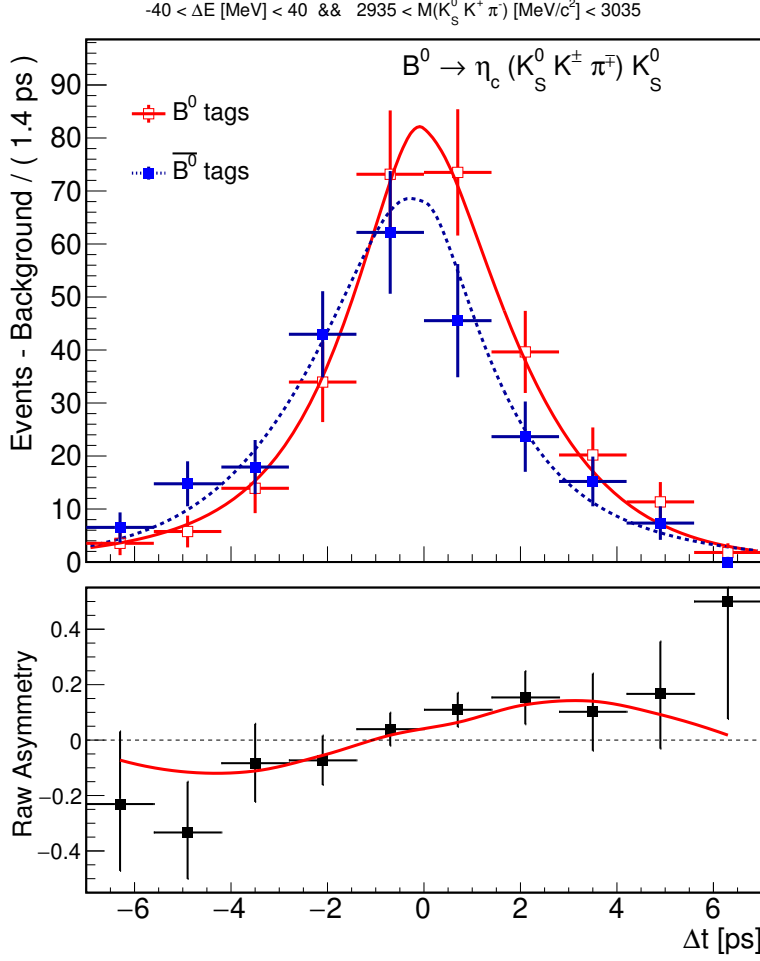
Figure 8.3: Projections of background-subtracted distributions of the estimated number of signal B^0 and $\overline{B}^0$-tagged events (top) and raw asymmetry $(N^+ - N^-)/(N^+ + N^-)$ of the number of total events (bottom). The distribution is integrated over all r-bins. Only the signal-enhanced region is projected. Curves show the fit model; points represent the data. SVD1 and SVD2 experiment data and model are combined.

uum $q\bar{q}$ MC, is compatible with the data. For this test, 60 fb^{-1} of off-resonance data are used and reconstructed with $\mathcal{CS}_{\text{cut}} = 0$ and $\mathcal{CS}_{\text{cut}} = 0.2$. From the number of passing events without and with the cut, the obtained background rejection is $(70 \pm 8)\%$ and $(63 \pm 4)\%$ for SVD1 and SVD2, respectively. These values seem slightly lower but statistically compatible with the MC background rejection rate of 74.9% and 70.0% for SVD1 and SVD2, respectively.

Second, we test for CPV effects stemming from interference with the non-resonant background. Around the η_c resonance peak, the imaginary part (phase) of the wave function changes as a function of the η_c resonant mass. This phase difference can act similarly to the strong phase and reveal mass-dependent CP-asymmetry. Based on the recommendation from the review committee, this effect is studied by redefinition of the CPV parameters as follows:

- We modify $\mathcal{S}_{CP} \to 0.59 - \Delta\mathcal{S}$ for $M < m_{\eta_c}$ and $\mathcal{S}_{CP} \to 0.59 + \Delta\mathcal{S}$ for $M > m_{\eta_c}$ and obtain $\Delta\mathcal{S} = -0.14 \pm 0.16$ (and $\mathcal{A} = 0.17 \pm 0.12$). The

hard-coded value (0.59) is the nominal fit result.

- We modify $\mathcal{S}_{CP} \to 0.7 - \Delta\mathcal{S}$ for $M < m_{\eta_c}$ and $\mathcal{S}_{CP} \to 0.7 + \Delta\mathcal{S}$ for $M > m_{\eta_c}$ and obtain $\Delta\mathcal{S} = -0.13 \pm 0.15$ (and $\mathcal{A} = 0.16 \pm 0.12$). Here world average (0.699) is used for the central (average) value.

- We modify $\mathcal{A}_{CP} \to -\Delta\mathcal{A}$ for $M < m_{\eta_c}$ and $\mathcal{A}_{CP} \to +\Delta\mathcal{A}$ for $M > m_{\eta_c}$ and obtain $\Delta\mathcal{A} = 0.04 \pm 0.12$ (and $\mathcal{S} = 0.60 \pm 0.17$). Here SM expectation (0.) is used for the central (average) value.

- We modify $\mathcal{A}_{CP} \to 0.16 - \Delta\mathcal{A}$ for $M < m_{\eta_c}$ and $\mathcal{A}_{CP} \to 0.16 + \Delta\mathcal{A}$ for $M > m_{\eta_c}$ and obtain $\Delta\mathcal{A} = 0.03 \pm 0.12$ (and $\mathcal{S} = 0.59 \pm 0.17$). The hard-coded value (0.16) is the nominal fit result.

Thus no statistically significant effect is observed in any of the cases.

Moreover, we again verified the B^0 lifetime estimation (with CPV parameters left floating) and obtained $\tau_{B^0} = (1.576 \pm 0.079)$ ps, consistent with the world average.

8.3 Systematic Uncertainties

This section gives an overview of all systematic uncertainties and some additional systematic cross-checks. While statistical errors dominate our measurements, the systematic uncertainties need to be evaluated carefully, primarily as we use quite a different fitting strategy than previous studies. We use the simultaneous determination of yields and CPV parameters in a multi-dimensional fit also to determine the branching fraction. This will generate some additional systematic errors with respect to a pure 2D fit. However, the full 4D fit retrieves a slightly smaller statistical error for the yields, and thus we prefer it. In contrast to previous analyses, there are no systematic errors associated to signal and background yields, as these are floating parameters, and the errors are properly (including correlations) included in the estimated statistical uncertainty.

In general, we follow the standard or recommended Belle procedures to evaluate most of the systematic errors. The input parameters of the analysis are varied by $\pm 1\sigma$ for values determined from data or external sources (like PDG or other studies) or $\pm 2\sigma$ for values coming from MC (like signal model parameters in $\Delta E \times M$). The maximal deviation from the nominal fit value is conservatively taken as a systematic error from each input parameter. Errors from all input parameters are summed in quadrature. For each category of the systematic errors, we evaluate asymmetric deviations and first sum in quadrature each side in the category. The maximal deviation to either side is used as the systematic uncertainty for each category. The errors in each category are discussed below.

Whenever a cut is changed, leading to different events entering the likelihood, the MC signal efficiencies are estimated again for such a configuration to estimate the branching fraction. In such cases, where signal and background yields change (sometimes significantly), one should not mistake statistical fluctuations with actual systematic effects. Nevertheless, effects from several such cut variations are used as this is a standard for Belle TDCPV analyses. In cases where the derived systematics is small, we usually take it 'as is.' If it is significant or even

becomes almost a major systematics, we do more studies. One should differentiate between systematic errors (where an effect is expected) and checks (where no effect is expected a priori). Changed analysis cuts usually fall into the category of checks as this often suffers from statistical fluctuations. If these are smaller than the statistical error, the check should pass without any assigned systematic error. This will be the case for cuts on M_{bc} and the continuum suppression variable.

The summary of the systematic uncertainties by category and their total sum in quadrature is shown in Table 8.1. Below, we discuss them one by one separately in detail.

Source	$\delta\mathcal{S}$ $[10^{-2}]$	$\delta\mathcal{A}$ $[10^{-2}]$	$\delta\mathcal{B}$ $[10^{-6}]$ (%)
SVD misalignment	0.24	0.41	–
Δz bias	0.39	0.50	–
Tag-side interference	0.70	3.30	–
Peaking bkg. CPV	5.00	1.90	–
η_c line shape	2.30	0.80	–
Pure fit bias	2.00	1.60	0.066 (0.67 %)
Punzi effect	2.70	1.20	0.080 (0.82 %)
Physics parameters	0.16	0.51	0.094 (0.97 %)
Track helix errors	0.12	0.35	0.006 (0.06 %)
$|\Delta t|$ range	0.04	0.12	0.041 (0.42 %)
Vertex quality	1.66	1.05	0.099 (1.01 %)
J/ψ excl. range	0.44	0.53	0.044 (0.45 %)
K^0_S efficiency	–	–	0.062 (0.64 %)
$K^\pm$ efficiency	–	–	0.188 (1.92 %)
$\pi^\pm$ efficiency	–	–	0.099 (1.01 %)
Tracking efficiency	–	–	0.069 (0.70 %)
TagV selections	0.32	0.13	–
Number of $B\overline{B}$	–	–	0.134 (1.37 %)
MC efficiency	–	–	0.049 (0.50 %)
Tagging w_l	0.08	0.21	0.001 (0.01 %)
Tagging Δw_l	0.27	0.06	0.001 (0.01 %)
$\mathcal{R}_{B\overline{B}}$	1.50	3.38	0.046 (0.47 %)
IP constraint	0.68	0.87	0.018 (0.19 %)
Fit model parameters	0.43	1.64	0.242 (2.47 %)
Total systematics	6.95	6.02	0.410 (4.19 %)
Statistical unc.	16.53	11.65	0.64 (6.6 %)

Table 8.1: Summary of the systematic uncertainties for the CPV parameters and branching fraction, their total sum in quadrature, and the estimated statistical error for reference and comparison.

Tag-Side Interference

The tag-side final state may be reachable by multiple diagrams, which can interfere. This is especially the case for kaon tags, where the dominant decay mode for $\overline{B}^0$ (via $b \to c\overline{u}d$, like $\overline{B}^0 \to D^+\pi^-$), which is CKM-favored, can be be also reached by a CKM-suppressed $\overline{b} \to \overline{u}c\overline{d}$ decay of B^0 [64]. The relative strength of the two amplitudes is approximately $r' = |(V_{ub}^\star V_{cd})/(V_{cb}V_{ud}^\star)| \sim 0.02$. We follow the standard Belle II procedure to evaluate the corresponding systematic errors [65], using inputs from an independent semileptonic $B^0 \to D^{\star-}\ell^+\nu$ control sample analysis. The latest update of this analysis yielded [66] $2r'\sin(2\phi_1 + \phi_3 + \delta') = +0.0096 \pm 0.0073$, and $2r'\sin(2\phi_1 + \phi_3 - \delta') = -0.0067 \pm 0.0073$, which can be used to modify the signal PDF, such that interference effect is included.

We generate 1000 pseudo-experiments using PDF-based toy MC using a signal PDF with interference included and fit it with the nominal model. The bias with respect to the nominal result for CPV parameters is used as a measure of the systematic error. We generated several different combinations of the PDFs with interference, varying the input parameters by $\pm 1\sigma$. The average of the absolute value of the biases is taken as the corresponding systematic error for $\mathcal{A}_{CP}$ as the bias is strictly positive (and has physical reasoning). For $\mathcal{S}_{CP}$, we use the maximal bias as usual. The errors are 0.007 for $\mathcal{S}_{CP}$, and 0.033 for $\mathcal{A}_{CP}$.

The systematic error for $\mathcal{A}_{CP}$ is somewhat larger than, for example, estimated by BaBar (0.014) [48]. On the other hand, this error is rather in agreement with the most recent Belle estimation from the full dataset for $J/\psi K_S$ [55] (+0.038— the maximal observed systematics for any of our toy MC tests was +0.037, lowest +0.028). Despite the tag-side interference is now the dominant systematics for $\mathcal{A}_{CP}$, it is still acceptable for the estimated statistical uncertainty of ~ 0.12.

Line Shape of the η_c Resonance

The interference of the signal and (non-resonant) background is neglected in the nominal fit. We use the results of the simplified interference study and generate PDF-based toy MC using signal shape in $M(K_S K^+\pi^-)$ as shown in Fig. 6.19 (right), but first, subtract the minimum of the PDF (and re-normalize it to the final signal region). As a result of the asymmetry in the generated signal distribution, the fit compensates for this by adjusting signal and background yields and slopes, leading to significant biases, e.g., in signal yields (around 10%) when the generated pseudo-experiments are fitted with the nominal model. Another test is performed for the toy sample where no peaking background is simulated. The maximal observed bias is used. For the CPV parameters, the corresponding systematic errors (taken as differences of the biases to nominal toy MC result) are $(-)0.0229$ and 0.001 for $\mathcal{S}_{CP}$ and $\mathcal{A}_{CP}$, respectively. The effect on $\mathcal{S}_{CP}$ is significant, as the fit trades some signal events for background events, which have a different generated CP asymmetry in $\mathcal{S}_{CP}$.

Note that the corresponding systematic error for the branching fraction will be taken from the full interference study and presented separately in the final result, as it is almost three times larger than the estimated statistical error.

Wrong Tag Fractions and Wrong Tag Fraction Differences

Wrong tag fractions and the wrong tag fractions differences, controlling the flavor dilution of the physics distribution, are taken (with their uncertainties) from dedicated Belle analyses [55], and can also be found in Ref. [40].

Fit Bias and Punzi Effect

Fit biases are evaluated using ensemble PDF- or GSIM-based tests of the final fitter. More significant biases are observed in the GSIM-based toy studies. We have studied the impact of the Punzi effect by sampling all event-dependent variables from the MC signal (also for all backgrounds). The maximal observed fit bias in these studies (coming in almost all cases from the GSIM-based toys) is then separated into the Punzi effect and a 'pure' fit bias (observed on the toys free of the Punzi effect), such that in quadrature, these sum to the maximal observed bias.

For $\mathcal{A}_{CP}$, the bias was very slightly larger (by 0.004±0.004) without the Punzi effect. Thus we simply assign the smaller of the two errors (to be summed in quadrature) to the Punzi effect and the larger one to the 'pure fit bias.' Interpretation of such a separation of these systematic effects becomes a bit questionable in this case.

Number of B Meson Pairs

The number of (charged and neutral) B-meson pairs $N_{B\overline{B}}$ only contributes to the branching fraction measurement and is determined together with its uncertainty by dedicated Belle studies [67], where the total number (SVD1 + SVD2) is $N_{B\overline{B}} = (771.581 \pm 10.566) \times 10^6$.

MC Signal Efficiency

The signal efficiency needed for branching fraction calculation is determined from a large MC signal sample. The statical errors of the determined efficiencies are used to evaluate the corresponding systematics.

Physics Parameters

Physics parameters taken from PDG are m_{η_c}, Γ_{η_c}, τ_{B^0} and Δm_d. These are varied by their respective uncertainties.

Peaking Background *CP* Violation

CP-violation in the peaking background was estimated both from MC and from a scan of a fit to the $M(K_S^0 K^+ \pi^-)$ sideband data. Multiple values are used to generate toy MC with *CP*-violating peaking background contribution and fitted with the nominal fit model. The largest observed deviation from the nominal toy results is used for the systematic estimate. For $\mathcal{A}_{CP}$, the largest bias comes from the MC estimates. For $\mathcal{S}_{CP}$, the maximal bias comes from the values obtained from the $M(K_S^0 K^+ \pi^-)$ sideband data fit.

Track Helix Errors

The track helix parameter errors estimated by the tracking are part of the event-dependent variables entering the resolution function. As these errors are not estimated perfectly (helix parameter measurements have non-unit pulls), additional smearing is applied to account for these differences. An older recommendation to switch these corrections off has been found to be too conservative, and instead, a newer method is used. This method exploits the observed correlation of the lifetime and mass difference fit results. The systematic errors are then evaluated by using different values for the lifetime and mass difference obtained in a dedicated study: $\tau_{B^0} = (1.5299 \pm 0.0029)$ ps and $\Delta m_\mathrm{d} = (0.5088 \pm 0.0019)$ ps^{-1} [66]. We also vary these inputs by their estimated errors and take the maximal observed bias as the corresponding systematics.

Time Difference Cut

The cut $|\Delta t| < 70$ ps is varied by ± 30 ps as recommended by the Belle TDCPV group.

Vertex Quality Cuts

Requirements on the vertex quality, namely the reduced χ^2 over the number of degrees of freedom $h^{\mathrm{rec,tag}} < 50$ is varied by ± 25. The cuts on the vertex z error $\sigma_z^{\mathrm{rec,tag}} < 200(500)$ μm for single (multi) track vertices are removed to estimate the systematics. We have also tried to vary the cuts by ± 100 μm. While the upward variations yield a small bias, the downward variation removes a significant amount of signal events for the multi-track vertices. This produces a much larger deviation but still less than half of the statistical error. By changing the cut only by 50 μm (or less to not affect the signal yield), the bias is much smaller, confirming it is only a statistical fluctuation. While in the $\eta_c \to p\bar{p}$ analysis [40], the variations of the cuts are used, we only use the option to remove the cuts on $\sigma_z^{\mathrm{rec,tag}}$ entirely. This is also a recommended procedure.

J/ψ Cut

The right limit on the $M(K_S^0 K^+ \pi^-)$ fit region to avoid the J/ψ peak was not considered as systematics for CPV parameters in the $\eta_c \to p\bar{p}$ analysis, only for the branching fraction. By changing the right fit window by approximately the detector resolution ± 5 MeV/c^2, the signal and background yields change only slightly. As we do not have any better idea on how to estimate the possible effect from interference with J/ψ from the data (and in MC, this is not simulated, we only set the fit window to limit the number of J/ψ events to only several ones based on generic $B\overline{B}$ MC studies), we include this systematic error for all parameters of interest.

M_{bc} and Continuum Suppression Cuts

The branching fraction, as well as the CPV parameters, should not a priory depend on the M_{bc} and continuum suppression cuts; only the background varies significantly (and signal slightly) when these cuts are changed. Thus we consider

those to be only systematic checks and the observed biases, significantly smaller than the estimated statistical errors (but otherwise somewhat large in the total systematics budget), are indications of passing these checks. Left M_{bc} signal window range was varied by approximately the corresponding detector resolution ± 2.5 MeV/c^2. To check the biases are only statistical, we also tested smaller and larger variations and observed only oscillations around the central value and not an obvious trend.

Similarly, we vary the continuum suppression cut from 0.2 by ± 0.05. We also tested a much wider range, up to 0.6. The CPV parameters and branching fractions seem to only oscillate well below statistical errors, despite large variations in yields. We also checked the estimated statical errors. From the nominal value of 0.165 for $\mathcal{S}_{CP}$, we in fact observed minimal value of 0.162 at $\mathcal{CS}_{\text{cut}} = 0.5$. This is in contrast to the charged mode, which we used to optimize the cut value to minimize the statistical error of $\mathcal{S}_{CP}$. In fact, we see a plateau from 0.2 to 0.5 of the cut value for the signal mode. However, this study was not done in its entirety. For large changes of the cut, the analysis should be, in fact, completely repeated, with less background (and also less signal). This will limit statistical precision on background parameters, namely in the Δt dimension. Also, the small change in the statistical error of 0.003 might be only a fluctuation and is not guaranteed after a fully repeated analysis. In conclusion, we do not find any good reason to change the background level and do not assign any additional systematics. We instead verified the robustness of the analysis and the predicted statistical errors and central value against the (continuum) background level.

K_S^0, Kaon ID and Pion ID Efficiency

Efficiency correction factors account for differences between K_S^0 identification and reconstruction efficiency, kaon ID efficiency, and pion ID efficiency in data and MC. These factors are evaluated using look-up tables, to which we fed transverse momenta and $\cos\theta$ of the K_S^0, K^+ or π^+ candidates. Both the kaon and pion ID efficiency calculation also requires the corresponding likelihood ratio cut used in the analysis. The correction factors, summarized in Table 8.2, are used to modify signal efficiency and thus branching fractions. Their uncertainties are used to estimate the corresponding systematic errors related to these efficiency corrections. The binned kaon and pion ID correction data are taken from Ref. [68] where the latest (2006 for SVD1 and 2010 for SVD2) corrections are used.

Category	SVD1 correction	SVD2 correction
K_S^0 efficiency	$0.985 \pm 0.004 \pm 0.006$	$0.985 \pm 0.004 \pm 0.006$
$K^\pm$ ID efficiency	1.0062 ± 0.0089	1.0304 ± 0.0214
$\pi^\pm$ ID efficiency	0.9829 ± 0.0055	0.9814 ± 0.0106

Table 8.2: Correction factors due to data/MC efficiency differences for K_S^0, $K^\pm$ and $\pi^\pm$.

Tracking Efficiency

A systematic error of 0.35 % per charged track (kaon and pion from η_c) is assigned to account for the difference between MC and data tracking efficiency [69]. No correction to the branching fraction is made.

IP Constraint

The IP tube constraint is used to improve the fitting of the B meson vertex by including the estimated interaction point position and size by the method of Lagrange multipliers added to the vertex fit. To account for the finite B meson flight length, the position in $R - \phi$ is additionally smeared in each event by a Gaussian distribution with zero mean and sigma of 21 μm. The standard procedure to estimate the related systematic errors is to vary this smearing by -10 μm and +20 μm.

As this requires repeating the reconstruction using the original Belle Analysis Software Framework (basf) with the changed smearing settings, we evaluate these systematics in a simplified but conservative way, which only uses the information stored in our n-tuples.

We modify Δz (used to determine Δt) for each event as follows:

$$\Delta z \to \Delta z' = \Delta z + 2\frac{p_t}{p_z} \times G(\mu, \sigma),$$

where p_t and p_z are the reconstructed transverse and longitudinal $B_{\rm sig}$ momenta, respectively. The factor of 2 accounts for symmetrization with $B_{\rm tag}$, which is smeared independently. Thus, in reality, the effect on the signal and tagging side should counter-act each other on average, while we conservatively take twice the effect. The correction is larger for larger transverse momenta of the signal B meson. The random smearing with mean μ and variance σ, which are supposed to act similarly as variations of the IP constraint smearing, is used to estimate the effect conservatively. In reality, the z vertex is not modified directly by the constraint, which is not exact but has an assigned uncertainty of the IP profile in $R - \phi$. Thus also, this part of the modification is conservative.

We have tried multiple combinations of (μ, σ), namely (0, 20 μm), (0, 40 μm), (0, 60 μm), (-10 μm, 41 μm), (-21 μm, 41 μm), (21 μm, 41 μm). We use the maximal deviation as the systematic error. The impact is relatively small, although individual Δz are modified by even tens of μm on average for some configurations. The reason is the nature of the asymmetry measurement, as the correction does not differentiate between the flavors.

TagV Selections (Tokyo Cut)

The selections for the B-tagging TagV algorithm [54, 62], which mainly deals with the pollution from long-lived particles or tracks with poor resolution, possibly biasing the tag vertex position, are also called a Tokyo cut. The algorithm requires the transverse impact parameter to be within 500 μm of the signal B vertex and the estimated uncertainty of the track longitudinal impact parameter within $\sigma_z < 500$ μm. The default procedure is to vary these criteria by ± 10 %. As this requires repeating the basf reconstruction, we use values obtained for the

full Belle $\sin 2\phi_1$ measurement, as these are expected to be mode-independent (because they are related to the tag side).

Resolution Function for B Mesons

The B meson resolution function $\mathcal{R}_{B\overline{B}}$ has many parameters determined in other studies, together with their estimated errors. The Belle package `tatami` [61] is used for resolution function computation and parametrization. We are using the latest set of parameters, all determined by independent high-statistics studies. We have observed some significant systematic for $\mathcal{A}_{CP}$ from the detector resolution part, which is far larger than in other studies. It is very likely that the simultaneous fit is more sensitive to variations in the resolution function. Anyway, these errors are still acceptable when compared to statistical errors. The problematic resolution function parameter is s_{rec}^1 (see [40]), which scales the linear dependence of resolution Gaussian width on the vertex quality h.

Model Parameters

All the remaining fit model parameters are varied by $\pm 1\sigma$ for values obtained from fits to data and by $\pm 2\sigma$ for values determined from fits to MC samples. The largest contribution to all parameters of interest arises due to the width of the peaking background Gaussians. In particular, when replacing the MC estimate for SVD1 with the data-driven estimate. Another notable systematics are parameters of signal shape in M, namely the fraction and width of the tail component, affecting mostly $\mathcal{A}_{CP}$ and $\mathcal{B}$.

SVD Misalignment and Δz Bias

The impact of SVD misalignment was evaluated for the full Belle $\sin 2\phi_1$ measurement [66, 55] by simulating geometries with different random local misalignments. In addition, as alignment validations revealed some biases, various global (SVD versus CDC) misalignment scenarios are produced to find a configuration resembling the observed biases. This configuration is then used to estimate the corresponding alignment systematics using simulation. As these studies are very computationally expensive and the corresponding systematic errors are strongly believed to be mode-independent, we use values from these estimations.

8.4 Summary

In summary, we obtain for $B^0 \rightarrow \eta_c K_S^0$, $\eta_c \rightarrow K_S^0 K^+ \pi^-$ using the full Belle dataset of 772×10^6 of $B\overline{B}$ events:

$$
\begin{aligned}
\mathcal{S}_{CP} &= 0.59 \pm 0.17 \pm 0.07 \\
\mathcal{A}_{CP} &= 0.16 \pm 0.12 \pm 0.06 \\
\mathcal{B} &= (9.79 \pm 0.64 \pm 0.41 \pm 2.25_{\mathrm{int}}) \times 10^{-6},
\end{aligned}
\tag{8.4}
$$

where the first uncertainty is statistical, the second systematic, and the last error on $\mathcal{B}$ accounts for the interference with the non-resonant peaking background (23%). For comparison, the SM and previous measurements predict $\mathcal{S}_{CP} = 0.70 \pm$

0.02, $\mathcal{A}_{CP} = 0.00 \pm 0.02$ [49] and $\mathcal{B} = (9.7 \pm 2.1) \times 10^{-6}$. The PDG uncertainty on the branching fraction does not include the interference effects, as it is obtained as a product of two branching fractions, see Sec. 6.2.3.

The statistical correlation coefficients for the floating fit parameters are summarized in Table 8.3. Notable is the correlation of $\mathcal{S}_{CP}$ to the peaking background fraction for the SVD2 experiment (15%), which justifies our approach of the simultaneous fit of signal and background yields. Furthermore, $\mathcal{B}$ has roughly 10% correlation to $\mathcal{S}_{CP}$ as can be estimated from the correlation to the signal yield in the SVD2 experiment (which dominates the branching fraction estimation).

Finally, to obtain the statistical significance of our result, we scan the likelihood for fixed values of $\mathcal{S}_{CP} = \mathcal{S}'$ in the range $[0, 1]$. At each point, the likelihood is minimized again, keeping $\mathcal{S}'$ fixed and floating the rest of the parameters of the final fit. The statistical significance then can be evaluated as $\sqrt{-2 \ln(\mathcal{L}_{\mathcal{S}'}/\mathcal{L}_{\max})}$, where $\mathcal{L}_{\mathcal{S}'}$ is the maximum likelihood for the given fixed value of $\mathcal{S}'$ and $\mathcal{L}_{\max}$ is the absolute likelihood maximum for the nominal estimated parameters. The null hypothesis of no CPV in $B^0 \to \eta_c K_S^0$ (for $\mathcal{S}' = 0$) is rejected at 3.4σ (99.93%) level, as can be seen in Fig. 8.4
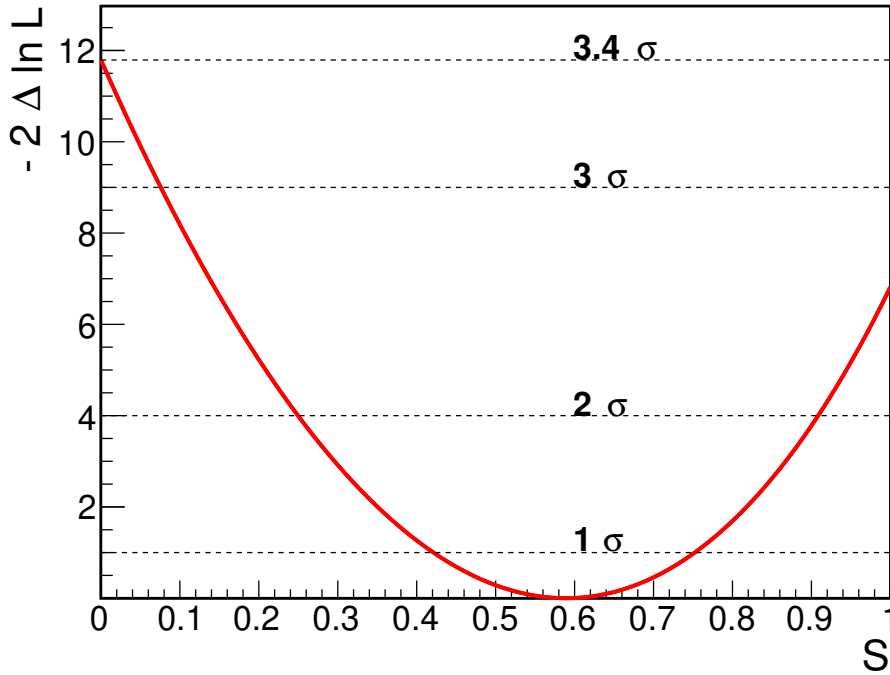
Figure 8.4: Scan of the likelihood function, maximized over all parameters except $\mathcal{S}_{CP}$ in steps of fixed $\mathcal{S}_{CP}$. Horizontal lines denote a change of the likelihood from the minimum corresponding to one, two, and three sigmas (68.27%, 95.45% and 99.73% confidence levels), as well as the observed statistical significance of our $\mathcal{S}_{CP}$ measurement. Only statistical uncertainties are considered.

Table 8.3: Statistical correlation coefficients among floating fit model parameters (in %). Parameters $c_1^{\Delta E(1,2)}$ and $c_1^{M(1,2)}$ control the background slopes in ΔE and $M(K_S^0 K^+\pi^-)$ for the SVD1 and SVD2 datasets, respectively.

	$c_1^{\Delta E(1)}$	$f_{\frac{\text{pb}}{\text{sig}}}^{(1)}$	$c_1^{M(1)}$	$N_{\text{sig}}^{(1)}$	$N_{\text{bkg}}^{(1)}$	$c_1^{\Delta E(2)}$	$f_{\frac{\text{pb}}{\text{sig}}}^{(2)}$	$c_1^{M(2)}$	$N_{\text{sig}}^{(2)}$	$N_{\text{bkg}}^{(2)}$	$\mathcal{S}_{CP}$
$f_{\frac{\text{pb}}{\text{sig}}}^{(1)}$	-0.8										
$c_1^{M(1)}$	0.9	22.8									
$N_{\text{sig}}^{(1)}$	0.3	-66.8	-24.7								
$N_{\text{bkg}}^{(1)}$	-0.2	47.6	17.6	-45.6							
$c_1^{\Delta E(2)}$	0.0	0.0	0.0	0.0	0.0						
$f_{\frac{\text{pb}}{\text{sig}}}^{(2)}$	0.0	-0.8	-0.3	0.7	-0.5	1.5					
$c_1^{M(2)}$	0.0	-0.3	-0.1	0.2	-0.2	1.2	23.2				
$N_{\text{sig}}^{(2)}$	0.0	0.6	0.2	-0.5	0.4	-1.2	-69.7	-23.2			
$N_{\text{bkg}}^{(2)}$	0.0	-0.5	-0.1	0.4	-0.3	0.9	51.4	17.1	-44.6		
$\mathcal{S}_{CP}$	-0.1	-4.7	-1.5	3.7	-2.7	0.2	15.2	5.4	-11.7	8.7	
$\mathcal{A}_{CP}$	0.0	-1.5	-0.6	1.5	-1.0	0.0	4.0	1.3	-3.5	2.6	-7.4

CHAPTER 9

Conclusion

In this part of the thesis, we have studied *CP* violation in the tree-dominated color-suppressed $B^0 \to \eta_c K_S^0$ decay. This decay is *CP*-odd and the SM predicts its direct ($\mathcal{A}_{CP}$) and mixing-induced ($\mathcal{S}_{CP}$) *CP* asymmetry are the same as for the most precisely measured *golden channel* $B^0 \to J/\psi K_S^0$. This mode allows measuring the sinus of the unitary angle ϕ_1, for which it holds $\sin 2\phi_1 = \mathcal{S}_{CP}$ within a small theoretical uncertainty in the SM. The measurement is using the full Belle dataset of 772×10^6 of $B\overline{B}$ events.

The BaBar Collaboration firmly established mixing-induced *CP* violation in this channel, in particular using the subsequent three-body decay $\eta_c \to K_S^0 K^+ \pi^-$ and 465×10^6 $B\overline{B}$ pairs [48]. Later studies of such decays [47] revealed significant interference effects, which were neglected in the previous measurements (for both *CP*V parameters and branching fraction). Our analysis exploits a more complex fitting approach, where the *CP*V parameters are estimated simultaneously with signal and background yields, offering better control over the related systematic effects. In addition, we performed a dedicated study to estimate the interference effect on the branching fraction measurement, which was done for the first time in the neutral mode. Such a study was already performed at Belle for our control mode, and we closely followed their interference model. The results of the dedicated interference studies were further used to estimate systematic effects due to the modified resonant η_c line shape resulting from the interference. Moreover, we check for possible dependence of the *CP*V parameters on the η_c mass, at least to the first order. As we observe no statistically significant effect, we further support the validity of the measurement of *CP*V parameters in $B^0 \to \eta_c(K_S^0 K^+ \pi^-)K_S^0$ decays, despite the significant presence of the non-resonant background, which is poorly understood both theoretically and experimentally. The statistical power of the available data is not enough to perform a full-scale angular analysis—such an endeavor can only be pursued in the future at Belle II. Needless to say, with (potentially) fifty times more data, the interference effects will have to be included in the data model, and possibly a much more advanced analysis will need to be performed.

One additional source of systematic uncertainties is the Punzi effect, not usually considered in such analyses, despite the Belle resolution function $\mathcal{R}_{B\overline{B}}$ depending on a number of conditional event-dependent variables. In our case, this effect becomes significant due to the simultaneous fit of yields and *CP*V parameters and high background level. Its correct treatment would require substantial additions to the data model, which would make it too slow. Given our current statistical precision, it is fully acceptable to consider these effects only as systematic uncertainties and not attempt to correct them. Especially as the dominant systematics for $\mathcal{S}_{CP}$ stems the possible *CP*V contributions to the peaking back-

ground. In this work, this effect is estimated directly from the sideband data for the first time. Previously, only MC estimates were available.

The last main difference from the previous analyses is the usage of beam-constrained mass $M_{\rm bc}$ for the best B candidate selection. Already for the $\eta_c \to p\bar{p}$ channel, it was recognized [40] that utilization of variables related to quality (goodness-of-fit) of the η_c decay vertex in the best candidate selection could introduce unwanted systematic biases.

The model is cross-validated on the charged control channel $B^+ \to \eta_c K^+$, $\eta_c \to K_S^0 K^+ \pi^-$, where no CP-violation is expected. We obtained

$$
\begin{aligned}
\mathcal{S}_{CP} &= -0.03 \pm 0.08 \\
\mathcal{A}_{CP} &= -0.05 \pm 0.06 \\
\mathcal{B} &= (2.49 \pm 0.09 \pm 0.40_{\rm int}) \times 10^{-5},
\end{aligned}
\tag{9.1}
$$

where the uncertainties are statistical only. For the branching fraction, not corrected for data/MC efficiency differences $(1-2\%)$, we also quote the estimated uncertainty due to the interference with the non-resonant background. These results are fully consistent with expectations and previous measurements.

After all cross-checks, validations, and evaluation of systematic uncertainties, the final CP violation and branching fraction measurement were performed. We obtained

$$
\begin{aligned}
\mathcal{S}_{CP} &= 0.59 \pm 0.17 \pm 0.07 \\
\mathcal{A}_{CP} &= 0.16 \pm 0.12 \pm 0.06 \\
\mathcal{B} &= (9.8 \pm 0.6 \pm 0.4 \pm 2.3_{\rm int}) \times 10^{-6}
\end{aligned}
\tag{9.2}
$$

for $B^0 \to \eta_c K_S^0$, $\eta_c \to K_S^0 K^+ \pi^-$. The uncertainties are statistical, systematic, and (for $\mathcal{B}$) to account for the interference with the non-resonant background, respectively.

The precision for $\mathcal{S}_{CP}$ is very slightly worse than for BaBar. There are two main reasons why the precision is not better despite the larger available data set. First, the error estimates have some uncertainty themselves. For example, in the ensemble studies, these are $\sim 10\%$ of the estimated error for $\mathcal{S}_{CP}$. Also, BaBar was 'lucky' to observe a large significance[1] for $\mathcal{S}_{CP}$. Second, our analysis attempts to take the interference with the non-resonant background into account for the first time. For this, our multi-dimensional fit allows for some shape compensations and, in particular, has floating parameters for signal and background yields. Thus, our statistical errors include the uncertainty and correlations stemming from the peaking background. We can see about 15% correlation of $\mathcal{S}_{CP}$ to the peaking background fraction in the SVD2 experiment (see Table 8.3), which inflates the error slightly. Although our method yields slightly larger statistical uncertainties, we believe these are estimated more carefully than in the traditional approach, where signal and background yields are fixed from a previous fit, and the correlations are ignored. In addition, the evaluation of systematic errors directly propagates possibly correlated effects from, e.g., changes in the signal shape distribution to the CPV parameters.

Fig. 9.1 shows a two-dimensional likelihood scan for the CPV parameters, considering statistical errors only. Note the slight 7% correlation of $\mathcal{S}_{CP}$ and $\mathcal{A}_{CP}$.

---

[1]Over 5σ, but with an upward deviation from the SM.
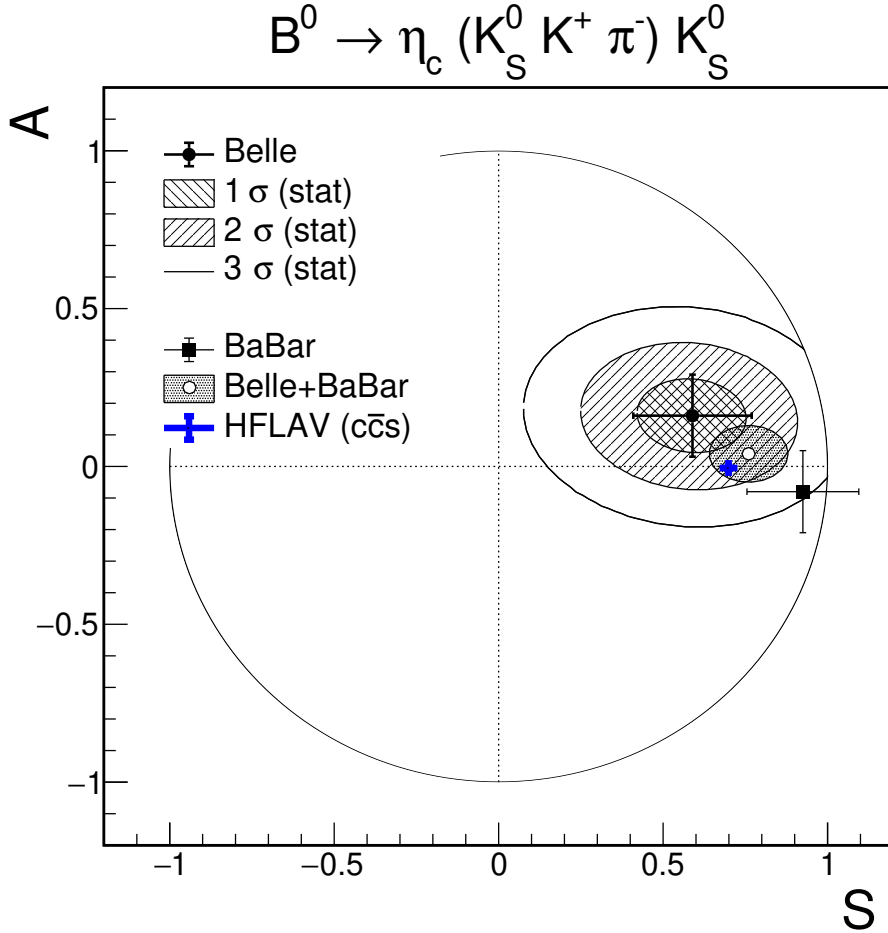
Figure 9.1: Overview of CPV parameter measurements in $B^0 \to \eta_c(K_S^0 K^+\pi^-)K_S^0$. The confidence regions corresponding to one, two, and three sigmas for two degrees of freedom (39.3%, 86.5%, 98.9%) are shown for our result, considering only the statistical uncertainty. The measurement with its total uncertainty is also compared to the BaBar measurement [48], and the world average for all charmonia [49]. In addition, a simple combination of our and BaBar measurement is shown with 1σ contour, ignoring any correlations. The circle with unit radius shows boundaries of the physically allowed region, where $\mathcal{S}^2 + \mathcal{A}^2 \leq 1$.

The points with error bars show measurements with their total uncertainty—for our result, the BaBar result, and the world average for all charmonia. Finally, a simple (ignoring any correlations) combination of our and BaBar results is shown, which yields

$$
\begin{aligned}
\mathcal{S}_{CP}(\text{Belle} + \text{BaBar}) &= 0.76 \pm 0.12 \\
\mathcal{A}_{CP}(\text{Belle} + \text{BaBar}) &= 0.04 \pm 0.09.
\end{aligned}
\tag{9.3}
$$

The average is now fully consistent with the SM prediction and the world average for all charmonia within $\sim 0.5\sigma$. It is natural to combine the Belle and BaBar results with very similar uncertainties and almost uncorrelated systematics. The only significant source of correlated systematic error is the tag-side interference effect on $\mathcal{A}_{CP}$ (around ~ 0.03) and the CP violation in the peaking background (around 0.05) for $\mathcal{S}_{CP}$. Note that the BaBar measurement was slightly deviating from the expectation for $\mathcal{S}_{CP}$, while our shows a slight departure for $\mathcal{A}_{CP}$. The

combination then suggests that these are only statistical fluctuations, and there are no signs of NP contributions to the CP-violation parameters in $B^0 \to \eta_c K_S^0$ at the current level of precision.

Part III

Belle II Detector Alignment

Introduction to Alignment

For most simulations, the positions of the detector elements are assumed to be known and exactly as designed. In reality, the detector is constructed with a limited mechanical precision at level $\mathcal{O}(0.1 \text{ mm})$ and the sensitive elements can also move in time. With detection precision at level $\mathcal{O}(0.01 \text{ mm})$ for silicon sensors, this would significantly degrade the physics performance of the vertex detector. While survey measurements prior to and after the construction and installation of the sub-detectors are often[1] of high importance to reach a reasonable starting geometry, the ultimate precision can usually only be achieved by utilizing a large amount of various in-situ measurements of charged tracks' trajectories recorded during the detector operation.

The procedure for determination of the geometrical configuration of the detector elements from tracks is called **track-based alignment** or just **alignment**. Sometimes, by alignment, one means the result of such a method, the **alignment constants**. The alignment constants are a subset of a much larger set of **calibration constants** that correspond to the various parameters needed for generic reconstruction of objects and calculation of aggregated information for high-level physics analyses. In this sense, alignment is part of the detector **calibration**.

Good alignment performance, namely for the vertex detector, is a key ingredient, for example, for time-dependent *CP*-violation measurements, like the analysis presented in Part II of this thesis. The physics motivations are briefly discussed in Sec. 10.1, followed by requirements on the method in Sec.10.2. Many calibrations are based on the same philosophy as alignment: as the underlying physics of propagation of charged particles in a magnetic field and detector material is well understood, we can construct a model for the particle trajectory and fit this model to the observed data. An imperfect detector description will result in a sub-optimal fit unless the calibration constants themselves can be set floating. Obviously, adding the degrees of freedom from, e.g., detector positions to a track fit would result in an underdetermined problem. Therefore in alignment, data from many tracks must be combined, making the problem overdetermined, such that many (usually thousands or more) measurements contribute to the determination of each single alignment parameter. The general formulation of the alignment problem is discussed in Sec. 10.3, after which the main challenges of this complex task are summarized in Sec. 10.4.

The rest of this part of the thesis is organized as follows. In Chapter 11, we describe the alignment method utilized at the Belle II detector. Specifics of

---

[1]At Belle II, such a complete survey was never available for the full vertex detector. Our alignment procedure derived the positions starting from the design geometry. Only for the drift chamber, survey measurements are utilized for starting geometry, end-plate deformation, and wire tensions [21].

the necessary mathematical formulation of the problem for over sixty thousand alignment constants of the Belle II vertex detector and central drift chamber are given in Chapter 12. The method is implemented in the official Belle II software and automated in the regular production of calibration constants, as presented in Chapter 13. The historical overview in Chapter 14 summarizes some of the most important results, problems, and solutions from the past, up to early physics data. We divide the detailed evaluation of the alignment method into two parts. In the first part, we discuss the alignment of all parameters, except CDC wires. This baseline alignment of about three thousand parameters allows for some extended studies of its systematic properties, presented in Chapter 15. Although the baseline alignment performed well, the experience and needs that resulted from real detector operation led to a significant update of the method to include more data samples and up to sixty thousand alignment parameters, including CDC wires. The path toward this extension is presented in Chapter 16. Finally, the performance of the baseline alignment and its new version, utilized for data reprocessing in 2021, is evaluated in Chapter 17. In Chapter 18, we discuss a number of possible extensions of the method left for further research. Chapter 19 is devoted to a critical discussion of the results, concluded in Chapter 20.

10.1   Physics Motivation

With the large dataset that Belle II aims for, many measurements, previously statistically limited at Belle, can become dominated by systematic uncertainties. Alignment constitutes a major systematic contribution to precision lifetime measurements. Belle II will play an important role in such studies thanks to its high-precision vertex detector, offering about twice better vertex resolution than in the previous generation of B-Factories, together with the clean environment and well-defined initial state of the electron-positron collisions. Such precision measurements are vital to pursuing the search for physics beyond the SM.

For example, the τ lifetime is an input to lepton flavor universality tests, a very active area of experimental research, especially due to several indications for deviations from SM predictions [12]. Precise measurements of lifetimes of charm mesons and baryons can serve as tests of QCD predictions. The flagship of the B-Factory measurements is time-dependent CP violation (TDCPV), which requires high vertex resolution to measure the distance between B meson decay vertices. TDCPV is another prime example that motivates advanced alignment procedures. At the end of the Belle II experiment, the unitary angles are expected to be measured at a precision of $1°$ or better ($0.4°$ for ϕ_1) [11], and any inconsistencies would point to physics beyond the SM. With such statistical precision, control over systematic uncertainties is crucial.

This means that a number of effects that were safe to neglect in the previous generation of experiments can now lead to significant performance degradation. For example, a more detailed description of silicon sensor deformations or higher frequency of realignment was expected to be necessary to fully exploit the physics potential of the new vertex detector. Apart from statistical precision of the alignment, which only contributes to a worse resolution on average, systematic misalignment can directly bias physics measurements and thus must be well under control. Furthermore, the vertex detector also needs to be aligned with respect to

the drift chamber, whose alignment itself contributes significantly to the physics performance. While these tasks were performed separately at Belle and BaBar, this raises possible concerns about the systematic effects stemming from such a decoupled procedure, where correlations among the different detectors can only be resolved by an iterative approach.

10.2 Alignment Requirements

The generic requirement is that the resolution observed for data is at most by 20% worse than for Monte Carlo (MC) simulations. Such a discrepancy is viewed as acceptable but might still need ad hoc scaling corrections at the analysis level. This includes all effects, including too optimistic simulation. Thus the alignment method should lead to a much smaller degradation than 20% in simulation studies to leave room for a wide range of additional effects, underestimations, or mis-modeling at both data and MC levels. In fact, our minimum target for resolution worsening due to alignment in simulations is only negligible degradation, less than 5%.

The physics requirements for the Belle II detector are, in principle, similar to its predecessor Belle SVD2 [16] or its competitor BaBar SVT [70]. Due to the reduced boost factor at Belle II, the design requirements estimated that approximately twice better vertex resolution in the boost direction than achieved for data at Belle is needed to not degrade the TDCPV measurement precision. In both previous B-Factory experiments, the alignment precision for the silicon sensors was required to be better than 10 μm on average. At Belle II, we derived the corresponding number in older preliminary studies concerning random misalignment [71, 72]. The general conclusion is that alignment precision better than 5 μm is needed for negligible (less than 5%) degradation. Under 3 μm, the effects are almost impossible to observe even in large-statistics samples and when reconstructed observables and MC truth information is compared on simulations.

A non-random, systematic misalignment could introduce unwanted physics biases and thus needs special care and studies to ensure it can be controlled with a sufficient level of precision. For example, changes in the total radial or longitudinal (z) scale of the detector could introduce bias into the measured absolute lifetimes of unstable particles. The total length of the VXD detector (about 50 cm) must be known significantly better than 50 μm for the outer SVD layer sensors in order to achieve absolute scale precision better than 10^{-4}. Note that for much shorter PXD layers, this value gets correspondingly smaller. Similarly, the outer SVD radius (14 cm) must be determined to better precision than 14 μm.

10.3 General Formulation

The purpose of the alignment procedure is to determine the (vector of) alignment parameters $\boldsymbol{a}$, respective their (small) corrections $\Delta\boldsymbol{a}$. In the track-based alignment, the track model predictions $\boldsymbol{p}_j(\boldsymbol{a}, \boldsymbol{q}_j)$ for each track j depend on $\boldsymbol{a}$, as well as on the set of track parameters $\boldsymbol{q}_j$. Along the track, multiple measurements can be compared to the predicted values. The difference between the i-th measured
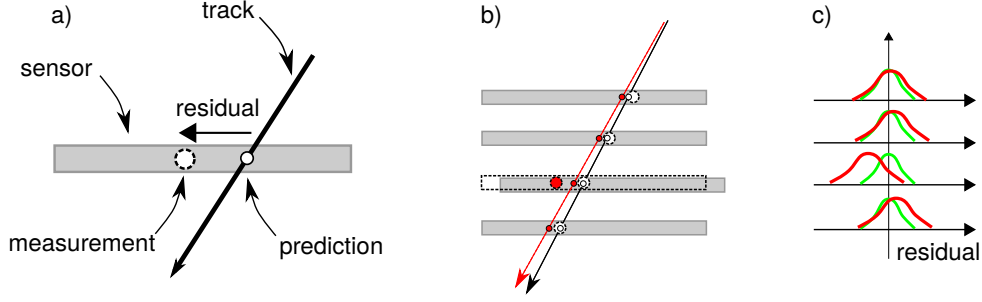
Figure 10.1: Definition of track-to-hit residual as a difference between the measurement and the prediction from the track fit (a). If one sensor's position differs from the assumed one (dashed), the reconstructed measurement will be shifted (red), pulling the track fit towards it (b). This will result in (c) residual distributions (red), which are biased and wider than ideal (green).

value m_{ij} on track j and the prediction (corresponding to i-th measurement) in the local frame of the measurement

$$r_{ij}(\boldsymbol{a}, \boldsymbol{q}_j) = m_{ij} - p_{ij}(\boldsymbol{a}, \boldsymbol{q}_j) \tag{10.1}$$

is called **residual**. Residuals should be normally distributed with zero mean and width given by the measurement uncertainty. If the assumed positions of the sensitive elements are incorrect, the residuals will appear biased (shifted) and wider as the track fit quality worsens. This is illustrated in Fig. 10.1. Optimal values for the alignment parameters can be obtained by minimizing the residuals from many tracks using the method of least squares. Each residual should be normalized (weighted) by the corresponding measurement uncertainty σ_{ij}. Such a quantity is usually called **pull** (r_{ij}/σ_{ij}). In general, these uncertainties may be correlated[2] among different measurements of the same track. The objective function to be minimized is the sum of the squares of the weighted residuals, which can be written as

$$\chi^2(\boldsymbol{a}, \boldsymbol{q}) = \boldsymbol{r}^\mathsf{T}(\boldsymbol{a}, \boldsymbol{q}) \boldsymbol{V}^{-1} \boldsymbol{r}(\boldsymbol{a}, \boldsymbol{q}), \tag{10.2}$$

where $\boldsymbol{q} = (\boldsymbol{q}_1 \ \boldsymbol{q}_2 \dots)^\mathsf{T}$ is the vector of all parameters of all tracks, $\boldsymbol{r}$ is the vector of all residuals and $\boldsymbol{V}$ is the covariance matrix of the measurements. We can denote $\boldsymbol{x} = (\boldsymbol{a} \ \boldsymbol{q})^\mathsf{T}$ and for a small change of the parameters, $\boldsymbol{x} = \boldsymbol{x}_0 + \Delta \boldsymbol{x}$, the residuals are linearized as

$$\boldsymbol{r}(\boldsymbol{x}) = \boldsymbol{r}_0 + \boldsymbol{J} \Delta \boldsymbol{x} \tag{10.3}$$

where $\boldsymbol{r}_0$ are the residuals for initial estimates of the alignment and track parameters and $\boldsymbol{J} = \frac{\partial \boldsymbol{r}}{\partial \boldsymbol{x}}$ is the Jacobian matrix. From equations 10.2 and 10.3, the system of normal equations for minimum of the linearized χ^2 is

$$\underbrace{\boldsymbol{J}^\mathsf{T} \boldsymbol{V}^{-1} \boldsymbol{J}}_{\boldsymbol{C}} \Delta \boldsymbol{x} = - \underbrace{\boldsymbol{J}^\mathsf{T} \boldsymbol{V}^{-1} \boldsymbol{r}_0}_{\boldsymbol{b}} . \tag{10.4}$$

---

[2]Multiple scattering effects correlate adjacent sensors in a track and multi-dimensional measurements can be correlated internally.

The matrix $\boldsymbol{C}^{-1}$ is the covariance matrix of the parameters $\boldsymbol{x}$, which is symmetric and positive-definite. The solution is then obtained formally by inversion as

$$\Delta\boldsymbol{x} = -\boldsymbol{C}^{-1}\boldsymbol{b}. \tag{10.5}$$

The practical issues with such an expression become evident when we consider the dimension of the $n \times n$ matrix $\boldsymbol{C}$. Its size is given as

$$n = \dim(\boldsymbol{a}) + \sum_j \dim(\boldsymbol{q}_j). \tag{10.6}$$

While the alignment parameters are counted in thousands, the number of individual tracks used for alignment can easily reach millions. The number of their local track parameters will be typically even larger than 10^7 in the problems discussed in this thesis. As the computational complexity of matrix inversion is $\mathcal{O}(n^3)$ and memory requirements scale as $\mathcal{O}(n^2)$, a direct solution is not practically feasible.

Different methods have been developed to deal with the computational complexity of the alignment problem. Two main approaches are used in HEP experiments:

- Local alignment, used, e.g., at BaBar [70], or Belle [16], where the χ^2 is minimized for each alignable element (like a silicon sensor) individually while keeping the track parameters fixed. Only inversion of small (usually 6×6) matrices is needed. This procedure needs to be iterated many times while the track parameters are updated after each iteration, which determines incremental corrections to the alignment parameters. In this approach, the correlations among different modules are neglected.

- Global alignment, used, e.g., at H1 [73], or CMS [74], which minimizes χ^2 with respect to all alignment and track parameters simultaneously. The special structure of the alignment problem is exploited, leading to significantly reduced computational complexity. The solution can be obtained without any further approximations. This method was chosen for the Belle II detector and is discussed in detail in the following sections.

One last approach to mention is alignment based on the Kalman Filter [75], where alignment parameters are updated incrementally with the addition of each new track. This is one of the alternative alignment methods developed for the CMS but never used in production [76] due to the computational requirements of the extensive bookkeeping necessary for the preservation of correlations among the alignment parameters.

10.4 Alignment Challenge

The major challenge of alignment based on track-to-hit residuals are linear combinations of alignment (and track) parameters, to which the χ^2 expression in Eq. 10.2 is weakly or not sensitive at all, known as **weak modes**. Certain movements or deformations of the detector can be absorbed in the track parameters and thus potentially bias physics measurements. Such deformations represent correlated movements of the detector elements, which transform particle trajectories into other valid trajectories with different parameters. Such distortions

might be present in the real detector, and the alignment procedure might not be able to correct them. Alternatively, the alignment procedure might unintentionally introduce such deformations, resulting in minimized residuals and working tracking but biased physics measurements.

In fact, the final alignment solution errors are typically entirely dominated by the (combination of) weak modes, understood as systematic errors of the procedure, as the statistical errors can usually be improved far beyond the systematic ones by simply using more data at the input. These systematic errors need to be studied, and their possible physics impact evaluated.

Weak modes result from a general lack of an absolute reference. There are no reference tracks with known parameters. However, for example, in the decay of an unstable particle, its decay products have exactly opposite momenta in the CM frame and originate from the same vertex point. Although we do not know the true vertex point, both parts of the detector, which registered the tracks, should provide compatible estimates. In such a way, different parts of the detector can be correlated by the tracks, which will result in a reduction of plausible track parameter biases. Similarly, cosmic-ray muons crossing the detector correlate very different combinations of modules/sensors than tracks originating from the collisions. Thus such additional data samples with various topologies or constraints can prevent a range of deformations, which are a weak mode for collision-only tracks. This is illustrated for two typical weak modes in Fig. 10.2.

The usage of several complementary datasets is the most basic ingredient for reducing weak modes and a good alignment, but it may not be sufficient. The alignment method should further support the inclusion of additional constraints or external measurements for the alignment parameters, take care of measurement outliers, and in particular, be able to provide the solution in a reasonable time. To utilize all available information, including all correlations, the global alignment approach is the most promising. One slight drawback is the lack of possibility to monitor the progress and spot problems easily on the fly. The global alignment is a *black box* approach, which (almost) always provides a solution in a single step. Such a solution is only as good as the underlying problem statement. The worst possible problem is shortcomings in detector modeling, for example, in the material budget or magnetic field. It is desirable for alignment to use the same detector material or magnetic field model as the official reconstruction for physics analyses. In such a way, no additional systematic effects need to be considered, and any future improvements in the detector model or reconstruction will automatically propagate to a more reliable alignment. The challenge is to combine all this information in a generic way to allow for alignment of the complete tracking system, with over 60×10^3 parameters.

Last but not least, many of the challenges are purely practical, resulting from constraints in real experiment operation. A typical problem is the lack of particular data samples. For example, very useful data without a magnetic field are only very limited. As time-dependent alignment might be needed, this could prevent frequent realignment ad hoc. This is also related to a problem specific to the experiment location—in Japan, one can occasionally expect earthquakes, which could alter the alignment at any time in unexpected ways.

This brings us to the single most challenging property of a real detector's align-
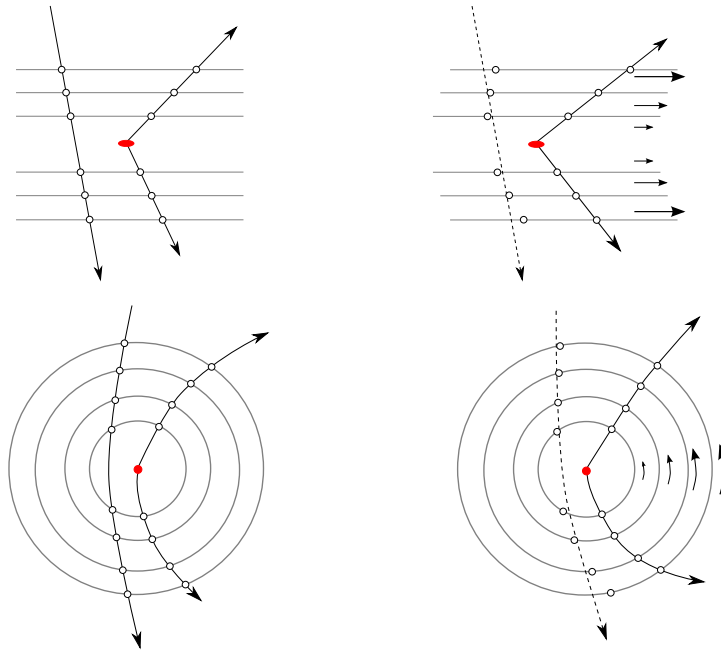
Figure 10.2: Illustration of tracks coming from the IP (red point) and a cosmic ray track in $y - z$ projection (top) and $R - \phi$ projection (bottom). For a 'telescope' ($\Delta z \propto R$) deformation (top right), tracks from the IP only change their direction without affecting residuals, while the cosmic track fit becomes bad due to the apparent *kink* in the measurements. For a 'curl' ($\Delta \phi \propto R$) deformation (bottom right), the curvature of tracks from the IP changes in the opposite way for positively and negatively charged tracks. For the illustrated cosmic trajectory, such a 'curl' deformation would result in incompatible curvature estimates for the upper and lower arm and sub-optimal residuals.

ment. The detector is not entirely stable in time for many reasons, from changes in temperature or magnet conditions to vibrations or even the earthquakes mentioned above. Thus calculating a single set of alignment constants for a couple of weeks or even days of experiment operation could significantly degrade physics performance. While alignment itself is the best tool to extract information about the detector positions, the time resolution is practically limited, e.g., by the rate of cosmic ray tracks available for alignment. While other monitoring methods are, in principle, possible (like continuous laser survey measurements), no such technique is utilized at Belle II, which would directly and independently monitor the sensitive detection elements, at least at a couple of reference points. Thus we have to extract the information about the time-dependence of the alignment constants from validation studies using reconstructed tracks to define proper intervals where alignment of all or some subset of degrees of freedom can be considered stable. The optimal solution to this problem is not explored in this thesis. The time dependence must be understood step by step, and this work will pave the way for such studies. However, we are still trying to correlate some of the changes to the detector and machine conditions, and this may take some time.

The truth is that understanding the origin of the time-dependence can be extremely difficult. While alignment can give some clues, only long-time experience and attempts to correlate alignment changes to detector operation conditions can reveal the source of some of the significant movements, typically happening at the level of larger structures rather than individual sensors. Thus individual sensors were not initially expected to need a frequent realignment. However, towards the end of this thesis, we will see that to fulfill the physics potential of the PXD, even such an endeavor needs to be undertaken.

CHAPTER 11

Alignment Method

The Millepede alignment method for the vertex detector was already proposed in the Belle II Technical Design Report [10]. The method is described in detail in Sec. 11.1. A key ingredient to formulate the alignment problem is a suitable global track model. This track model or track parametrization needs to properly take into account all the knowledge about particle propagation in the detector. For this purpose, we integrated the General Broken Lines (GBL) track model and refit into a generic track fitting toolkit GENFIT2. This effort was a subject of the previous work of the author [77], where a full description with more details about the implementation can be found.

The track model relates a set of track parameters to the predictions at the measurements. The track parameters are usually initially defined at some reference point and need to be extrapolated to other points along the trajectory. The necessary machinery to perform such extrapolations of the track state, handled by the GENFIT2 toolkit, is briefly explained in Sec. 11.2. Sec. 11.3 to 11.5 describe the GBL track model and the generic way to parametrize the detector material and multiple scattering effects, as well timing-related corrections. In Sec. 11.6 we describe an extension of the track model formulation from single trajectories to combined objects, representing decays of unstable particles. Finally, in Sec. 11.7, the data samples used for alignment MC studies as well as for real data are summarized.

11.1 Millepede II

The global alignment approach is implemented in the Millepede method, first introduced by V. Blobel for the H1 experiment [73, 78], which exploits the special structure of the matrix equation 10.4. The dimension of the matrix can be reduced to the number of alignment parameters using block matrix algebra. It allows solving the problem in a single step, without external[1] iterations, and any approximations beyond the initial linearization.

Since then, Millepede has been used in a number of HEP experiments. The development of its current version, Millepede II [79, 80], adjusted to the needs of large LHC experiments, namely the CMS (Compact Muon Solenoid) [81], started in 2005, and the tool has been maintained since then by the Statistics Tools group of the Analysis Center of the Helmholtz Alliance. Millepede II is independent of any experiment and publicly available at the DESY GitLab server [82].

---

[1]There can be internal iterations in the algorithm, but external iterations with a complete recalculation of the inputs are not needed in many cases.

For the Millepede formalism, uncorrelated measurements are assumed, and the χ^2 from Eq. 10.2, linearized around the initial parameter estimates, can be written as

$$\chi^2(\Delta \boldsymbol{a}, \Delta \boldsymbol{q}_1 \dots) = \sum_j^{\text{tracks}} \sum_i^{\text{meas}} \frac{1}{\sigma_{ij}^2} \left(r_{ij}(\boldsymbol{a_0}, \boldsymbol{q_{0}}_j) + \frac{\partial r_{ij}}{\partial \boldsymbol{a}} \Delta \boldsymbol{a} + \frac{\partial r_{ij}}{\partial \boldsymbol{q}_j} \Delta \boldsymbol{q}_j \right)^2, \quad (11.1)$$

where σ_{ij} is the measurement uncertainty, and $\boldsymbol{a_0}$ and $\boldsymbol{q_{0}}_j$ are the initial estimates of alignment and track parameters, respectively. Two categories of parameters are recognized:

- Local parameters $\boldsymbol{q}_j$ of the individual tracks, which influence only a limited subset of measurements.

- Global parameters $\boldsymbol{a}$, corresponding to the alignment parameters, which influence many measurements (residuals) and are correlated by the tracks crossing various parts of the detector.

For the alignment, the interest is only in the global parameters. The local parameters are usually anyway recomputed after the alignment and additional calibrations for official data reconstruction. To formulate the alignment problem, one needs, in addition to the initial measurement residuals r_{ij} and errors σ_{ij}, the derivatives of the residuals with respect to the local parameters $\frac{\partial r_{ij}}{\partial \boldsymbol{q}_j}$ (**local derivatives**) and the global alignment parameters $\frac{\partial r_{ij}}{\partial \boldsymbol{a}}$ (**global derivatives**). The normal equations 10.4 can be written in a block matrix form as

$$\begin{pmatrix} \sum_j \boldsymbol{C_1}_j & \dots & \boldsymbol{G}_j & \dots \\ \vdots & \ddots & 0 & 0 \\ \boldsymbol{G}_j^{\mathsf{T}} & 0 & \boldsymbol{\Gamma}_j & 0 \\ \vdots & 0 & 0 & \ddots \end{pmatrix} \cdot \begin{pmatrix} \Delta \boldsymbol{a} \\ \vdots \\ \Delta \boldsymbol{q}_j \\ \vdots \end{pmatrix} = - \begin{pmatrix} \sum_j \boldsymbol{g_1}_j \\ \vdots \\ \boldsymbol{\beta}_j \\ \vdots \end{pmatrix}. \quad (11.2)$$

For each GBL trajectory (simple or combined), there is a block with a matrix connecting only local parameters

$$(\boldsymbol{\Gamma}_j)_{kl} = \sum_i^{\text{meas}} \frac{1}{\sigma_{ij}^2} \left(\frac{\partial r_{ij}}{\partial q_{j,k}} \right) \left(\frac{\partial r_{ij}}{\partial q_{j,l}} \right). \quad (11.3)$$

On the two transposed borders, matrices

$$(\boldsymbol{G}_j)_{kl} = \sum_i^{\text{meas}} \frac{1}{\sigma_{ij}^2} \left(\frac{\partial r_{ij}}{\partial a_k} \right) \left(\frac{\partial r_{ij}}{\partial q_{j,l}} \right) \quad (11.4)$$

relate global and local parameters. Finally, in the upper left corner sits a sum of matrices

$$(\boldsymbol{C_1}_j)_{kl} = \sum_i^{\text{meas}} \frac{1}{\sigma_{ij}^2} \left(\frac{\partial r_{ij}}{\partial a_k} \right) \left(\frac{\partial r_{ij}}{\partial a_l} \right) \quad (11.5)$$

over all tracks j, each with a size corresponding to the number of global (alignment) parameters. The big matrix multiplies the vector of unknown corrections

164

to the alignment and track parameters. On the r.h.s is a vector from the gradient of the objective function with block entries

$$(\boldsymbol{g_{1j}})_k = \sum_i^{\text{meas}} \frac{r_{ij}}{\sigma_{ij}^2}\left(\frac{\partial r_{ij}}{\partial a_k}\right), \quad (\boldsymbol{\beta}_j)_k = \sum_i^{\text{meas}} \frac{r_{ij}}{\sigma_{ij}^2}\left(\frac{\partial r_{ij}}{\partial q_{j,k}}\right) \tag{11.6}$$

formed from (normalized) residuals times the corresponding global or local derivatives. The matrix 11.2 is huge (easily over $10^8 \times 10^8$, corresponding to one million tracks with 100 degrees of freedom); in particular, the block diagonal part occupied by the $\boldsymbol{\Gamma}_j$ matrices. These are inverses of the covariance matrices for local parameters of each trajectory, see Sec. 11.3 and Eq. 11.20 (with $\boldsymbol{A} = \boldsymbol{\Gamma}$).

11.1.1   Reduction of Matrix Size

The idea of matrix reduction and removal of unwanted local parameters can be illustrated using a simple example. A (block) matrix equation

$$\begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{pmatrix} \cdot \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{pmatrix} = \begin{pmatrix} \boldsymbol{a} \\ \boldsymbol{b} \end{pmatrix} \tag{11.7}$$

may be written as two equations for unknown (vectors) $\boldsymbol{x}$ and $\boldsymbol{y}$. If we are interested only in $\boldsymbol{x}$, we can simply express $\boldsymbol{y}$ from the second equation as

$$\boldsymbol{y} = \boldsymbol{D}^{-1}(\boldsymbol{b} - \boldsymbol{C}\boldsymbol{x}) \tag{11.8}$$

under the condition that $\boldsymbol{D}$ is invertible. Substitution back into the first equation yields a problem reduced to the determination of $\boldsymbol{x}$:

$$(\boldsymbol{A} - \boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C})\boldsymbol{x} = \boldsymbol{a} - \boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{b}. \tag{11.9}$$

The solution is then obtained by inverting the matrix on the l.h.s., called Schur complement.

Repeated application of the matrix reduction with the Schur complement will remove the bordered and diagonal part and update the vector on the r.h.s. of Eq. 11.2, yielding a matrix and a vector

$$\boldsymbol{C} = \sum_j \boldsymbol{C_{1j}} - \sum_j \boldsymbol{G}_j \boldsymbol{\Gamma}_j^{-1} \boldsymbol{G}_j^{\mathsf{T}}, \quad \boldsymbol{g} = \sum_j \boldsymbol{g_{1j}} - \sum_j \boldsymbol{G}_j \boldsymbol{\Gamma}_j^{-1} \boldsymbol{\beta}_j \tag{11.10}$$

and normal equations in the form

$$\boldsymbol{C}\Delta\boldsymbol{a} = -\boldsymbol{g}. \tag{11.11}$$

The matrix $\boldsymbol{C}$ is called **global matrix**, and we will only consider this reduced form for further discussion. The corrections to the alignment parameters $\Delta\boldsymbol{a}$ can be obtained formally by inversion as

$$\Delta\boldsymbol{a} = -\boldsymbol{C}^{-1}\boldsymbol{g}. \tag{11.12}$$

For linear problems, this yields an exact solution in a single step, with all correlations preserved in the solution, reached without any approximations. The whole solution may need to be iterated to go beyond linear problems, with a repeated evaluation of the local and global derivatives from the track model, which are fixed inputs in the formalism. Such iterations are typically only really needed for initial alignment. Typical alignment corrections in practice are small enough that a single iteration can be sufficient.

11.1.2 Linear Equation Solvers

Millepede II offers several solution methods for the linear Eq. 11.11. We will explicitly discuss those directly used in this work:

- **Inversion** is only suitable for problems with up to a couple of thousand global parameters as its solution time scales as $\mathcal{O}(n^3)$. While this means a couple of minutes (including other computations) for, e.g., 3×10^3 parameters, with $\sim 60 \times 10^3$ parameters, over half a day of computation using ten threads was needed in our tests.

- **Diagonalization** is about ten times slower than inversion but allows access to the eigenvalues of the matrix C and extraction of eigenvectors corresponding to the smallest eigenvalues in the spectrum. These eigenvectors represent linear combinations of parameters to which the χ^2 is only weakly or not sensitive at all. This result can be used to study the weak modes of the solution.

- **Cholesky decomposition**, which is faster and numerically more stable than inversion [80]. It has even a faster variant of **LAPACK factorization** utilizing highly optimized external LAPACK libraries [83], e.g., with Intel MKL implementation [84]. This last method can be an order of magnitude faster than the default Cholesky. These methods do not compute parameter errors by default, in contrast to the inversion and diagonalization.

- **MINRES-QLP**, or generalized minimization of residuals, is an approximate fast iterative solution method minimizing $|C\Delta a - g|$, suitable for very large problems with hundreds of thousands of alignment parameters, heavily utilized at the CMS experiment [74, 85]. This method works well for sparse matrices, which result from a limited number of connected global and local parameters. Some or all correlations may be lost in this method.

11.1.3 Treatment of Outliers

The method of least squares assumes that the measurement errors are Gaussian-distributed. Non-Gaussian tails of multiple-scattering or bad measurements with large normalized residuals (pulls) can significantly influence the χ^2 and thus distort the solution. Several methods can be used in Millepede II to suppress these outliers, and we regularly use a combination of all of them in the configuration described below.

As initially, it is not known if the large residuals are caused by misalignment or outliers; only measurements with extremely large pulls are rejected before the global solution. Additional internal iterations are necessary to reject bad data. The full solution is not performed again in these iterations (by default), only incrementally updated.

Entire tracks can be rejected if they have too large χ^2 divided by the number of degrees of freedom (NDF). Successively more strict cuts on χ^2/NDF are used, with the final iteration removing all tracks with χ^2/NDF corresponding to more than three standard deviations.

In addition, individual measurements can be down-weighted with the method of M-estimators, which requires repeated local track fits. We are using two iterations that employ the Huber function, where the influence of large pulls is constant on the χ^2. Additional iterations, not used regularly at Belle II, utilize the Cauchy function, where the weight even decreases for large pulls. Tracks with a large fraction (> 10 % in this work) of down-weighted measurements are also removed.

11.1.4 Removal of Degrees of Freedom and Constraints

It is generally necessary to remove some degrees of freedom from the solution of the alignment problem, which is otherwise underdetermined. The trivial undefined degrees of freedom in residual minimization correspond to three global shifts and three rotations of the complete detector. If no restrictions are posed, e.g., using information from other sub-detectors, one could, for example, fix alignment parameters of some detection element, like a sensor. This sensor will then become an alignment reference. Millepede II allows for the fixing of arbitrary parameters and defining initial values and uncertainties for them.

But in general, it is more practical to fix some linear combinations of the parameters, for example, the average of shifts and rotations of all sensors. This can be represented as six linear equations—each defining one linear equality constraint. The set of all constraints poses an additional condition

$$\boldsymbol{A}\Delta\boldsymbol{a} = \boldsymbol{m}, \tag{11.13}$$

where $\{\boldsymbol{A}\}_{ij}$ is the j-th coefficient of the i-th constraint. Each constraint is formulated in the form

$$\sum c_j \Delta a_j = m_i \tag{11.14}$$

where m_i is usually set to zero and c_j are the constraints' coefficients.

Millepede II allows for two possible ways of including linear constraints into the solution. The classical method of Lagrange multipliers, where the global matrix is enlarged, is accompanied by an elimination method, where instead, the dimension of $\boldsymbol{C}$ is reduced before the solution. The elimination method is used by default.

11.2 Track Fitting with GENFIT2

Track reconstruction and parameter estimation (fitting) in modern HEP detectors is a complex task. Measurements of various types need to be combined in the fits and compared to predictions obtained by extrapolations in detector material and magnetic field. The GENFIT toolkit was developed as a generic framework for track fitting in HEP experiments [86]. It was adopted for the Belle II track fitting, and after a major upgrade, it is now known as GENFIT2 [87].

GENFIT2 has a modular design. Its functionality can be roughly divided into three main categories:

- Generic representation of reconstructed hits dealing with their different dimensionality and expressing non-planar measurements in a suitable way.

- Track representation and tools for particle extrapolation in a magnetic field and detector material using the Runge-Kutta method.

- Track fitting algorithms, which combine the above information to estimate the track parameters. At Belle II, a Deterministic Annealing Filter (DAF) based on Kalman Filter is the default track fitting method [88].

The complete separation of the track fitting algorithms from details of the underlying measurements is achieved by the method of virtual planes. After each extrapolation, the track state is expressed in some virtual plane. For planar detectors, this plane coincides with the sensor plane, while for wire hits in the CDC, the plane is constructed as depicted in Fig. 11.1. Without any measurement that would define a sensible preferred plane's orientation, a co-moving frame of the track (with a normal in the track direction) is constructed.
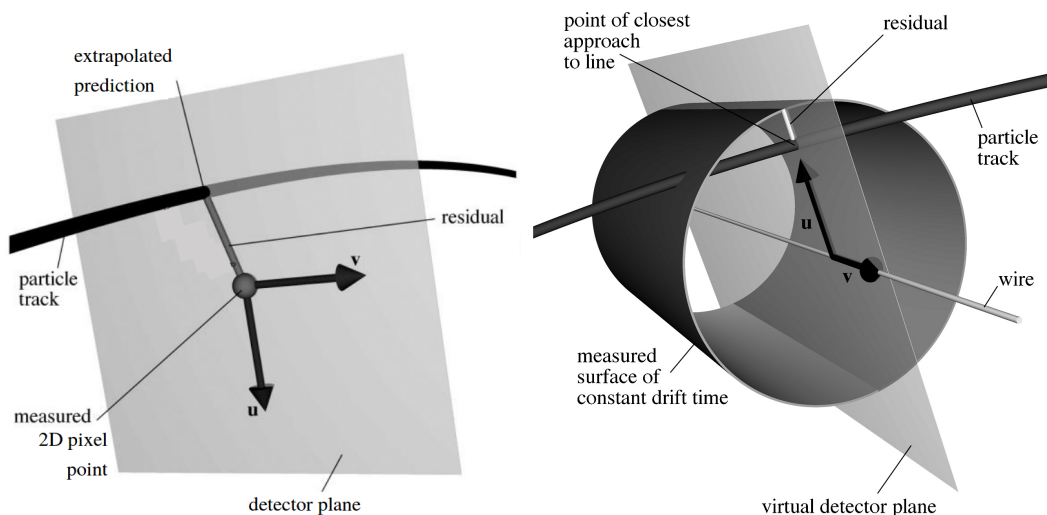
Figure 11.1: Construction of virtual planes and definition of the track-to-hit residual for planar hits (left) and wire hits in the drift chamber (right).

Each virtual plane defines the local frame in which the state of the track is represented by a five-dimensional vector $\boldsymbol{q} = (q/p, \boldsymbol{u}' = (\frac{du}{dw}, \frac{dv}{dw}), \boldsymbol{u} = (u, v))^\mathsf{T}$, where q/p is the signed track curvature, $\boldsymbol{u}'$ is a vector of the track slopes and $\boldsymbol{u}$ is a vector of positions (offsets). The dimensionality of the measurement is encoded in a (projection) $\boldsymbol{H}$-matrix, which reduces the local track state (the prediction) to the local measurement dimension, such that the local measurement $\boldsymbol{m}$ has a residual defined as

$$\boldsymbol{r} = \boldsymbol{m} - \boldsymbol{H}\boldsymbol{q}. \tag{11.15}$$

11.2.1 Helix Parameter Definitions

The track fit's final output is the estimate of the tracks' parameters, typically extrapolated to the point of closest approach (POCA) to a given reference (pivot). The tracks are extrapolated to the POCA, starting typically from a local track state at the first measurement after the fit, propagating the particle backward.
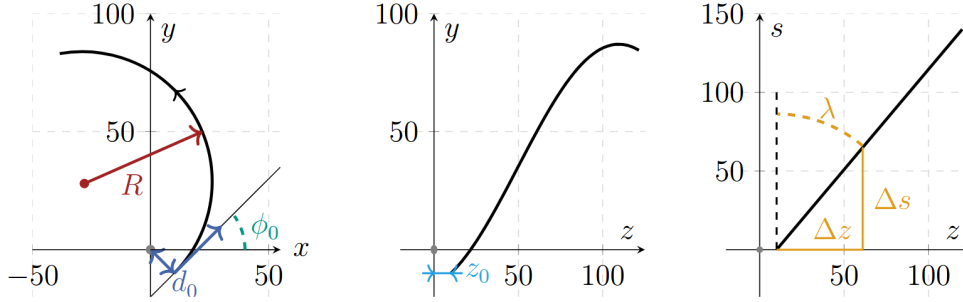
Figure 11.2: Definition of the Belle II helix parameters at the POCA in a schematic projection of a helix trajectory in the $x-y$ (left), $y-z$ (middle), and $s-z$ space (right). [88]

The trajectory of a charged particle propagating in a magnetic field can be locally described by a helix with five free parameters. At Belle II, the pivot point by default corresponds to the origin of the coordinate system. The definition of the five helix parameters is sketched in Fig. 11.2. In the $x-y$ plane, where the POCA is defined, the parameters are

- signed distance of the POCA to the origin d_0,

- the angle ϕ_0, between the transverse momentum at the POCA and the x-axis,

- signed curvature ω (sign depending on charge), given as $|\omega| = 1/R$, where R is the radius of the helix.

For positively charged particles, the sign of d_0 is positive (negative) if the pivot point lies outside (inside) the circle given by the helix projection in the $x-y$ plane. For negatively charged tracks, this definition is reversed [8]. In the $y-z$ and $s-z$ space, where s is the path length along the trajectory, the parameters are

- distance of the POCA from the origin in the z-direction, z_0 and

- the tangent of the dip angle $\tan \lambda$.

In addition, variables dr and dz, also used at Belle, represent the transverse and longitudinal POCA with respect to (using as a pivot) the calibrated IP position.

11.3 General Broken Lines Track Model

The just discussed helix with five free parameters is only an approximate global description of the particle trajectories. In reality, the real magnetic field is not homogeneous; the particle undergoes energy loss and, in particular, travels through detector material resulting in multiple Coulomb scattering and deviations from a simple helical trajectory. General Broken Lines (GBL) [89, 90] is a track model and re-fit which explicitly incorporates multiple scattering effects in the parametrization. Its publicly available implementation [91] is maintained together with Millepede II.
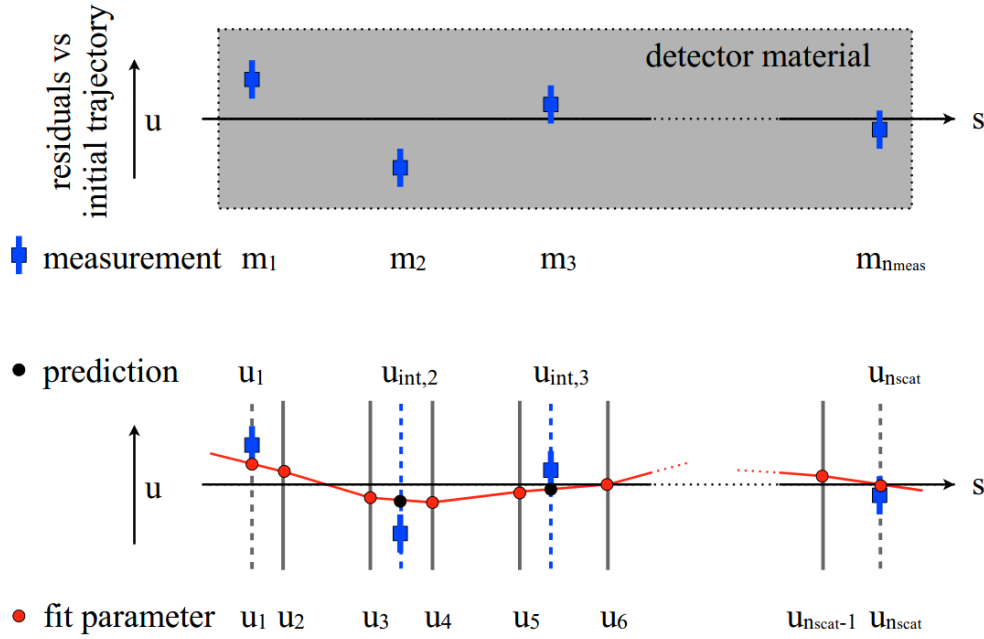
Figure 11.3: Reference trajectory propagating in the detector material and measurement residuals along arc-length s (top) and the GBL trajectory with offsets u_i at each point with a thin scatterer (solid vertical lines) and predictions at the measurement points $u_{\text{int},i}$, interpolated from offsets at scatterers (bottom). The fitted broken trajectory, reducing the residuals, is shown as a red line.

Initial trajectory

The GBL trajectory is composed of individual points. At each point, one defines a local coordinate system (u, v, w). The track is locally described using $\boldsymbol{q} = (q/p, \boldsymbol{u'}, \boldsymbol{u})$ as the track model is linearized around an initial (reference) trajectory, which corresponds to the particle path in the inhomogeneous magnetic field plus the average energy (and thus momentum) loss effects. This trajectory is seeded by the result of a previous fit performed by the standard reconstruction (see Sec. 13.2), and the extrapolation is performed using GENFIT2, which provides the point-to-point Jacobians

$$\boldsymbol{J}_{i \to i+1} = \frac{\partial (q/p, \boldsymbol{u'}, \boldsymbol{u})_{i+1}}{\partial (q/p, \boldsymbol{u'}, \boldsymbol{u})_i} \tag{11.16}$$

and the list of steps in the detector material. The effects of multiple scattering are added to the initial trajectory as depicted in Fig. 11.3. One introduces additional points where the initial path is allowed to be *broken*. These points represent idealized scatterers with zero thickness.

Kinks

Small changes in the track parameters are propagated from point to point using elements of the propagation Jacobians. With three adjacent points having offsets $\boldsymbol{u}_{i-1}, \boldsymbol{u}_i, \boldsymbol{u}_{i+1}$, two slopes can be determined at the central point: $\boldsymbol{u'}_{i,-}$ (left

slope) and $\boldsymbol{u}'_{i,+}$ (right slope). Their difference

$$\boldsymbol{k}_i = \boldsymbol{u}'_{i,+} - \boldsymbol{u}'_{i,-} \tag{11.17}$$

is called a kink. Its expected value is zero[2], and its variance is related to the distribution of radiation length along the path of the particle. An idealized thin scatterer will be a source of variance of track slopes θ_0^2 in each (orthogonal) direction. This variance is diagonal in the co-moving frame. In general, the coordinate frame at the point is not the co-moving frame[3]. In such a case, the covariance for the kinks becomes

$$\boldsymbol{V}_k = \frac{\theta_0^2}{(1 - c_1^2 - c_2^2)^2} \begin{pmatrix} 1 - c_2^2 & c_1 c_2 \\ c_1 c_2 & 1 - c_1^2 \end{pmatrix}, \tag{11.18}$$

where $c_i = \boldsymbol{e}_{\text{track}} \cdot \boldsymbol{e}_{\boldsymbol{u}_i}$ are projections of the unit track direction into the unit vectors of the local axes u and v.

Measurements

To obtain the broken trajectory predictions $\boldsymbol{u}_{\text{int}}$ at the points with measurement, the offsets at the scatterers are interpolated[4] as sketched in Fig 11.3. The expected values for these offsets are the measurements $\boldsymbol{m}_i$ with a diagonal[5] measurement covariance $\boldsymbol{V}_{m,i}$. These measurements correspond to the residuals with respect to the initial trajectory in the GBL formalism.

Least Squares Formulation

The parameters that fully describe the broken trajectory are a common correction to the track curvature and all offsets at the scatterers (and offsets at the first and the last measurements, if those have no scatterer):

$$\boldsymbol{x} = (\Delta q/p, \boldsymbol{u}_1, ..., \boldsymbol{u}_{n_{\text{scat}}}).$$

The parameters are obtained by means of the least-squares method, where the residuals from measurements and from kinks are minimized together using the following χ^2 expression:

$$\chi^2(\boldsymbol{x}) = \sum_{i=1}^{n_{\text{meas}}} (\boldsymbol{H}_{m,i}\boldsymbol{x} - \boldsymbol{m}_i)^\intercal \boldsymbol{V}_{m,i}^{-1}(\boldsymbol{H}_{m,i}\boldsymbol{x} - \boldsymbol{m}_i) + \sum_{i=2}^{n_{\text{scat}}-1} (\boldsymbol{H}_{k,i}\boldsymbol{x})^\intercal \boldsymbol{V}_{k,i}^{-1}(\boldsymbol{H}_{k,i}\boldsymbol{x}), \tag{11.19}$$

---

[2]In the case that the initial trajectory is based on the General Broken Lines model itself, non-zero kinks may be inherited from the previous iteration. In this case, the expectation is non-zero (the previously determined kink). This option is implemented in the software but is not used for alignment by default because no improvement was observed with this more detailed and computationally expensive approach.

[3]We construct a co-moving frame in the case of points with only a scatterer. For points with measurement, the natural coordinate system of the measurement is constructed.

[4]Later we will see that at Belle II, the first of the two scatterers in between every two measurements coincides with the measurement, so the interpolation is trivial.

[5]A non-diagonal covariance matrices (for measurement or a scatterer) must be internally diagonalized by GBL with the necessary modification of the measurements/kinks and precisions) [90].

where the (sparse) projection matrices $\boldsymbol{H}_{m,i} = \frac{\partial \boldsymbol{u}_{\text{int},i}}{\partial \boldsymbol{x}}$ and $\boldsymbol{H}_{k,i} = \frac{\partial \boldsymbol{k}_i}{\partial \boldsymbol{x}}$ from the large vector of track parameters $\boldsymbol{x}$ to the measurements and kinks are constructed from the propagation Jacobians, see Ref. [92]. The solution is formulated as a linear equation system

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \tag{11.20}$$

with

$$\boldsymbol{A} = \sum_{i=1}^{n_{\text{meas}}} \boldsymbol{H}_{m,i}^{\intercal} \boldsymbol{V}_{m,i}^{-1} \boldsymbol{H}_{m,i} + \sum_{i=2}^{n_{\text{scat}}-1} \boldsymbol{H}_{k,i}^{\intercal} \boldsymbol{V}_{k,i}^{-1} \boldsymbol{H}_{k,i}$$
$$\boldsymbol{b} = \sum_{i=1}^{n_{\text{meas}}} \boldsymbol{H}_{m,i}^{\intercal} \boldsymbol{V}_{m,i}^{-1} \boldsymbol{m}_i. \tag{11.21}$$

$\boldsymbol{A}$ is a bordered band matrix, as every offset depends only on the adjacent points (the band part with a size of five), and one common curvature correction parameter (bordered part with a border size of one) is included[6] The special structure of this matrix allows for avoiding matrix inversion in the solution of $\boldsymbol{x}$ [90].

In fact, the step of parameter estimation is not necessary for the alignment formulation, as Millepede II performs the solution internally. The ingredients for the local fit in Millepede II are the projection matrices $\boldsymbol{H}$, covariance matrices $\boldsymbol{V}$ and the vector of measurements (residuals) $\boldsymbol{m}$, possibly after diagonalization is performed in case of non-diagonal covariance matrices. GBL has a direct interface to store this information in binary files used by Millepede II.

Composed Trajectories

So far, only trajectories of single particles have been discussed. GBL also supports the concept of composed trajectories which can represent decays of unstable particles. The vector of fit parameters is expanded to contain information on multiple trajectories as

$$\boldsymbol{x} = (\boldsymbol{\omega}, \boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_n), \tag{11.22}$$

where $\boldsymbol{\omega}$ denotes a set of common parameters, which extend the border size of $\boldsymbol{A}$ by $\dim(\boldsymbol{\omega})$. The band part of $\boldsymbol{A}$ is now composed of n blocks, each for one trajectory. The individual trajectories are correlated via the elements of the common bordered part. A transformation from the set of common parameters $\boldsymbol{\omega}$ of the combined trajectory to the local parameters at the first point of every single trajectory

$$\boldsymbol{H}_{\omega,i} = \frac{\partial (q/p, u', v', u, v)_{1,i}^{\intercal}}{\partial \boldsymbol{\omega}} \tag{11.23}$$

must be defined. A small change in the common parameters then propagates to the individual trajectories and the overall vector of track parameters of the combined object as

$$\Delta \boldsymbol{x} = \sum_i \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{q}_{1,i}} \boldsymbol{H}_{\omega,i} \Delta \boldsymbol{\omega}. \tag{11.24}$$

One can constrain such a change by an optional external measurement (residual) $\boldsymbol{m}_\omega$ of the external parameters with covariance $\boldsymbol{V}_\omega$, which leads to an additional contribution to the χ^2 of the form $(\frac{\partial \boldsymbol{\omega}}{\partial \boldsymbol{x}} \boldsymbol{x} - \boldsymbol{m}_\omega)^{\intercal} \boldsymbol{V}_\omega^{-1} (\frac{\partial \boldsymbol{\omega}}{\partial \boldsymbol{x}} \boldsymbol{x} - \boldsymbol{m}_\omega)$. Such an external

---

[6]We are also using a second common parameter for drift chamber tracks: correction to track time. In such a case, the border has a size of two.

measurement can be a prediction from another detector, known invariant mass of the decaying particle, or average interaction point position estimated previously from many tracks.

11.4   Material Parametrization

Up to this point, we have not discussed where to put the additional scattering points in the trajectory, and what value of θ_0^2 in Eq. 11.18 should be used to scale the covariance of the kinks. Let us, for simplicity, consider a situation without a magnetic field and only in two dimensions, as illustrated in Fig. 11.4. We can parametrize the track state by a two-dimensional vector (θ, y) of slope and offset at any arc-length s. Assuming an idealized scatterer of zero thickness with scattering variance θ_i^2 positioned at arc length s_i, a small deflection $\Delta\theta$ will propagate to a small change in the offset Δy at arc-length s as $\Delta y \approx (s - s_i)\Delta\theta$. The covariance matrix propagated to $s > s_i$ will thus have the form

$$\boldsymbol{V}_i(s) = \begin{pmatrix} (\Delta\theta)^2 & \Delta\theta\Delta y \\ \Delta\theta\Delta y & (\Delta y)^2 \end{pmatrix} = \theta_i^2 \begin{pmatrix} 1 & s - s_i \\ s - s_i & (s - s_i)^2 \end{pmatrix}. \qquad (11.25)$$

A realistic material can be understood as a sum of (infinitely) many scatterers, each at a particular arc-length s_i and with a scattering covariance (as seen at arc-length s) given by Eq. 11.25. The total effect is contained in the sum of the individual contributions and can be written as

$$\boldsymbol{V}(s) = \theta^2 \begin{pmatrix} 1 & s - \bar{s} \\ s - \bar{s} & (s - \bar{s})^2 + \Delta s^2 \end{pmatrix}, \qquad (11.26)$$

where we have denoted

$$\theta^2 = \sum \theta_i^2, \quad \bar{s} = \frac{1}{\theta^2} \sum \theta_i^2 s_i, \quad \Delta s^2 = \frac{1}{\theta^2} \sum \theta_i^2 (s_i - \bar{s})^2. \qquad (11.27)$$

These three quantities can be understood as an overall normalization factor of the scattering variance and mean and variance of the density of the material in units of radiation length. By reproducing these three basic properties of the material distribution, equivalent multiple scattering effects can be achieved.
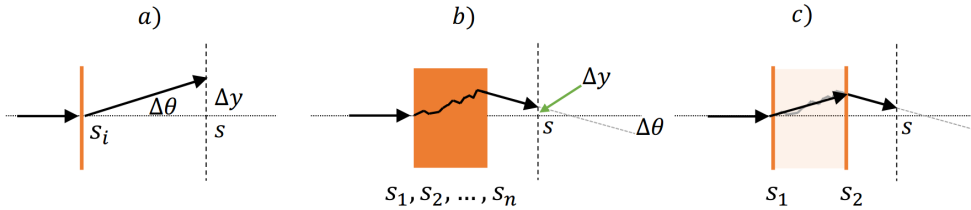

Figure 11.4: Schematic depiction of multiple scattering deflection $\Delta\theta$ from an idealized thin scatterer (a) and from a realistic material distribution (b), also leading to offset Δy at some larger arc-length s. The total effect at s can also be achieved by placing two thin scatterers (with appropriate scattering variances) at specific points s_1, s_2 (c).

In between each two (virtual) measurement planes, the steps retrieved during the extrapolation of the initial trajectory in the detector material with local radiation length X_0 are used to compute the corresponding integrals:

$$X = \int \frac{1}{X_0(s)} ds, \quad \bar{s} = \frac{1}{X} \int \frac{s}{X_0(s)} ds, \quad \Delta s^2 = \frac{1}{X} \int \frac{(s - \bar{s})^2}{X_0(s)} ds \qquad (11.28)$$

with total uncertainty of the track angle due to multiple scattering evaluated using the Highland formula as [93, 57]

$$\theta_0 = \frac{13.6 \text{MeV}}{\beta cp} z \sqrt{X/X_0} [1 + 0.038 \ln(X/X_0)]. \qquad (11.29)$$

The triplet $(\theta_0^2, \bar{s}, \Delta s^2)$, representing the material distribution (a thick scatterer) is translated into a doublet of thin scatterers with positions and variances $(s_1 = 0, \theta_1^2)$ and (s_2, θ_2^2). The position of the first scatterer is conveniently chosen to be exactly at the point of the measurement. From the first scatterer (measurement), the extrapolation is performed by arc-length s_2, and a virtual plane in the co-moving frame is constructed for the second scatterer, which is placed between every two measurements. One can check that for distribution of the form

$$\frac{1}{X_0(s)} = \delta(s)\theta_1^2 + \delta(s - s_2)\theta_2^2,$$

where $\delta(x)$ is the Dirac delta function, the normalization $X = \theta_0^2$, mean $\bar{s}$, and variance Δs^2 in expressions 11.28 are obtained for the following choice of parameters of the two thin scatterers:

$$s_2 = \frac{\Delta s^2 + \bar{s}^2}{\bar{s}}, \quad \theta_1^2 = \frac{\theta_0^2 \Delta s^2}{\Delta s^2 + \bar{s}^2}, \quad \theta_2^2 = \frac{\theta_0^2 \bar{s}^2}{\Delta s^2 + \bar{s}^2}. \qquad (11.30)$$

The values of θ_1^2 and θ_2^2 are used to scale the covariance of the kinks in Eq. 11.18 at the measurement plane and at arc-length distance s_2. This process is repeated for each segment of the trajectory in between two measurements.

11.5 Parametrization of Track Time Offset

Next to the traditional five, GBL allows the inclusion of additional local track parameters. We are using this method to allow for optimization of the track time offset. The event time is determined by combining information from several sub-detectors in a dedicated module, using information from all tracks in the event. The time for track propagation in the detector is taken into account in the CDC hit reconstruction. The track arrives (if it goes from IP) later into the outer layers, and this delay is corrected in drift time estimation and subsequent drift length and residual determination. If the track time relative to the event time changes by Δt_0, the prediction in the only sensitive local u-coordinate of every CDC hit i on the track changes by

$$\Delta u_i = -s^i v_{\text{drift}}^i \Delta t_0 \qquad (11.31)$$

where v_{drift}^i is the estimate of the drift velocity from the $x - t$ relation used to reconstruct hit i, and s^i is a sign factor determining whether the track passes on the left or right side of the wire, which produced this hit.

Therefore an additional local derivative $\frac{\partial u_i}{\partial \Delta t_0} = -s^i v^i_{\text{drift}}$ is provided to the GBL track model construction for each point with a CDC measurement, which allows us to correct for this single parameter per each track (or decay) during the alignment solution.

This correction is only included for drift times in the range $(20, 200)$ ns and for drift velocity smaller than 10^{-2} cm/ns. These cuts were determined in studies of wire time calibration, and their primary purpose is to protect against poorly described data entering the alignment procedure. However, a small dedicated test did not reveal any impact on alignment parameters discussed in this thesis when these cuts are removed, as bad data are generally rejected in Millepede II automatically.

11.6 Parametrization of Decays

Tracks originating from decays of unstable (or virtual) particles provide additional knowledge about correlations of the emerging particles' properties. Namely, we know they must come from exactly the same point in space—the decay vertex. If this decay vertex corresponds to the luminous region, its position (with its uncertainty) can be introduced as a constraint. The coordinates of the IP can even be added to the alignment parameters.

Furthermore, invariant mass formed from the momenta and energies of the particles should match the mass of the primary particle (or invariant mass of beams $\sqrt{s}$ for di-muons). In the case of two-body decays, the originating tracks must be back-to-back in the rest frame of the mother particle. This information then could be exploited in alignment. However, the invariant mass is known only when no significant initial-state (ISR) of final-state radiation (FSR) occurs. If this is not the case, some four-momentum can be lost due to undetected (or ignored) photons. Thus using the known invariant mass (or even primary four-momentum) as a constraint might require dedicated studies to ensure no biases are introduced. Another complication is the necessary CM frame transformation, which itself utilizes quantities (like calibrated beam boost vector) that are only known to limited precision and can be time-dependent.

To introduce kinematic constraints and possibly external measurements like known invariant mass, we implemented the support for two-body decays of type $V^0 \to f^+ f^-$, where the mother V^0 particle has mass M and the daughter particles have mass m. As this feature is not used anywhere in this thesis nor data production, we leave out the details of this topic. We just note that the set of common parameters is $\boldsymbol{\omega} = (x, y, z, p_x, p_y, p_z, \theta, \phi, M)^\mathsf{T}$, where x, y, z are coordinates of the decay vertex, p_x, p_y, p_z is mother momentum (in laboratory frame) and θ and ϕ are angles of the momenta of the daughters in the rest frame of the mother. Four-momentum conservation then implies a decay model which correlates the daughter momenta and vertices $(x, y, z, p_x, p_y, p_z)_\pm$ in the laboratory frame to $\boldsymbol{\omega}$. We followed the formalism in Ref. [94] and used an analytical derivative of the decay model. We, however, still deem two-body decay parametrization and related features to be experimental.

The daughter particles' trajectories need to be included in a single combined GBL trajectory fit object to include correlations among the two tracks in the

alignment problem, see Sec. 11.3. For that purpose, we have to define the vector $\boldsymbol{\omega}$ of the parameters of the combined object and projections $\boldsymbol{H}_{\omega,i}$ for each daughter trajectory.

Before the `MillepedeCollector` module (see Sec. 13.4), the candidate decay particle list has to be created and fitted with a vertex fitter, followed by an update of the daughter momenta and positions. After this vertex fit, the common vertex (now exactly the same for both tracks) is known and can be used as a seed for GBL reference trajectory propagation by inserting it as the first point of every single trajectory. This is particularly important as the individual trajectories typically start at the first PXD layer, which is at a larger radius than the beam pipe, usually representing the major material contribution seen along the path of the tracks. By starting from the common vertex, usually inside the beam pipe, its material is then taken into account in each individual GBL trajectory.

We construct the local measurement system for each trajectory with axes $\boldsymbol{u} = (u_x, u_y, u_z)^\intercal$, $\boldsymbol{v} = (v_x, v_y, v_z)^\intercal$ and a normal $\boldsymbol{n}$ at the point of the common fitted vertex $\boldsymbol{v}^{\text{vtx}}$ such that the normal coincides with the track direction. This defines a local frame at each initial point.

In case only the geometric constraint for the common track origin is required, the set of common parameters that fully defines the state of (for example) two daughter tracks at the vertex is

$$\boldsymbol{\omega} = (x, y, z, q/p_{(1)}, u'_{(1)}, v'_{(1)}, q/p_{(2)}, u'_{(2)}, v'_{(2)})^\intercal, \tag{11.32}$$

where x, y, and z are coordinates of the vertex, $q/p_{(i)}$ is curvature and $u'_{(i)}$ and $v'_{(i)}$ are track slopes at the vertex for daughter i. The implementation allows for an arbitrary number of daughters, but only two daughters have been used so far in practice.

The projections from the common set of parameters to the local parameters of each trajectory at the first point, defined in Eq. 11.23, then require as input the matrices

$$\frac{\partial(u, v)^\intercal}{\partial(x, y, z)^\intercal} \tag{11.33}$$

for each first point of each trajectory, discussed further in Sec. 12.5. These are, in fact, obtained from the transformations $\frac{\partial(q/p, u', v', u, v)^\intercal}{\partial(x, y, z, p_x, p_y, p_z)^\intercal}$, regularly used in GEN-FIT2 to convert track states between local and global frames.

11.7 Alignment Data Samples

A rich topology of tracks used for alignment is crucial to suppress possible weak modes. Our alignment software framework supports a wide variety of inputs, from simple and cosmic ray tracks to two-body decays with invariant mass or even primary four-momentum constraint[7]. These have been implemented before the first data with somewhat optimistic expectations on the machine luminosity. In the future, some of the advanced features could become interesting but will require more thorough testing. Instead, for practical reasons, the list of data sample categories is limited to the following three basic ones:

---

[7]A preliminary version of the primary four-momentum constraint was implemented before the new interface to the beam boost, and invariant mass calibration was introduced. So it will require a minor update to be usable for real data.

- cosmic tracks in a magnetic field,

- single tracks from IP or outside of IP (off-IP), and

- IP-constrained decays.

The three most important of them are illustrated with real recorded events using the Belle II event display in Fig. 11.5.
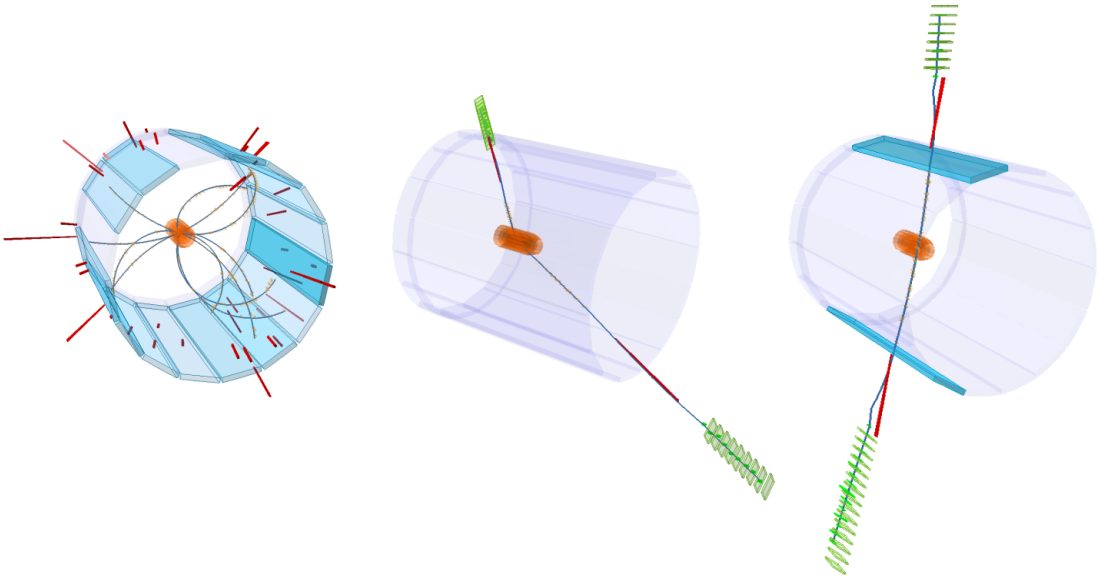

Figure 11.5: Event display of a typical event from a hadron (left), di-muon (middle), and cosmic (right) skim recorded by the Belle II detector. The tracks are fitted with GBL. The blue bars are TOP modules that enclose the CDC. The CDC wires are not shown. The red spikes represent energy deposits in the ECL crystals, and the green boxes are KLM hits.

Cosmic ray data without a magnetic field are not listed. This is usually a primary alignment sample and was used a lot for alignment at the start of the data-taking, see Chapter 14. But tight schedules of the data-taking resulted in a very limited amount of data without a magnetic field. Only a couple of hours are usually devoted to taking this data at the start and end of each data-taking period, i.e., four times a year. In addition, CDC calibration is not being derived automatically for this data. Thus there is no recent data with a good calibration to show in this section, and in turn, for the alignment itself.

The IP-constrained decays are entirely represented by $e^+e^- \rightarrow \mu^+\mu^-$ events. These can only be used in later alignment stages when the detector is already well-calibrated. Only then a reliable time-dependent IP calibration can be provided to be used as a (possibly floating) constraint. Nevertheless, the di-muon events are also used in the initial stages, only as single muon tracks.

Other single tracks are usually taken from a hadronic sample with multi-track events. All tracks are used in this case. In addition, tracks originating outside of the interaction (off-IP events) are mixed into the sample.

The basic data samples collected in specialized skims, and their MC equivalents will be briefly described in the following subsections. The comparison to

MC simulation is important as we want to do some alignment performance estimates based on simulation studies. The aim is, however, not to match data and simulation as close as possible. We even do not have an established simulation for the off-IP events. But as long as the tracks are reconstructed correctly, more information and track topologies should only improve the alignment. Thus, in general, one should only check that MC covers at least the same phase space as the data. A possible issue may be if one would, for example, assume that cosmic tracks with a large d_0 are recorded, while, in fact, these would be absent, e.g., due to some trigger limitations. Then MC could possibly overestimate alignment performance. The general philosophy of all alignment performance studies is to provide conservative estimates for physics analyses, with upper bounds on possible systematics. Thus, richer topology (e.g., momentum spectrum for cosmic rays) in real data samples with respect to MC simulation is not an issue, in contrast to physics analyses, where one naturally aims for as close data/MC correspondence as possible.

11.7.1   Cosmic Ray Tracks

Cosmic ray muons traversing the detector are essential for a reliable alignment. They connect opposite halves of the detector, in contrast to tracks coming from the IP. Their momentum spectrum goes far beyond what is possible in collision events which makes them insusceptible to multiple scattering. In addition, they are not correlated to the IP position, cover a wide range of phase space, and in general, break the symmetry of events from the IP. On the other hand, there are practically no horizontal tracks in the underground experimental hall, and the angular asymmetry of the cosmic rays can cause issues in the alignment (if not combined with other samples).

In the early stages of data collection, cosmic ray data were recorded in between luminosity runs. Prior to this, cosmic rays were the only available data for alignment. However, due to the erratic availability of these data, many concerns arose. Since Autumn 2019, a dedicated cosmic trigger and corresponding skim have been introduced. In this skim, the cosmic ray tracks recorded during the collisions are preserved and separated for further calibration and analysis. This ensures all alignment data is recorded at the same time and therefore has the same misalignment.

The cosmic software trigger looks for events with exactly two reconstructed charged tracks. For each track, the related ECL cluster energy E corresponding to the photon hypothesis is accumulated. For each track, it requires $p_t > 0.5$ GeV/c, $E < 1$ GeV (to suppress electrons that have a large energy deposit in the ECL), and z_0 and d_0 in the region of a well understood magnetic field. With a condition

$$z_0 < 57 \text{ cm or } |d_0| > 26.5 \text{ cm} \tag{11.34}$$

the forward region near the beamline with the complex field of the QCS and the compensating magnets is avoided. Finally, the two tracks must be matched in the polar and azimuthal angles within two degrees. As the real data sample is still contaminated by tracks from collisions, an additional requirement $|z_0| > 0.4$ cm is used beyond the skim requirements.
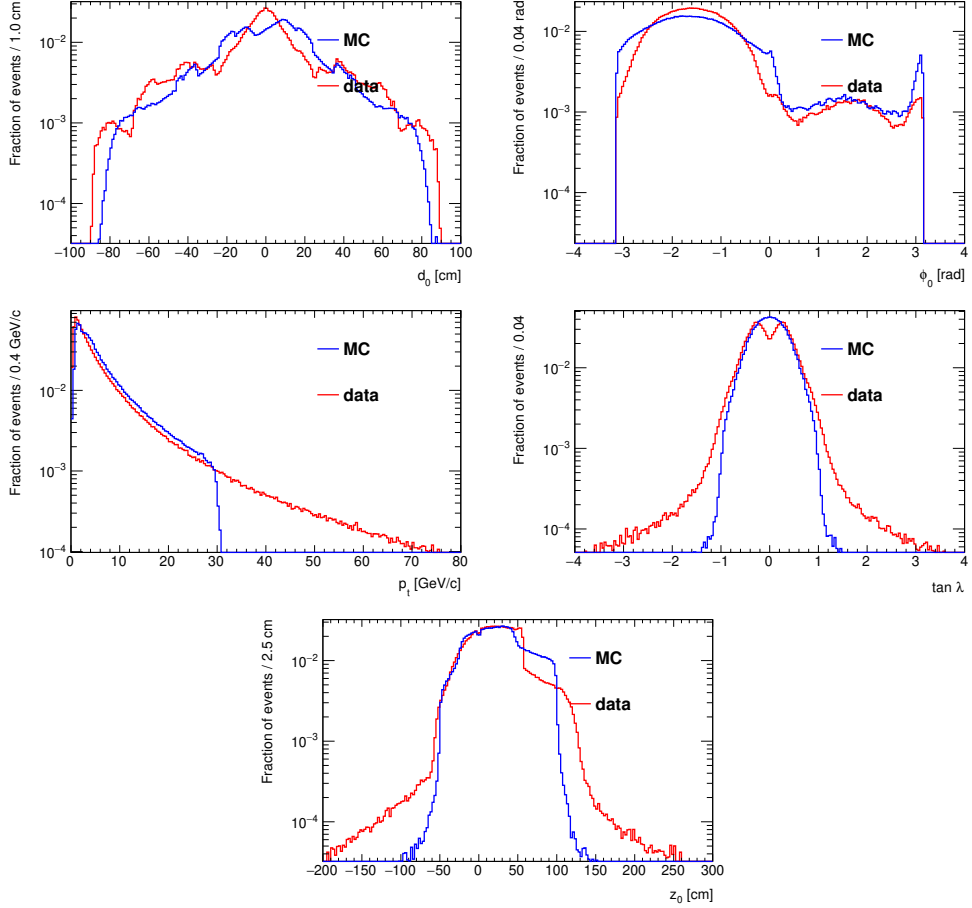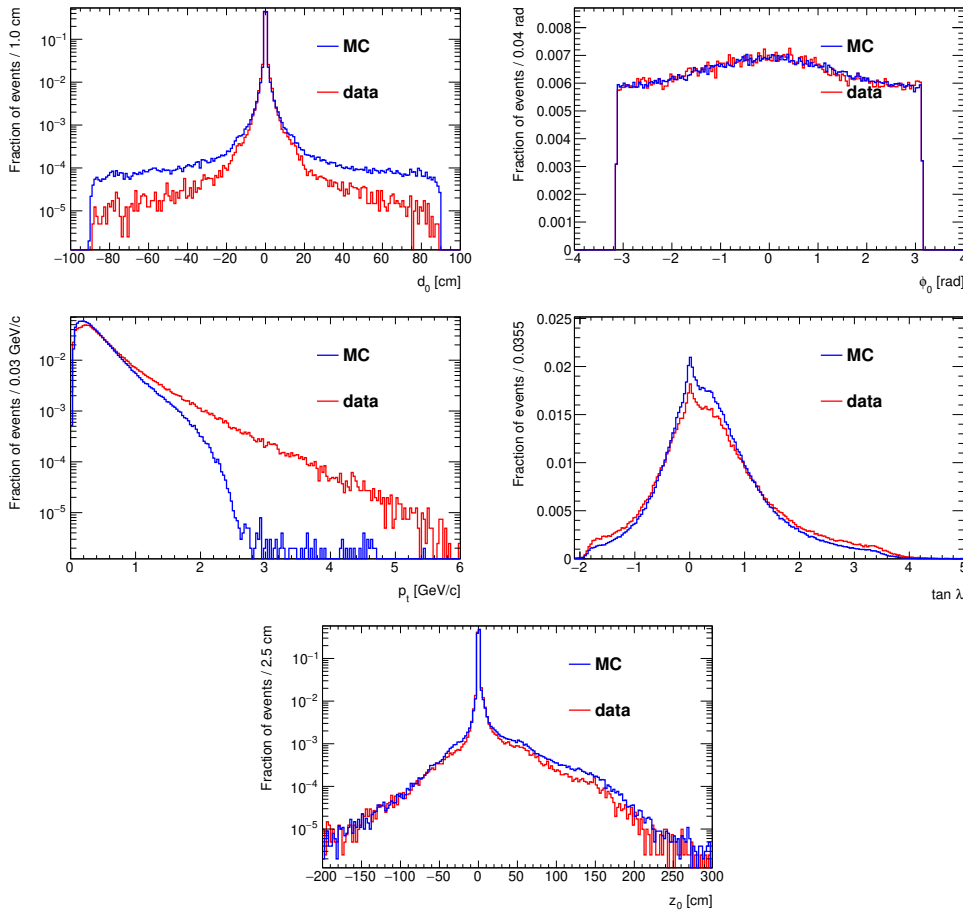
Figure 11.6: Distributions of helix parameters d_0, ϕ_0, p_t, $\tan\lambda$ and z_0 for the cosmic ray simulation (blue) and cosmic data skim (red).

The initial reconstruction of cosmic ray tracks for trigger output calculation on the HLT is provided by the standard reconstruction chain, and thus cosmic particles are reconstructed as two separate tracks, split at the point of the closest approach to the origin. For alignment, these tracks are always used as merged and correctly oriented into a single list of hits, later used in the fits and for GBL trajectory construction.

For MC simulations, two generators for cosmic ray events are available in the Belle II software packages. A more realistic and advanced CRY [95] generator was used for some studies, but it was generally not used for any alignment tests because of its long-standing issues and complicated configuration. Instead, a generator imported from Belle is generally used, which is more efficient and sufficiently realistic for our purposes.

The helix parameters of the reconstructed merged cosmic tracks are compared to the simulation in Fig. 11.6. The phase space observed in the data is well covered by the MC simulation, although some very high momentum tracks are missing. The most important feature for alignment is the very wide acceptance of the triggers in z_0 and d_0. On the other hand, the spectrum of angles is quite limited. Note that the region with $\phi_0 > 0$ corresponds to tracks for which the orientation becomes ambiguous as they do not come from top to bottom but almost horizontally, entering and exiting primarily in the upper half of the CDC.

11.7.2 Hadron Skim and Generic $B\overline{B}$ MC

A hadronic event candidate may be easily identified by the presence of many charged tracks. In addition, a so-called Bhabha veto is imposed. The hadron skim at Belle II [96] contains $q\overline{q}$ continuum, $\tau^+\tau^-$ events, and $B\overline{B}$ events. It is the baseline for most physics analyses. While this sample contains a lot of low momentum tracks, which are generally not considered interesting for alignment, there is a priori no reason to reject any correctly reconstructed and modeled trajectory. For MC studies, we are only using generic $B\overline{B}$ events, which have, on average, an even softer spectrum of particles from B meson decays. The hadronic skim and our default MC to simulate it is compared in Fig. 11.7. As we did not observe any issues on MC with the inclusion of the low-momentum tracks, we do not impose any additional selections. In fact, in the initial stages of the experiment, we simply used all reconstructed charged tracks. Later, it became necessary to separate the events into individual skims to apply pre-scales for processes with very high yields, like Bhabha scattering.

Figure 11.7: Distributions of helix parameters d_0, ϕ_0, p_t, $\tan\lambda$ and z_0 for the generic $B\overline{B}$ simulation (blue) and hadron data skim (red).

11.7.3 Off-IP Tracks

Tracks stemming from accelerator operations but not originating from the luminous region have many different sources and are in most cases only considered to be a background. One example is beam-gas events, where the beam interacts with residual gas molecules in the beam pipe. Another example is beam halos that interact with the material of the beam pipe or regions further upstream or downstream of the beamline. Also, secondary particles from such interactions produce such off-IP tracks.
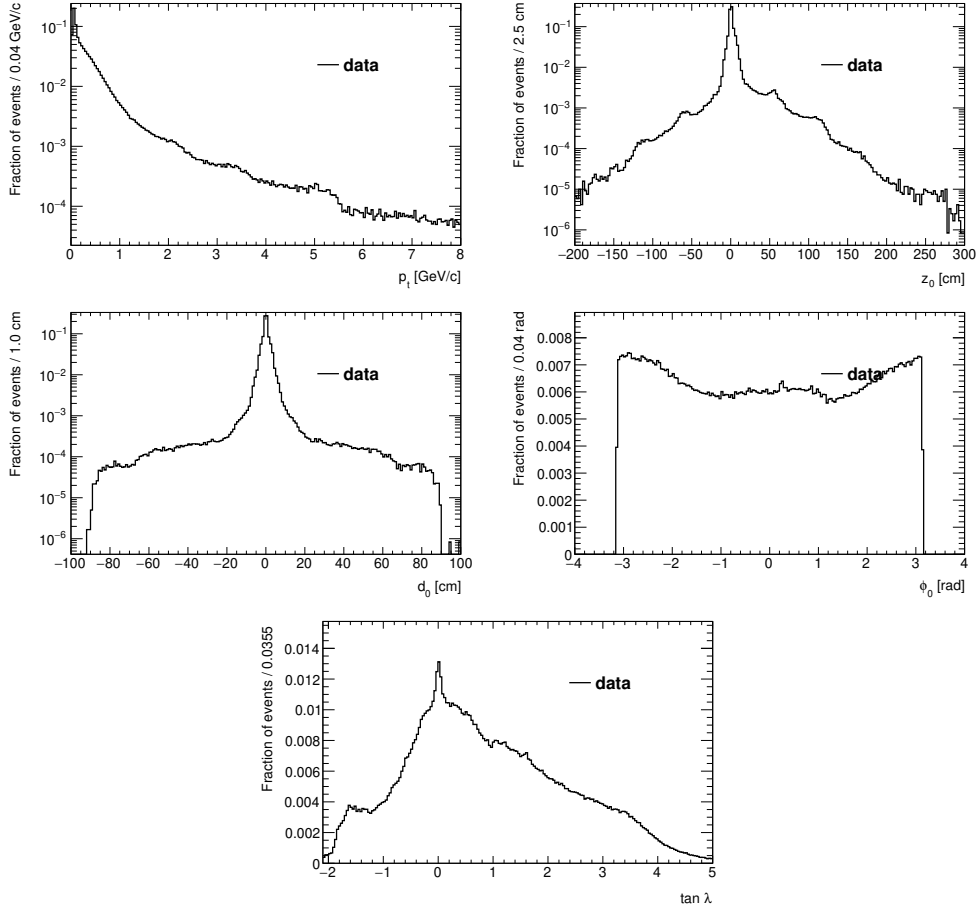
Figure 11.8: Distributions of helix parameters p_t, z_0, d_0, ϕ_0 and $\tan\lambda$ for the off-IP data skim.

Following the alignment philosophy to utilize as much (reliable) information as possible, we include off-IP events into the alignment data sample, although no definitive positive effect was ever observed. On the other hand, no tensions have been found when this sample is added/removed. There is no readily available simulation of the combination of such backgrounds that could be easily used. In addition, also the data skim itself may be further optimized. The distributions of helix parameters from this skim are shown in Fig. 11.8. The software HLT decision requires at least one track in the event to have ≥ 3 SVD hits and ≥ 20 CDC hits and to satisfy $|z_0| > 2$ cm. Currently, also criterion $|d_0| > 1$ cm is included, which is, however, not optimal and shall be removed.

Comparing Fig. 11.7 and 11.8, one can see a large overlap between the hadron and off-IP skims (also seen in average high track multiplicity of the off-IP events),

still mostly populated around the IP. In the future, this sample should be made more orthogonal to the hadron sample. Note that in the z_0 distribution, one can identify so-called hot spots. These are special places where the beamline cross-section is limited by structures providing detector shielding and beam collimation. Secondary particles from these material interactions then *shine* into the detector. These structures can be identified e.g. around $z_0 \sim -120, -70, 60$ and 110 cm.

11.7.4 Di-muon Events

The process $e^+e^- \to \mu^+\mu^-$, called di-muon or $\mu\mu$ events, is an essential channel for many calibrations with a clear experimental signature. In the preliminary stages of the alignment, the two muon tracks are used individually, as usually, a fine calibration of the IP is not yet available. In later alignment stages, the tracks are combined into a single object to utilize the knowledge of their common origin and its location and uncertainty in an IP-constrained fit.

Currently, the so-called tight di-muon skim is used for all alignment productions. Its previous version was more efficient but significantly contaminated, especially by electrons from Bhabha events, whose tracks pass through the gaps between the barrel and the end-caps of the ECL and can thus mimic a di-muon process [96]. The remnants of this problem can still be observed in the angular distributions of tracks from the di-muon sample, which are compared to MC simulation in Fig. 11.9. Note the peaks in $\tan\lambda$ at $\sim 30°$ and $\sim 130°$, which exactly correspond to the gaps, where the Bhabha (electron) veto is not efficient. Because in the current alignment, only the geometric information is used, and no kinematic assumptions are made on the two tracks, the small electron contamination is not expected to cause any issues as the tracks still originate from the primary luminous region.
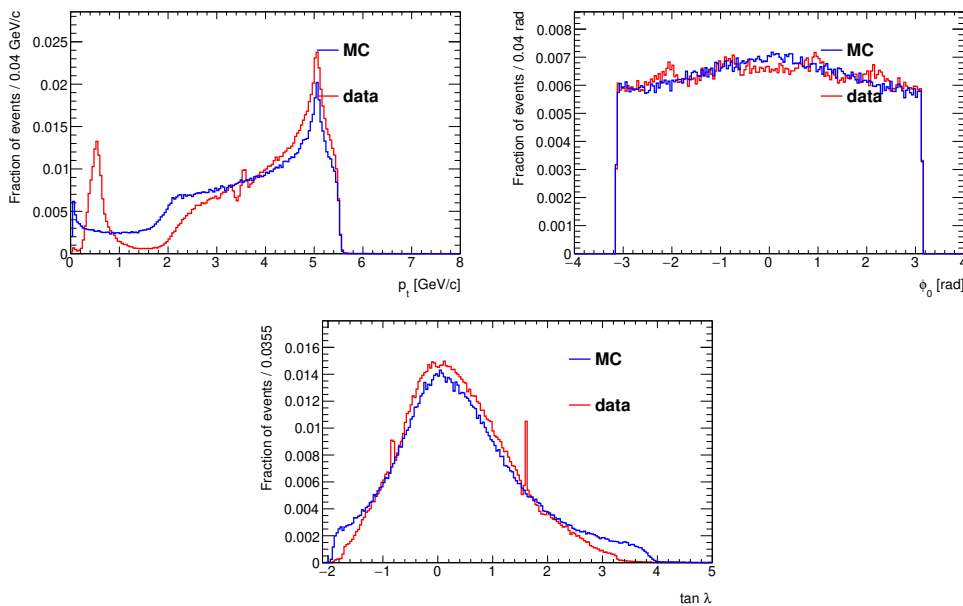
Figure 11.9: Distributions of helix parameters p_t, ϕ_0 and $\tan\lambda$ for di-muon simulation (blue) and tight di-muon data skim (red).

For the di-muon skim, two tracks with $|d_0| < 2$ cm, $|z_0| < 4$ cm, associated

deposited ECL cluster energy $E < 0.5$ GeV, $p_t > 0.2$ GeV/c and CM momentum $p^\star > 0.5$ GeV/c are required. The total deposited ECL energy must be less than 2 GeV, and the polar and azimuthal angles of the two tracks in the CM frame must match within 10°.

For MC, the KKMC generator is being used, and no further selections are imposed on the reconstructed events as there is naturally no contamination from other backgrounds like Bhabha events.

11.8 Selection of Hits and Tracks

Only limited pre-selection of hits and tracks is performed. For each hit in the CDC, we require a converged DAF fit (see Sec.13.2) with reasonable weights assigned to left / right measurements by the standard reconstruction and remove bad hits from tracks. We reject tracks with more than 80% of removed CDC hits. To avoid region where the description of the magnetic field is not reliable, we only accept cosmic ray tracks which satisfy $z_0 < 57$ cm or $|d_0| > 26.5$ cm. Finally, we only use tracks with goodness-of-fit p-value greater than 1×10^{-5} for data (and 1×10^{-4} for MC). No further selections are used. We rely on the internal outliers down-weighting by Millepede II. With these conditions, about 1% (3%) of the remaining tracks provided to Millepede II are rejected on MC (data).

Alignment Parametrization, Derivatives, and Constraints

In this chapter, we properly define all alignment parameters and derive analytical expressions for the derivatives of the residuals with respect to them, $\partial r_{ij}/\partial \boldsymbol{a}$, as defined in Sec. 11.1. Following the Millepede II conventions, we call these global derivatives; in contrast to the local derivatives of the residuum with respect to the track model parameters, automatically provided from a constructed GBL trajectory, discussed in Sec. 11.3. At the end of this chapter, the linear equality constraints on the global parameters will be discussed.

Although similar derivations can be found in literature [81, 74, 97], which had been many times our inspiration, we are not aware of a fully consistent set of results to which we could refer. The true value of the derivations in this chapter is that most have been validated with simulations and real experiment data and cross-checked with the software implementation.

12.1 Global and Local Frames

Any sensitive element records the best estimate of the track intersection position in its natural local coordinate system, while the coordinates of, for example, a pixel or strip used as input for track reconstruction are given in the global reference frame. To translate from pixel or strip coordinates to a point in global coordinates, one needs to know the position and orientation of the sensor itself.

If the local frame origin in the global frame is $\boldsymbol{r_0}$ and the local frame is rotated with respect to the global frame by rotation matrix $\boldsymbol{R}$, a point in the global frame $\boldsymbol{r}$ can be obtained from the point in the local frame $\boldsymbol{q}$ as

$$\boldsymbol{r} = \boldsymbol{Rq} + \boldsymbol{r_0} \tag{12.1}$$
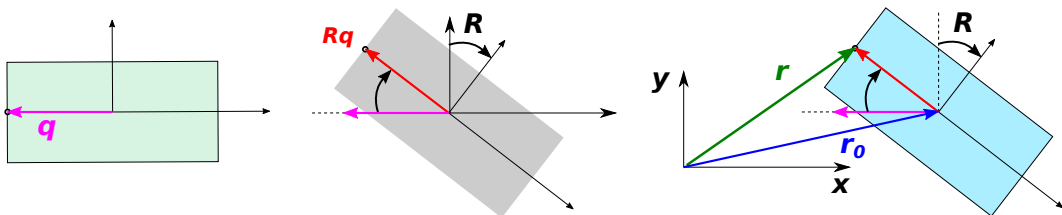
Figure 12.1: Transformation of a point with position vector $\boldsymbol{q}$ in the local reference frame (left) to a point $\boldsymbol{r}$ in global reference (right) is performed by rotation in the local frame by $\boldsymbol{R}$ (middle) followed by a shift of the local frame origin by $\boldsymbol{r_0}$.
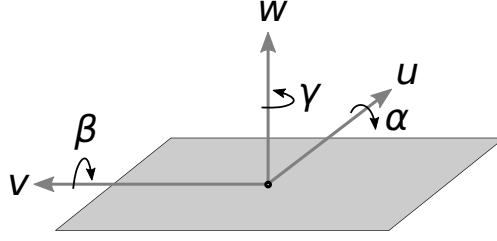
Figure 12.2: Definition of local frame axes u, v, w and rotations α, β, γ.

as shown in Fig. 12.1. In fact, the transformation between the local and global systems are so common that a unified interface is created for the VXD sensors, which stores the combined transformations $\boldsymbol{T}(\boldsymbol{R}, \boldsymbol{r_0})$ for each sensor.

In practice, the alignment consists of finding corrections to these transformations. These corrections are usually expressed in the local system, such that under a rotation $\boldsymbol{\Delta R}$ and shift[1] $\boldsymbol{\Delta q}$ a point $\boldsymbol{q}$ changes to

$$q \to q' = \boldsymbol{\Delta R q} + \boldsymbol{\Delta q}. \tag{12.2}$$

Consequently, in the global frame, the updated point from Eq. 12.1 reads

$$r \to r' = \boldsymbol{R}(\boldsymbol{\Delta R q} + \boldsymbol{\Delta q}) + \boldsymbol{r_0}. \tag{12.3}$$

At this point, a proper definition of the rotation matrices, parametrized by rotation angles, should be given. An important remark is that $\boldsymbol{R}$ and $\boldsymbol{\Delta R}$ are using different angle conventions. For $\boldsymbol{R}$, used for geometry construction and detector placements, Euler angles are used. At the alignment level, the actual definition used to construct $\boldsymbol{R}$ is not relevant as we directly read these placement matrices and extract the elements of the rotation matrix $\boldsymbol{R}$, which are invariant to angular conventions. Truly relevant are the alignment corrections $\boldsymbol{\Delta R}$, which are using Tait-Brian angles. The angles are defined counter-clockwise, and the (active) rotation (in a right-handed coordinate system) is performed by angle α around the first axis, followed by rotation by angle β around the (new) second axis, followed by rotation by angle γ around the (new) third axis. Therefore

$$\boldsymbol{\Delta R} = \boldsymbol{R_3}(\gamma)\boldsymbol{R_2}(\beta)\boldsymbol{R_1}(\alpha), \tag{12.4}$$

where

$$\boldsymbol{R_1}(\alpha) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}, \quad \boldsymbol{R_2}(\beta) = \begin{pmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 0 & 0 \\ -\sin\beta & 0 & \cos\beta \end{pmatrix},$$

$$\tag{12.5}$$

$$\boldsymbol{R_3}(\gamma) = \begin{pmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

If the local system is the same as the global system, $\boldsymbol{R} = 1$ and $\boldsymbol{r_0} = 0$ and the three axes are $(1, 2, 3) = (x, y, z)$. One usually then denotes the angle γ as

---

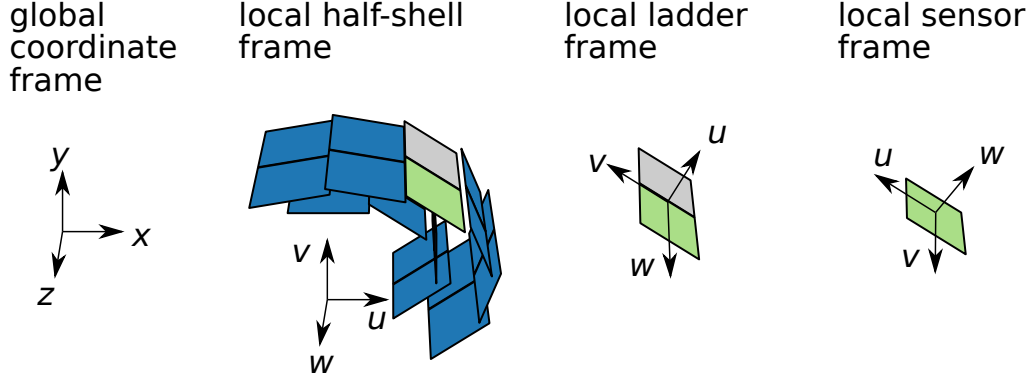[1]In the formalism in this section, $\boldsymbol{\Delta R}$ and $\boldsymbol{\Delta r}$ do not need to be small.

Figure 12.3: Global and local half-shell, ladder, and sensor frames in the VXD hierarchy.

ϕ. We will use this notation later for the alignment of the CDC, where we do not construct local alignment systems. In all other cases, we define the local alignment frame with axes $(1, 2, 3) = (u, v, w)$ and the respective rotation angles as in Fig. 12.2. In the case of silicon sensors, the local system is naturally placed with its origin at the sensor center. The v-coordinate is then chosen to point along the long sensor side, i.e., in the global z-direction (for non-slanted sensors).

12.1.1 VXD Mechanical Hierarchy

In the real implementation of the geometry in ROOT [98] and GEANT4 [99], the transformations by rotations, translations, and scales are expressed as one combined matrix operation, which can be achieved formally by defining 4×4 transformation matrices of general homogeneous transformation $\boldsymbol{T}(\boldsymbol{R}, \boldsymbol{r_0})$. The advantage of this approach is that one can collect the chain of transformations as matrix multiplications [100].

$\boldsymbol{T}(\boldsymbol{R}, \boldsymbol{r_0})$ transforms the local sensor coordinate $\boldsymbol{q}$ to the global point $\boldsymbol{r}$ as

$$\boldsymbol{r} = \boldsymbol{T}(\boldsymbol{R}, \boldsymbol{r_0})\boldsymbol{q}, \tag{12.6}$$

performing the same transformation as in Eq. 12.1. Originally only $\boldsymbol{T}(\boldsymbol{R}, \boldsymbol{r_0})$ was needed for the reconstruction. For the purposes of alignment, we introduced virtual intermediate local frames or **hierarchy levels**. These local frames directly correspond to the geometry placements[2] of the VXD half-shells to the global system $\boldsymbol{T}_{\text{shell}\rightarrow\text{global}}$, ladders into half-shell systems $\boldsymbol{T}_{\text{ladder}\rightarrow\text{shell}}$, and sensors into ladders' systems $\boldsymbol{T}_{\text{sensor}\rightarrow\text{ladder}}$. The orientation of the local frames is depicted in Fig. 12.3. Note that two $90°$ rotations are needed to transform from ladder to sensor frame. Thus we decompose the full sensor transformation in Eq. 12.6 as

$$\boldsymbol{T} = \boldsymbol{T}_{\text{shell}\rightarrow\text{global}}\boldsymbol{T}_{\text{ladder}\rightarrow\text{shell}}\boldsymbol{T}_{\text{sensor}\rightarrow\text{ladder}} \tag{12.7}$$

and can determine corrections in the local systems of the three hierarchy levels, updating the transformation matrices as

$$\boldsymbol{T}_{\text{B}\rightarrow\text{A}} \rightarrow \boldsymbol{T}_{\text{B}\rightarrow\text{A}}\Delta\boldsymbol{T}_{\text{B}\rightarrow\text{A}}(\boldsymbol{\Delta R}_B, \boldsymbol{\Delta r}_B), \tag{12.8}$$

---

[2]The geometry placements are using matrices $\boldsymbol{T}^{-1}$, chained to descend into deeper and deeper levels of the geometry. We store their inverted counter-parts to revert the geometry transformations and get hits in global coordinates.
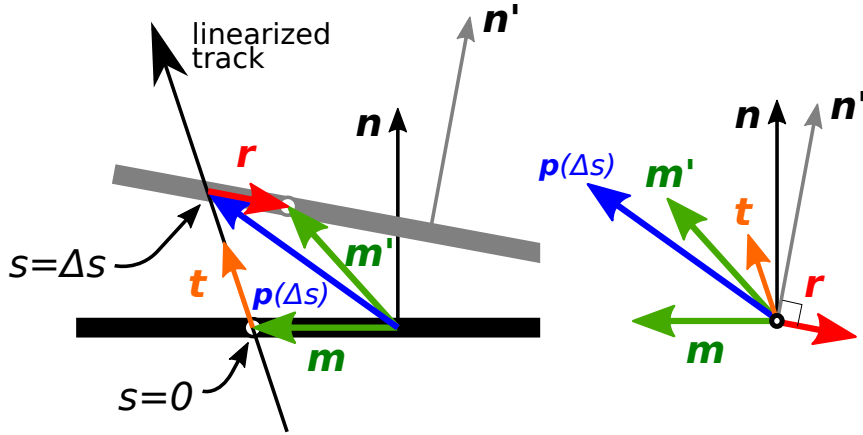
Figure 12.4: Illustration of the geometric situation to formulate the vector equation for the residual $\boldsymbol{r}$ of a distorted measurement $\boldsymbol{m'}$ for a track with direction $\boldsymbol{t}$ defined by the prediction $\boldsymbol{p}$ parameterized by the arc-length s. The normal to the original (moved) measurement plane is denoted as $\boldsymbol{n}$ ($\boldsymbol{n'}$), and the (new) residual as $\boldsymbol{r}$. The quantities in the figure are further described in the text.

where the transformation between each two levels A and B depends on the corrections to rotations and shifts of the lower level B, denoted as $\boldsymbol{\Delta R}_B$, $\boldsymbol{\Delta r}_B$, with $B \in \{\text{sensor}, \text{ladder}, \text{shell}\}$.

12.2 Rigid Body Alignment

In this section, the alignment of shifts and rotations of rigid bodies in different frames is discussed, and the corresponding formulas are derived for global derivatives. This covers all alignment parameters, except sensor deformations, reviewed in the next section.

12.2.1 Residual for a Displaced Measurement

For a small displacement vector $\boldsymbol{\Delta d}$ and small rotation matrix $\boldsymbol{\Delta R}$ applied to a rigid body, a measured point $\boldsymbol{m}$ at some original plane with normal $\boldsymbol{n}$ moves to point $\boldsymbol{m'}$ as

$$m \rightarrow m' = \boldsymbol{\Delta R}\,m + \boldsymbol{\Delta d}. \tag{12.9}$$

The direction of the normal vector is rotated as $\boldsymbol{n} \rightarrow \boldsymbol{n'} = \boldsymbol{\Delta R}\,\boldsymbol{n}$. To make the situation simpler, let us assume the initial track-to-hit residual of the hit is zero (we will be interested only in differences) and the track with unit direction vector $\boldsymbol{t}$ described by prediction $\boldsymbol{p}(s)$, linearized at the original intersection with the plane is expressed as

$$p(s) = m + s\boldsymbol{t} \tag{12.10}$$

such that $\boldsymbol{p}(0) = \boldsymbol{m}$. The situation is depicted in Fig. 12.4. With the new, displaced plane, the track has to be propagated to $s = \Delta s$ such that the new residual

$$r = m' - p(\Delta s) \tag{12.11}$$

188

lies in the new plane (is perpendicular to $\boldsymbol{n}'$). By construction, solving

$$0 = (\boldsymbol{m}' - \boldsymbol{p}(\Delta s)) \cdot \boldsymbol{n}' \tag{12.12}$$

for Δs, one obtains the expression for the residual on the displaced plane

$$\boldsymbol{r} = \boldsymbol{m}' - \boldsymbol{m} - \frac{(\boldsymbol{m}' - \boldsymbol{m}) \cdot \boldsymbol{n}'}{\boldsymbol{t} \cdot \boldsymbol{n}'}\, \boldsymbol{t}. \tag{12.13}$$

The residual depends on the small displacement through both $\boldsymbol{m}'$ (depending on rotations and shifts) and $\boldsymbol{n}'$ (depending only on rotations). One can check that by replacing $\boldsymbol{n}'$ with $\boldsymbol{n}$, we neglect only terms of second order in the small parameters[3]. Thus, in the small misalignment approximation, the residual depends on the displacement only through $\boldsymbol{m}'$ as

$$\left\{ \frac{\partial \boldsymbol{r}}{\partial \boldsymbol{m}'} \right\}_{ij} \approx \delta_{ij} - \frac{t_i n_j}{\boldsymbol{t} \cdot \boldsymbol{n}}, \tag{12.14}$$

where $i, j = 1, 2, 3$. This 3×3 matrix will become useful for the following sections.

12.2.2 Alignment in Local Frame

A special case arises when the alignment parameters influencing a measurement are expressed in the same local frame as the local track frame constructed at the measurement. We use this convenient parameterization for planar detectors. The local system and definition of rotation angles are shown in Fig. 12.2 and Eqs. 12.4, 12.5 .

The track direction in the local coordinates is $\boldsymbol{t} = (u', v', 1)^{\mathsf{T}}/\sqrt{1 + u'^2 + v'^2}$, where $u' = \frac{du}{dw} = \tan \psi$ and $v' = \frac{dv}{dw} = \tan \zeta$. With a normal $\boldsymbol{n} = (0, 0, 1)^{\mathsf{T}}$, the expression 12.14 evaluates to

$$\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{m}'} = \begin{pmatrix} 1 & 0 & -u' \\ 0 & 1 & -v' \\ 0 & 0 & 0 \end{pmatrix}. \tag{12.15}$$

For small rotation angles $\Delta\alpha$, $\Delta\beta$, $\Delta\gamma$ and small displacement vector $(\Delta u, \Delta v, \Delta w)^{\mathsf{T}}$ in the local frame, Eq. 12.9 can be explicitly written using local coordinates for $\boldsymbol{m} = (u, v, 0)^{\mathsf{T}}$ as

$$\boldsymbol{m}' = \begin{pmatrix} 1 & -\Delta\gamma & \Delta\beta \\ \Delta\gamma & 1 & -\Delta\alpha \\ -\Delta\beta & \Delta\alpha & 1 \end{pmatrix} \begin{pmatrix} u \\ v \\ 0 \end{pmatrix} + \begin{pmatrix} \Delta u \\ \Delta v \\ \Delta w \end{pmatrix}. \tag{12.16}$$

For a vector of alignment parameters $\boldsymbol{a} = (\Delta u, \Delta v, \Delta w, \Delta\alpha, \Delta\beta, \Delta\gamma)^{\mathsf{T}}$, one gets

$$\frac{\partial \boldsymbol{m}'}{\partial \boldsymbol{a}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & -v \\ 0 & 1 & 0 & 0 & 0 & u \\ 0 & 0 & 1 & v & -u & 0 \end{pmatrix}, \tag{12.17}$$

---

[3]$\Delta \boldsymbol{R} = \boldsymbol{1} + o(\texttt{small angles})$.

which is combined with Eq. 12.15 using the chain rule to obtain an expression for the global derivative of the local residual $\boldsymbol{r} = (r_u, r_v, r_w = 0)^\intercal$ as $\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{a}} = \frac{\partial \boldsymbol{r}}{\partial \boldsymbol{m'}} \frac{\partial \boldsymbol{m'}}{\partial \boldsymbol{a}}$. Therefore the matrix of global derivatives explicitly reads

$$\frac{\partial (r_u, r_v)^\intercal}{\partial (\Delta u, \Delta v, \Delta w, \Delta \alpha, \Delta \beta, \Delta \gamma)^\intercal} = \begin{pmatrix} 1 & 0 & -u' & -u'v & u'u & -v \\ 0 & 1 & -v' & -v'v & v'u & u \end{pmatrix}. \quad (12.18)$$

The (omitted) last row $\partial r_w / \partial \boldsymbol{a}$ of the matrix is identically zero, as the local w-residual r_w is zero for a planar measurement.

12.2.3 Alignment in Local Hierarchy Frames

The global derivatives for the parameters at the different hierarchy levels are discussed in this section. If the hierarchy levels L_1, L_2 ... L_N are organized as

$$\texttt{global frame} \leftarrow L_1 \leftarrow L_2 \leftarrow ... \leftarrow L_N(\texttt{sensitive elements}) \quad (12.19)$$

one can obtain the transformation between rigid body parameters $\boldsymbol{a}_A = (\Delta u, \Delta v, \Delta w, \Delta \alpha, \Delta \beta, \Delta \gamma)^\intercal_A$ and $\boldsymbol{a}_B = (\Delta u, \Delta v, \Delta w, \Delta \alpha, \Delta \beta, \Delta \gamma)^\intercal_B$ of any two levels L_A and L_B ($A < B$) using 6×6 matrices derived in Appendix D:

$$\frac{\partial \boldsymbol{a}_A}{\partial \boldsymbol{a}_B} = \begin{pmatrix} \boldsymbol{R} & \boldsymbol{DR} \\ 0 & \boldsymbol{R} \end{pmatrix}, \quad \frac{\partial \boldsymbol{a}_B}{\partial \boldsymbol{a}_A} = \begin{pmatrix} \boldsymbol{R}^\intercal & -\boldsymbol{R}^\intercal \boldsymbol{D} \\ 0 & \boldsymbol{R}^\intercal \end{pmatrix}, \quad (12.20)$$

where the $\boldsymbol{R}$ is the rotation matrix of the local system of the level B with respect to the level A. The matrix $\boldsymbol{D}$ is defined as

$$\boldsymbol{D} = \begin{pmatrix} 0 & z & -y \\ -z & 0 & x \\ y & -x & 0 \end{pmatrix} \quad (12.21)$$

where $(x, y, z)^\intercal = \boldsymbol{r_0}$ is the center of the B frame in coordinates of the level A, such that the local-to-'global' (lower-to-upper; $B \rightarrow A$) transformation equivalent to Eq. 12.1 of a local point $\boldsymbol{q}$ in the frame B to point $\boldsymbol{r}$ in the frame A goes as $\boldsymbol{r} = \boldsymbol{Rq} + \boldsymbol{r_0}$. The necessary rotation matrix and position is extracted from the stored geometry transformations $\boldsymbol{T}(\boldsymbol{R}, \boldsymbol{r_0})$, which include sum of previous alignment corrections, as in Eq. 12.8.

While our implementation allows for arbitrary hierarchy (recursively traversed), only $N = 3$ hierarchy levels with $L_1 = $ shell, $L_2 = $ ladder and $L_3 = $ sensor are implemented for the VXD geometry. The desired global derivatives are obtained using the chain rule, which for the half-shells evaluates to

$$\frac{\partial (r_u, r_v)^\intercal}{\partial \boldsymbol{a}_{\text{shell}}} = \frac{\partial (r_u, r_v)^\intercal}{\partial \boldsymbol{a}_{\text{sensor}}} \frac{\partial \boldsymbol{a}_{\text{sensor}}}{\partial \boldsymbol{a}_{\text{ladder}}} \frac{\partial \boldsymbol{a}_{\text{ladder}}}{\partial \boldsymbol{a}_{\text{shell}}}, \quad (12.22)$$

where $\frac{\partial (r_u, r_v)^\intercal}{\partial \boldsymbol{a}_{\text{sensor}}} = \frac{\partial (r_u, r_v)^\intercal}{\partial (\Delta u, \Delta v, \Delta w, \Delta \alpha, \Delta \beta, \Delta \gamma)^\intercal}$ as defined in Eq. 12.18. Similarly, for ladders, we have

$$\frac{\partial (r_u, r_v)^\intercal}{\partial \boldsymbol{a}_{\text{ladder}}} = \frac{\partial (r_u, r_v)^\intercal}{\partial \boldsymbol{a}_{\text{sensor}}} \frac{\partial \boldsymbol{a}_{\text{sensor}}}{\partial \boldsymbol{a}_{\text{ladder}}}. \quad (12.23)$$

Therefore, if ladders and half-shells are aligned together with the sensors, the matrix of global derivatives for each planar measurement will have $6 + 6 + 6 = 18$ columns and 2 rows:

$$\frac{\partial(r_u, r_v)^\intercal}{\partial \boldsymbol{a}} = \left(\frac{\partial(r_u, r_v)^\intercal}{\partial \boldsymbol{a}_{\text{sensor}}} \quad \frac{\partial(r_u, r_v)^\intercal}{\partial \boldsymbol{a}_{\text{ladder}}} \quad \frac{\partial(r_u, r_v)^\intercal}{\partial \boldsymbol{a}_{\text{shell}}} \right) \tag{12.24}$$

For the case of only one-dimensional measurement available from any SVD hit, only the corresponding row is effectively used.

One can also exclude a hierarchy level from the alignment. In particular, one can align only half-shells together with the sensors. In such case, the sensors are (virtually) placed directly into the half-shells, and the necessary transformation is obtained by multiplying the matrices in the chain:

$$\frac{\partial \boldsymbol{a}_{\text{sensor}}}{\partial \boldsymbol{a}_{\text{shell}}} = \frac{\partial \boldsymbol{a}_{\text{sensor}}}{\partial \boldsymbol{a}_{\text{ladder}}} \frac{\partial \boldsymbol{a}_{\text{ladder}}}{\partial \boldsymbol{a}_{\text{shell}}}. \tag{12.25}$$

It should be noted that by introducing the additional alignment degrees of freedom for ladders and half-shells, one has to include a set of linear equality constraints. For example, any movement of a ladder can be canceled by an opposite coherent movement of all sensors in the ladder. So, two sets of parameters correspond to the same alignment configuration. Obviously, this makes the problem underdetermined. To resolve this issue, one has to fix the average movement of the sensors in every ladder such that all correlated movements are absorbed only in the ladder parameters. We discuss the necessary constraints in Sec. 12.6.1.

12.2.4   Alignment in Global Frame

For alignment in the global reference frame, the measurement is expressed as $\boldsymbol{m} = (x, y, z)^\intercal$. The alignment parameters $\boldsymbol{a} = (\Delta x, \Delta y, \Delta z, \Delta \alpha, \Delta \beta, \Delta \gamma)^\intercal$ are directly the shifts and rotations in the global coordinate system. The distorted measurement from Eq. 12.9 evaluates to

$$\boldsymbol{m}' = \begin{pmatrix} 1 & -\Delta\gamma & \Delta\beta \\ \Delta\gamma & 1 & -\Delta\alpha \\ -\Delta\beta & \Delta\alpha & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \end{pmatrix} \tag{12.26}$$

and its derivative with respect to the alignment parameters is

$$\frac{\partial \boldsymbol{m}'}{\partial \boldsymbol{a}} = \begin{pmatrix} 1 & 0 & 0 & 0 & z & -y \\ 0 & 1 & 0 & -z & 0 & x \\ 0 & 0 & 1 & y & -x & 0 \end{pmatrix}. \tag{12.27}$$

While the alignment parameters are expressed in the global system, the alignment input is the change of the local residual $\boldsymbol{r}_l = (r_u, r_v, r_w)^\intercal$ defined in the local track frame at the measurement. If the residual in the global coordinates is $\boldsymbol{r} = (r_x, r_y, r_z)^\intercal$, it can be projected to a local frame with unit axes $\boldsymbol{u} = (u_x, u_y, u_z)^\intercal$, $\boldsymbol{v} = (v_x, v_y, v_z)^\intercal$ and normal $\boldsymbol{n} = \boldsymbol{w} = (w_x, w_y, w_z)^\intercal$, where $||u|| = ||v|| = ||w|| = 1$. One has $r_u = \boldsymbol{r} \cdot \boldsymbol{u}$, $r_v = \boldsymbol{r} \cdot \boldsymbol{v}$, $r_w = \boldsymbol{r} \cdot \boldsymbol{w}$ and thus

$$\frac{\partial \boldsymbol{r}_l}{\partial \boldsymbol{r}} = \begin{pmatrix} u_x & u_y & u_z \\ v_x & v_y & v_z \\ w_x & w_y & w_z \end{pmatrix} = \boldsymbol{R}^\intercal. \tag{12.28}$$

The desired derivatives of the local residual with respect to the alignment parameters in the global reference frame are obtained using the chain rule as

$$\frac{\partial(r_u, r_v, r_w)^\intercal}{\partial(\Delta x, \Delta y, \Delta z, \Delta\alpha, \Delta\beta, \Delta\gamma)^\intercal} = \frac{\partial \boldsymbol{r}_l}{\partial \boldsymbol{a}} = \frac{\partial \boldsymbol{r}_l}{\partial \boldsymbol{r}}\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{m}'}\frac{\partial \boldsymbol{m}'}{\partial \boldsymbol{a}}, \tag{12.29}$$

where for $\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{m}'}$ one should use Eq. 12.14. We will need only some elements of this matrix of derivatives for the alignment of CDC wires and layers.

An equivalent expression can be obtained using the previous result for alignment in the local frame $\partial \boldsymbol{r}_l / \partial \boldsymbol{a}_{\text{local}}$ from Eq. 12.18 and the global-to-local rigid body parameter transformation $\partial \boldsymbol{a}_{\text{local}} / \partial \boldsymbol{a}_{\text{global}}$:

$$\frac{\partial \boldsymbol{r}_l}{\partial \boldsymbol{a}} = \frac{\partial \boldsymbol{r}_l}{\boldsymbol{a}_{\text{local}}}\frac{\partial \boldsymbol{a}_{\text{local}}}{\partial \boldsymbol{a}_{\text{global}}}, \tag{12.30}$$

where

$$\frac{\partial \boldsymbol{p}_{\text{local}}}{\partial \boldsymbol{p}_{\text{global}}} = \begin{pmatrix} \boldsymbol{R}^\intercal & -\boldsymbol{R}^\intercal \boldsymbol{D} \\ 0 & \boldsymbol{R}^\intercal \end{pmatrix} \tag{12.31}$$

is defined in Sec. 12.2.3 (using $A = \texttt{global}$, $B = \texttt{local frame}$) and the rotation matrix $\boldsymbol{R}^\intercal$ is in fact the transformation matrix from Eq. 12.28, already silently denoted above.

12.3 Sensor Surface Deformations

The brief history of alignment of the surface deformations of the silicon sensors is discussed in Chapter 14. The main message is that, in reality, the silicon sensors are not flat but deformed at a similar scale as the thickness of the sensors—about $50\,\mu\text{m}$ for PXD and sometimes over $300\,\mu\text{m}$ for SVD sensors. Neglecting this fact in the reconstruction and alignment can lead to significant degradation of the sensor resolution.

We parametrize each silicon sensor surface by two-dimensional Legendre polynomials up to the fourth order. The one-dimensional Legendre polynomials of interest, shown in Fig. 12.5, are defined as

$$\begin{aligned} L_0(x) &= 1 \\ L_1(x) &= x \\ L_2(x) &= \frac{1}{2}(3x^2 - 1) \\ L_3(x) &= \frac{1}{2}(5x^3 - 3x) \\ L_4(x) &= \frac{1}{8}(35x^4 - 30x^2 + 3). \end{aligned} \tag{12.32}$$

They are orthogonal in the interval $x \in [-1, 1]$. The orthogonality is important for properly separating the different degrees and can be guaranteed in 2D mathematically if the sensor is illuminated uniformly in at least one of the two coordinates.
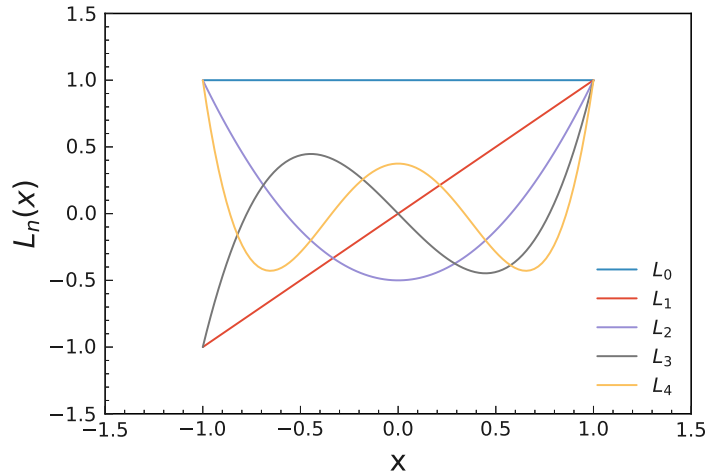
Figure 12.5: Legendre polynomials up to fourth order.

At each point with local coordinates (u, v) on the sensor, we define the normalized, relative sensor coordinates as

$$u^r = 2u/W$$
$$v^r = 2v/L,$$

(12.33)

where W is the width and L is the length of the sensor. At each point, the local w-coordinate ($w = 0$ for a flat sensor) is then generally given as

$$\sum_{i=2}^{n} \sum_{j=0}^{i} p_{i-j,j} L_{i-j}(u^r) L_j(v^r),$$

(12.34)

where n is the order of the parametrization. This expression evaluates explicitly for the third order to

$$
\begin{aligned}
w(u^r, v^r) = {} & p_{20} L_2(u^r) + p_{11} L_1(u^r) L_1(v^r) + p_{02} L_2(v^r) \\
& + p_{30} L_3(u^r) + p_{21} L_2(u^r) L_1(v^r) + p_{12} L_1(u^r) L_2(v^r) + p_{03} L_3(v^r) \\
& + p_{40} L_4(u^r) + p_{31} L_3(u^r) L_1(v^r) + p_{22} L_2(u^r) L_2(v^r) + p_{13} L_1(u^r) L_3(v^r) \\
& + p_{04} L_4(v^r),
\end{aligned}
$$

(12.35)

where the 3+4+5 parameters p_{ij} are the coefficients of the surface decomposition into the Legendre base, shown in Fig. 12.6, and correspond to the actual alignment parameters.

In the simulation, track finding, and extrapolation, the sensors are assumed to be flat. To reduce the measurement residual, we shift the local measurement coordinates during the GENFIT2 measurement construction using a simple geometric correction. The situation is depicted for the u-dimension in Fig. 12.7. This approximate method assumes that

- the track curvature can be neglected, and

- the change of the length of the sensor due to the deformation can be neglected, in the sense that the point H in Fig. 12.7 can be identified with the local coordinate u_h on a theoretically un-deformed sensor by simply moving it along w.
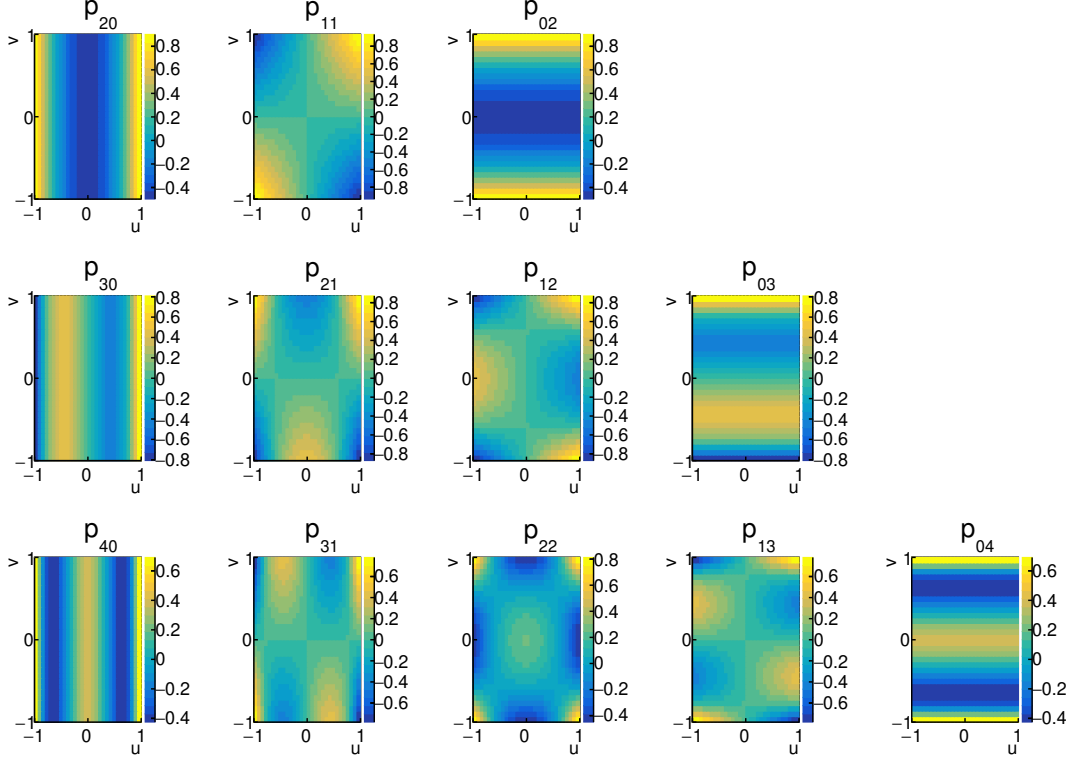
Figure 12.6: Base 2D Legendre polynomials corresponding to the decomposition coefficients and alignment parameters p_{ij}. The rows show from top to bottom quadratic, cubic, and quartic deformations.

At the aligned flat sensor surface, the track is locally parametrized with parameters $(\frac{q}{p}, \frac{du}{dw}, \frac{dv}{dw}, u_p, v_p)$, where the purpose of the subscript in u_p, v_p is to differentiate the estimates from the track prediction and from the actual local hit coordinates (u_h, v_h). For SVD hits with only one coordinate measured, the prediction from the track is used instead of the hit coordinate. The angles of the track with respect to the (flat) sensor surface $\frac{du}{dw} = \tan\psi$, $\frac{dv}{dw} = \tan\zeta$ can be used to correct the local hit positions to first order as

$$
\begin{aligned}
u_h \rightarrow u_h' &= u_h + w(u_h^r, v_h^r)\tan\psi \\
v_h \rightarrow v_h' &= v_h + w(u_h^r, v_h^r)\tan\zeta,
\end{aligned}
\tag{12.36}
$$

where the superscript r denotes the relative normalized local coordinates.

From the above Eqs. 12.35 and 12.36, the derivatives of the corrected track-to-hit residual vector $\boldsymbol{r} = (r_u = u_h' - u_p, r_v = v_h' - v_p)^\intercal$ with respect to any of the surface alignment parameters p_{ij} is

$$
\frac{\partial(r_u, r_v)^\intercal}{\partial p_{ij}} = \begin{pmatrix} L_i(u_h^r)L_j(v_h^r)\tan\psi \\ L_i(u_h^r)L_j(v_h^r)\tan\zeta \end{pmatrix}.
\tag{12.37}
$$
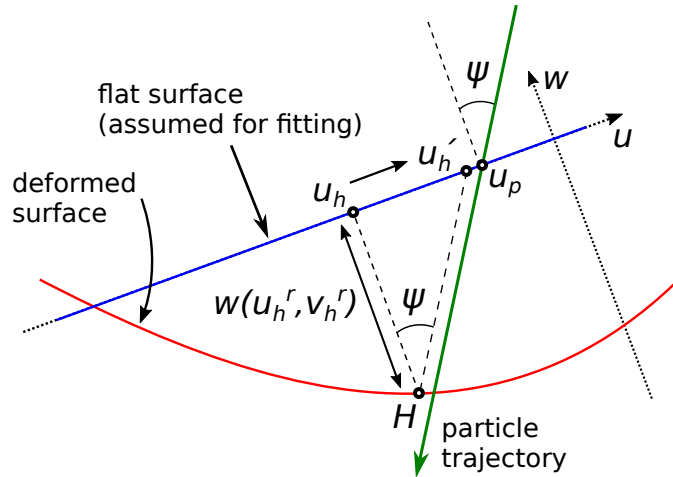
Figure 12.7: A charged track (green) crosses the real deformed sensor surface (red) and creates a nearby hit at point H, with local u-coordinate u_h. The track prediction on the assumed flat sensor surface (blue) has coordinate u_p. By shifting $u_h \to u'_h$, the track-to-hit residual can be reduced using the track incidence angle ψ and the local sensor deformation at the hit position $w(u^r_h, v^r_h)$, using normalized relative coordinates, i.e. $u^r_h = 2u_h/W$, where W is the width of the sensor.

12.4   CDC Layers and Wires

In the CDC, the sensitive elements are the $\sim 14 \times 10^3$ sense wires. The wires are, to a good approximation, straight-line segments[4]. These are defined by two three-dimensional points—one at each wire end at the respective backward or forward CDC end-plate. The results of mechanical surveys of the chamber during construction and installation [21] are used for the initial positions of the wires as displacements (see Sec. 13.6.1). For each wire, the displaced end-point positions are denoted (x^0_b, y^0_b, z^0_b) and (x^0_f, y^0_f, z^0_f) below.

For the purposes of the CDC alignment, we introduced additional degrees of freedom to describe alignment corrections for a higher level of mechanical structures: the layers. Each layer has six alignment parameters $(\Delta x_b, \Delta y_b, \Delta\phi_b, \Delta x_{f-b}, \Delta y_{f-b}, \Delta\phi_{f-b})$. The three former are general shifts and rotation of the layers at the backward end-plate; the three latter are the differences between the forward and backward end-plate shifts and rotations.

The alignment corrections for the individual wires are applied on top of displacements and layer alignment corrections. The track-based alignment cannot be performed for the wire z-positions[5], and thus there are four alignment parameters per each wire: $(\Delta x^w_b, \Delta y^w_b, \Delta x^w_f, \Delta y^w_f)$.

The layer rotation is performed before shifts due to layer and wire alignment,

---

[4]The wire tensions were measured after CDC construction to estimate the effect of gravitational wire sagging in the reconstruction, where for the final estimate, the wire is parametrized as a parabola.

[5]At least not with this basic method. In general, the resolution in z-direction is for CDC worse than mechanical survey precisions, and thus the wire z-positions should be fixed to the best knowledge from the survey.
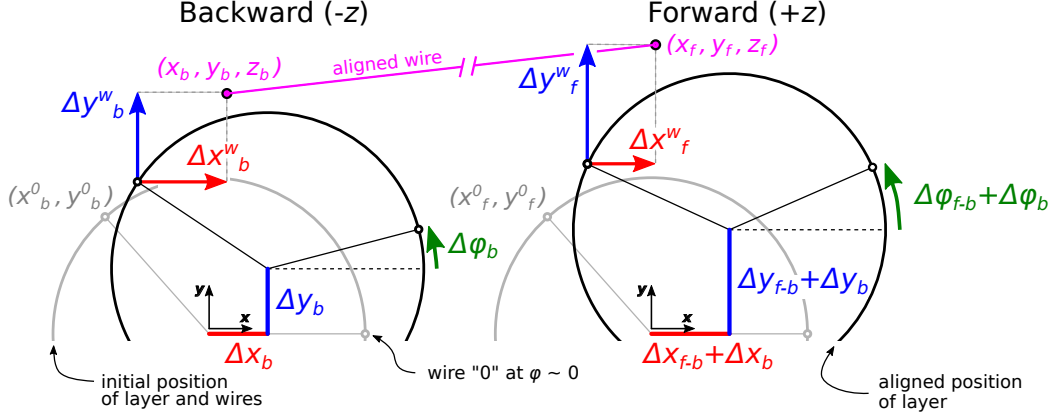
Figure 12.8: CDC layer and wire alignment parameters. The aligned wire (magenta) is obtained from the initial (nominal + displacement) wire and layer positions (gray) after application of shifts and rotations due to the layer alignment parameters $(\Delta x_b, \Delta y_b, \Delta \phi_b, \Delta x_{f-b}, \Delta y_{f-b}, \Delta \phi_{f-b})$, followed by application of the wire shift parameters $(\Delta x_b^w, \Delta y_b^w, \Delta x_f^w, \Delta y_f^w)$ for each wire-end on the backward (left) and forward (right) CDC end-plate.

such that wires are rotated in-layer and not around the origin, although the difference is fully negligible for small layer rotations and shifts. To be explicit, the aligned wire end positions $(x_b, y_b, z_b, x_f, y_f, z_f)$ are defined as[6]:

$$x_b = +x_b^0 \cos \Delta \phi_b - y_b^0 \sin \Delta \phi_b + \Delta x_b + \Delta x_b^w$$
$$y_b = +x_b^0 \sin \Delta \phi_b + y_b^0 \cos \Delta \phi_b + \Delta y_b + \Delta y_b^w$$
$$z_b = +z_b^0$$

$$x_f = +x_f^0 \cos(\Delta \phi_b + \Delta \phi_{f-b}) - y_f^0 \sin(\Delta \phi_b + \Delta \phi_{f-b}) + \Delta x_b + \Delta x_{f-b} + \Delta x_f^w$$
$$y_f = +x_f^0 \sin(\Delta \phi_b + \Delta \phi_{f-b}) + y_f^0 \cos(\Delta \phi_b + \Delta \phi_{f-b}) + \Delta y_b + \Delta y_{f-b} + \Delta y_f^w$$
$$z_f = +z_f^0$$

(12.38)

In the track parametrization at a given wire hit, the local coordinate system is constructed as depicted in Fig. 11.1. Thus the wires are not aligned in the same coordinate system as the one used for the local parametrization, unlike planar sensors. The local track frame is constructed at each wire hit such that the local (insensitive) v-coordinate points from backward to forward wire-end and the normal points along the track direction. Therefore the (only) sensitive coordinate is u.

For the vector $(\Delta x_b, \Delta y_b, \Delta \phi_b, \Delta x_{f-b}, \Delta y_{f-b}, \Delta \phi_{f-b})^\intercal$ of layer alignment parameters, the first three affect the residual equivalently on the backward and forward end-plate. The influence of the latter three parameters is proportional to

---

[6]Note that the angles $\Delta \phi_b$, $\Delta \phi_{f-b}$ are by mistake defined clock-wise in the software, same as rigid body angles. This wrong sign is then compensated by a wrong sign of the global derivatives with respect to rotations. The same sign flip is also present for VXD alignment. However, this non-standard definition (which gets applied after reading constants from payloads) has been kept for several years for backward compatibility. One must keep this in mind when manually interpreting the raw parameter values in the alignment payloads.

the relative z-position of the hit. Therefore, using the elements of the full general matrix of derivatives $\frac{\partial r_l}{\partial p}$ from Eq. 12.29, the derivative of the local u-residual can be written as

$$\frac{\partial r_u}{\partial(\Delta x_b, \Delta y_b, \Delta\phi_b, \Delta x_{f-b}, \Delta y_{f-b}, \Delta\phi_{f-b})^{\mathsf{T}}} =$$
$$= \begin{pmatrix} \frac{\partial r_u}{\partial\Delta x} & \frac{\partial r_u}{\partial\Delta y} & \frac{\partial r_u}{\partial\Delta\gamma} & z^r\frac{\partial r_u}{\partial\Delta x} & z^r\frac{\partial r_u}{\partial\Delta y} & z^r\frac{\partial r_u}{\partial\Delta\gamma} \end{pmatrix}$$

(12.39)

where $z^r \in [0, 1]$ is the relative normalized z-coordinate of the hit with z-prediction z_p from the track, defined as

$$z^r = \frac{z_p - z_b}{z_f - z_b}$$

(12.40)

with z_b and z_f being the backward and forward wire-end z-position, respectively.

For the vector $(\Delta x_b^w, \Delta y_b^w, \Delta x_f^w, \Delta y_f^w)^{\mathsf{T}}$ of the wire alignment parameters, we align directly the shifts at both wire-ends. Thus similarly to the layers, the derivatives of the local u-residual for the wire alignment are

$$\frac{\partial r_u}{\partial(\Delta x_b^w, \Delta y_b^w, \Delta x_f^w, \Delta y_f^w)^{\mathsf{T}}} = \begin{pmatrix} (1 - z^r)\frac{\partial r_u}{\partial\Delta x} & (1 - z^r)\frac{\partial r_u}{\partial\Delta y} & z^r\frac{\partial r_u}{\partial\Delta x} & z^r\frac{\partial r_u}{\partial\Delta y} \end{pmatrix}$$

(12.41)

where z^r is defined in Eq. 12.40.

When wires are aligned together with layers, the additional degrees of freedom introduced for layer alignment will result in a singular problem, as one can, for example, compensate layer rotation by an equivalent coherent rotation of the wires in the layer. Thus, similarly to the VXD alignment hierarchy, a set of basic wire constraints needs to be introduced, which will be discussed in Sec. 12.6.3

12.5   IP Position Alignment

The IP position measured previously from many tracks can (together with its covariance) provide an additional strong constraint to the alignment problem. But this position is itself correlated to the alignment, and thus it might be necessary to determine corrections also for the IP position in the alignment procedure.

An initial estimate of the IP position $\boldsymbol{b} = (b_x, b_y, b_z)^{\mathsf{T}}$ and the reconstructed prediction of a common vertex $\boldsymbol{v}^{\text{vtx}} = (v_x^{\text{vtx}}, v_y^{\text{vtx}}, v_z^{\text{vtx}})^{\mathsf{T}}$ of (usually two) tracks form a global 3D measurement residual defined as

$$\boldsymbol{r} = \boldsymbol{b} - \boldsymbol{v}^{\text{vtx}}.$$

(12.42)

We construct the local measurement system for each track with unit axes $\boldsymbol{u} = (u_x, u_y, u_z)^{\mathsf{T}}$, $\boldsymbol{v} = (v_x, v_y, v_z)^{\mathsf{T}}$ and a normal $\boldsymbol{n}$ at the point of the vertex $\boldsymbol{v}^{\text{vtx}}$ such that the normal coincides with the track direction. The residual in the global system $\boldsymbol{r} = (x, y, z)^{\mathsf{T}}$ can be projected to the local system $\boldsymbol{q} = (u = \boldsymbol{r} \cdot \boldsymbol{u}, v = \boldsymbol{r} \cdot \boldsymbol{v}, w = 0)^{\mathsf{T}}$ using the following matrix

$$\boldsymbol{P} = \frac{\partial(u, v)^{\mathsf{T}}}{\partial(x, y, z)^{\mathsf{T}}} = \begin{pmatrix} u_x & u_y & u_z \\ v_x & v_y & v_z \end{pmatrix}.$$

(12.43)

The local residual is determined as

$$(r_u, r_v)^{\mathsf{T}} = \boldsymbol{P}\,\boldsymbol{r}$$

(12.44)

and its global derivatives with respect to the beam spot (IP) alignment parameters $(\Delta b_x, \Delta b_y, \Delta b_z)^\intercal$ are

$$\frac{\partial(r_u, r_v)^\intercal}{\partial(\Delta b_x, \Delta b_y, \Delta b_z)^\intercal} = \boldsymbol{P}. \tag{12.45}$$

In addition, the local external measurement (2D) covariance is

$$\boldsymbol{V}_{(r_u, r_v)^\intercal} = \boldsymbol{P}\,\boldsymbol{V_b}\boldsymbol{P}^\intercal, \tag{12.46}$$

where $\boldsymbol{V_b}$ is the (3D) IP position covariance matrix in the global coordinates. This is a completely external input, which must be provided by an estimate[7] of the size of the luminous region ellipsoid and its exact orientation in space.

12.6 Alignment Constraints

In this section, the linear equality constraints on the alignment parameters, introduced in Sec. 11.1.4, are discussed. The constraints basically fall into three categories:

- Trivial (or basic) constraints remove additional degrees of freedom introduced for alignment of high-level structures. These constraints become necessary for VXD if ladders and/or half-shells are aligned together with the sensors and for the CDC if wires are aligned together with the layers.

- Reference constraints define the global reference frame and remove the basic six undefined degrees of freedom that describe coherent movements of the whole tracking system. Most of the constraints on the parameters of the CDC layers fall into this category, as we decided to use the CDC to define the Belle II alignment reference frame.

- Special constraints remove weakly defined degrees of freedom, which the alignment procedure cannot sufficiently constrain. In this category, we have a constraint on the twist of the CDC and various types of constraints on the parameters of CDC wires, mainly fixing their radial movements.

12.6.1 VXD Constraints

The constraints removing the additional degrees of freedom for alignment of ladders and half-shells can be retrieved from the hierarchy of geometry transformations. We have already discussed how the rigid-body alignment parameters of consecutive hierarchy levels are related. A change in a higher-level A will propagate to all daughter elements of a lower-level B as

$$\Delta\boldsymbol{a}_{\mathrm{B}} = \frac{\partial\boldsymbol{a}_{\mathrm{B}}}{\partial\boldsymbol{a}_{\mathrm{A}}}\Delta\boldsymbol{a}_{\mathrm{A}}. \tag{12.47}$$

---

[7]At least a preliminary estimate is needed to assign reasonable weights to the external IP position measurement (via its uncertainty).

One can invert this relation and extract the coherent part of translations and rotation of lower-level structures by simply[8] taking an average over all elements B which belong to A as:

$$\Delta \boldsymbol{a}_\mathrm{A} = \frac{1}{N} \sum_{B \in A}^{N} \frac{\partial \boldsymbol{a}_\mathrm{A}}{\partial \boldsymbol{a}_\mathrm{B}} \Delta \boldsymbol{a}_\mathrm{B}. \tag{12.48}$$

Such correlated movements must be removed from the solution as we already represent them by the parameters $\boldsymbol{a_\mathrm{A}}$ Thus the six constraints that remove the redundant degrees of freedom of the higher level A will read

$$\sum_{B \in A}^{N} \frac{\partial \boldsymbol{a}_\mathrm{A}}{\partial \boldsymbol{a}_\mathrm{B}} (\Delta u, \Delta v, \Delta w, \Delta \alpha, \Delta \beta, \Delta \gamma)_\mathrm{B}^\mathsf{T} = \boldsymbol{0}. \tag{12.49}$$

These constraints are generated automatically from the known structure of the hierarchy and stored transformation matrices $\frac{\partial \boldsymbol{a}_\mathrm{A}}{\partial \boldsymbol{a}_\mathrm{B}}, \frac{\partial \boldsymbol{a}_\mathrm{B}}{\partial \boldsymbol{a}_\mathrm{A}}$. Therefore if both ladders and half-shells are aligned with the sensors, there will be six constraints for each half-shell fixing the ladders inside the half-shells. There will be six constraints needed for each ladder, fixing the average movements of sensors in the ladder.

12.6.2 CDC Layer Constraints

In this section, we will mainly discuss the method of fixing CDC as a rigid body in space. For this, coherent movements have to be extracted from the combinations of alignment parameters and set to zero. To fix movements along x and y and rotation in the $x - y$ plane, one needs the following three constraints:

- Δx_b-constraint (shift along x):

$$\sum_{l=0}^{55} \Delta x_{b,l} = 0 \tag{12.50}$$

- Δy_b-constraint (shift along y):

$$\sum_{l=0}^{55} \Delta y_{b,l} = 0 \tag{12.51}$$

- $\Delta \phi_b$-constraint (rotation in $x - y$ plane, angle ϕ or γ):

$$\sum_{l=0}^{55} \Delta \phi_{b,l} = 0 \tag{12.52}$$

---

[8]This is simple only to first order as rotation and translation generally do not commute. One can, however, iterate this relation by back-propagating higher-level updates to lower-level structures using Eq. 12.47. One may even need to update the linearized transformations between the local systems for substantial shifts and rotations. In practice, we have needed only a couple of iterations with the former approach so far.
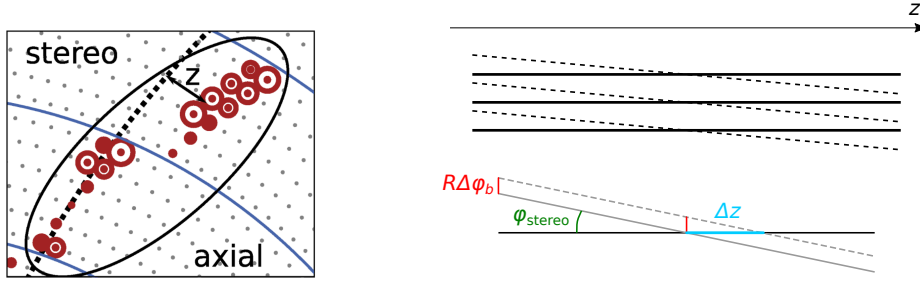
Figure 12.9: Left: reconstruction of the z-coordinate from stereo and axial hit pattern at $z = 0$ projection in the CDC [88]. Right: schematic projection of axial wires and wires with a stereo angle (top right) and how the apparent crossing point of stereo and axial wires (assumed in reconstruction) moves with changing alignment parameter $\Delta\phi_b$ (bottom right), controlling coherent layer rotation at both end-plates.

Similarly, the rotation around angles α and β[9] can be fixed by constraining the average difference of the relative shifts in y and x between the backward and forward end-plate. In terms of the already introduced alignment parameters, this evaluates to:

- Δy_{f-b}-constraint (rotation in $y - z$ plane, angle α):

$$\sum_{l=0}^{55} \Delta y_{f-b,l} = 0 \tag{12.53}$$

- Δx_{f-b}-constraint (rotation in $x - z$ plane, angle β):

$$\sum_{l=0}^{55} \Delta x_{f-b,l} = 0 \tag{12.54}$$

Note that all these constraints could be removed if, for example, a fixed VXD could be used as an alignment reference. Similarly, if, for example, a fixed IP position is introduced as an alignment constraint through IP-constrained decays, one could remove constraints for shifts in x, y, and z (discussed in the following subsection). These variations amount to different definitions of alignment reference frames.

Z-Offset Constraint

So far, we have only provided constraints that fix the CDC movements as a rigid body in five parameters. To understand the method to constrain effective movement in the z-coordinate, it is best to see how the local approximate z-position of a track can be inferred from the hit pattern in alternating axial and stereo layers, as sketched in Fig. 12.9 (left). The stereo hits appear as shifted

---

[9]These rotation angles are defined to be around the center of the coordinate system, while the layer alignment parameters Δx_{f-b}, Δy_{f-b} rather correspond to rotations around the backward side at negative z and thus amount to a combined rotation and shift (of the CDC center of mass), which are both unwanted.

because the track hits different stereo wires along its path. Thus the z-coordinate of a track can be locally estimated from the shift of axial wire hits in the projection at $z = 0$. If a continuous hit pattern is left, the track is close to vertical ($\tan \lambda \sim 0$) and with $z_0 \sim 0$. Only tracks that are passing near the *crossing points* of stereo and axial wires, as sketched in Fig. 12.9 (top right), will leave the hit pattern undisturbed. In the design geometry, this crossing point occurs for all wires at $z = 0$.

By inducing a specific rotation pattern to the axial layers, one can synchronously move the crossing points in z by Δz. The rotation must occur by the same amount at both end-plates and thus is related to the layer alignment parameter $\Delta \phi_b$. Consider such a small rotation by $\Delta \phi_b$ in a specific axial layer with stereo angle ϕ_{stereo} and radius R. As sketched in Fig. 12.9 (bottom right), the shift of the crossing point is related to the layer alignment as

$$\Delta z = R \Delta \phi_b \tan(\phi_{\text{stereo}}). \tag{12.55}$$

The same Δz must be produced in all stereo layers[10]. In that case, the total effect is an effective shift of the center of the CDC in the z-coordinate. If we wish to remove the combination of alignment parameters resulting in such CDC shifts in z, we have to introduce the following **z-offset constraint**:

$$\sum_{l=0}^{55} R_l \tan(\phi_{\text{stereo},l}) \Delta \phi_{b,l} = 0 \tag{12.56}$$

to which actually only the stereo layers contribute as $\phi_{\text{stereo}} = 0$ for axial layers.

Twist Constraint

This last constraint for the CDC layers belongs to the special category and may not be needed if the track sample contains the necessary information to prevent momentum biases for oppositely charged tracks. A simple set of tracks from the IP will not provide such a constraint, and the CDC could be easily twisted (backward and forward end-plate rotated relative to each other), changing the momenta of the reconstructed tracks as a function of the polar angle and charge.

Such a kind of twist deformation can be removed from the spectrum of possible alignment solutions by introducing the **twist constraint**:

$$\sum_{l=0}^{55} \Delta \phi_{f-b,l} = 0. \tag{12.57}$$

12.6.3 CDC Wire Constraints

The main set of constraints involving alignment parameters of the wires $(\Delta x_b^w, \Delta y_b^w, \Delta x_f^w, \Delta y_f^w)$ is needed when wires are aligned together with the layers. As we introduced the additional layer alignment parameters describing correlated wires shifts and rotations, these movements at the wire level have to be removed from the alignment solution. The correlated wire shifts are removed by two

---

[10]Otherwise, the tracks would not be smooth.

constraints at each end-plate per every layer L:

$$\texttt{backward:} \sum_{\text{wire}\in L} \Delta x_{b,\text{wire}}^{w} = 0 \ (\texttt{x-shift}), \qquad \sum_{\text{wire}\in L} \Delta y_{b,\text{wire}}^{w} = 0 \ (\texttt{y-shift}),$$
$$\texttt{forward:} \sum_{\text{wire}\in L} \Delta x_{f,\text{wire}}^{w} = 0 \ (\texttt{x-shift}), \qquad \sum_{\text{wire}\in L} \Delta y_{f,\text{wire}}^{w} = 0 \ (\texttt{y-shift}),$$

$$(12.58)$$

To remove average rotations of wires in the layers, the vector of wire alignment parameters $(\Delta x_{f/b}^{w}, \Delta y_{f/b}^{w})^{\intercal}$ has to be projected to the $R - \phi$ coordinate using wire angle $\phi_{b/f,\text{wire}} = \text{atan2}(y_{b/f,\text{wire}}^{0}, x_{b/f,\text{wire}}^{0})$ at the respective end-plates. This results in two additional constraints per layer:

$$\texttt{backward:} \sum_{\text{wire}\in L} -\sin(\phi_{\text{wire}}^{b})\Delta x_{b,\text{wire}}^{w} + \cos(\phi_{\text{wire}}^{b})\Delta y_{b,\text{wire}}^{w} = 0 \ (\phi\texttt{-rotation}),$$
$$\texttt{forward:} \sum_{\text{wire}\in L} -\sin(\phi_{\text{wire}}^{f})\Delta x_{f,\text{wire}}^{w} + \cos(\phi_{\text{wire}}^{f})\Delta y_{f,\text{wire}}^{w} = 0 \ (\phi\texttt{-rotation}),$$

$$(12.59)$$

Radial Constraints

The radial constraints belong to the category of special constraints, as it is generally very difficult to align the wires in the radial coordinate. A typical physics track moves in an approximately radial direction, and thus by construction, the residuals of the wire measurements do not depend on the radial shift of the wires. We will study the impact of the absence of any radial constraints in the full alignment with wires in Chapter 16. Let us here only summarize the different types of radial constraints one may introduce.

Similarly to removing rotations of wires in layers, one can project the wire alignment parameters to the radial direction. Using again the wire angle $\phi_{\text{wire}}^{b/f} = \text{atan2}(y_{b/f,\text{wire}}^{0}, x_{b/f,\text{wire}}^{0})$ at the respective end-plates, we define **radial constraint** for a set of wires X at backward (b) or forward (f) side as:

$$\text{C}^{\text{rad}}(X; \text{side}) := \sum_{\text{wire}\in X} +\cos(\phi_{\text{wire}}^{\text{side}})\Delta x_{\text{side,wire}}^{w} + \sin(\phi_{\text{wire}}^{\text{side}})\Delta y_{\text{side,wire}}^{w} \qquad (12.60)$$

Using this notation, we can define a set of two **layer radius constraints** for some layer L

$$\text{C}^{\text{rad}}(\text{layer } L; b) = 0$$
$$\text{C}^{\text{rad}}(\text{layer } L; f) = 0$$

$$(12.61)$$

or for more layers (two constraints per each layer). We also define **CDC radius constraints**, fixing the average change of the CDC radius at both sides. In such case, the sums run over all wires in the CDC:

$$\text{C}^{\text{rad}}(\text{CDC}; b) = 0$$
$$\text{C}^{\text{rad}}(\text{CDC}; f) = 0$$

$$(12.62)$$

One last version of radial constraints is discussed in the next subsection, where we split layers into hemispheres.

Hemisphere Constraints

The last, truly special constraints have been introduced to directly suppress an observed weak mode in the full alignment with wires (see Chapter 16). These are, in fact, variations on the already presented hierarchy constraints fixing average rotations of wires and average radial movement of wires in a layer. Instead of fixing the average over all wires in the layer, the wires are split into hemispheres, for the rotations to left and right and for radial shifts to the upper and lower hemispheres. Obviously, if such constraints are introduced for a layer L, the previous constraints for the averages over the whole layer are automatically fulfilled and thus must be removed if the hemisphere constraints are used.

Using the notation from previous sections, let us first define **circular constraint** for a set of wires X as:

$$\mathrm{C}^{\mathrm{cir}}(X; \mathrm{side}) := \sum_{\mathrm{wire} \in X} -\sin(\phi_{\mathrm{wire}}^{\mathrm{side}})\Delta x_{\mathrm{side,wire}}^{w} + \cos(\phi_{\mathrm{wire}}^{\mathrm{side}})\Delta y_{\mathrm{side,wire}}^{w} \qquad (12.63)$$

and using the definition of the radial constraint in Eq. 12.60, we define a set of eight hemisphere constraints for a layer L:

$$\mathrm{C}^{\mathrm{cir}}(\{\mathrm{wire} \in L, \ \cos(\phi_{\mathrm{wire}}^{b}) \leq 0\}; \ b) = 0$$
$$\mathrm{C}^{\mathrm{cir}}(\{\mathrm{wire} \in L, \ \cos(\phi_{\mathrm{wire}}^{f}) \leq 0\}; \ f) = 0$$
$$\mathrm{C}^{\mathrm{cir}}(\{\mathrm{wire} \in L, \ \cos(\phi_{\mathrm{wire}}^{b}) > 0\}; \ b) = 0$$
$$\mathrm{C}^{\mathrm{cir}}(\{\mathrm{wire} \in L, \ \cos(\phi_{\mathrm{wire}}^{f}) > 0\}; \ f) = 0$$

$$(12.64)$$

$$\mathrm{C}^{\mathrm{rad}}(\{\mathrm{wire} \in L, \ \sin(\phi_{\mathrm{wire}}^{b}) \leq 0\}; \ b) = 0$$
$$\mathrm{C}^{\mathrm{rad}}(\{\mathrm{wire} \in L, \ \sin(\phi_{\mathrm{wire}}^{b}) \leq 0\}; \ f) = 0$$
$$\mathrm{C}^{\mathrm{rad}}(\{\mathrm{wire} \in L, \ \sin(\phi_{\mathrm{wire}}^{f}) > 0\}; \ b) = 0$$
$$\mathrm{C}^{\mathrm{rad}}(\{\mathrm{wire} \in L, \ \sin(\phi_{\mathrm{wire}}^{f}) > 0\}; \ f) = 0.$$

CHAPTER 13

Alignment Software

In this chapter, the alignment method will be described in terms of the software implementation. This topic will be discussed only briefly, without aiming to provide any actual documentation. Just one code example will demonstrate the user interface to the alignment. However, the importance of this topic should not be underestimated. In fact, any serious experimental physics work nowadays requires a lot of time spent on software development. After implementing all the concepts presented in the previous sections, extensive testing and validation on simulations and later real data were required. For this work to be practically useful for physics, it cannot be done separately but needs to be fully integrated into the complete workflow of the experiment and automated. The author was the leading alignment developer since the first lines of code, with occasional contributions from other collaboration members.

The Belle II core software framework will be introduced in the following sections. Its most crucial part for the alignment, the track finding, and reconstruction, will be described briefly. As the alignment is only one of many detector calibrations, the general framework for automating all calibration tasks will be introduced before describing the actual alignment software.

In addition, alignment validation and misalignment will be included in this chapter. These are slightly less software-oriented topics but had to be implemented along with the alignment procedure itself to exercise it on simulations, validate its results with MC or real data or estimate its related systematic effects for physics analyses.

13.1 Belle II Software Framework

The Belle II Analysis Software Framework (basf2) [101] provides a foundation for practically all software-related tasks of the Belle II experiment, from GEANT4 [99] simulation and detector response modeling (digitization) to reconstruction and physics analysis. It has a modular design, and its core features are written in compiled C++ (C++17 standard). The user interfaces and many additional functionalities use Python code with bindings to the compiled layer. The basf2 is divided into external software, utility tools, and more than 30 packages with experiment-related code. Millepede II is part of the external software provided with the framework.

The main task of HEP frameworks is efficient event-level processing, where large amounts of experimental data are split into individual triggered events, which need to be simulated, reconstructed, calibrated, and analyzed. In the event loop, the building blocks providing well-defined tasks, called **modules**, are

executed in the order given by the user. This order does not need to be strictly linear. The user puts the modules in a **path**, and depending on an optional return value of any of the modules; this path can be split such that some modules are executed only conditionally. This is illustrated in Fig. 13.1. The modules exchange mutable data via a `DataStore` based on ROOT I/O functionalities providing object persistence, streaming, and storage.
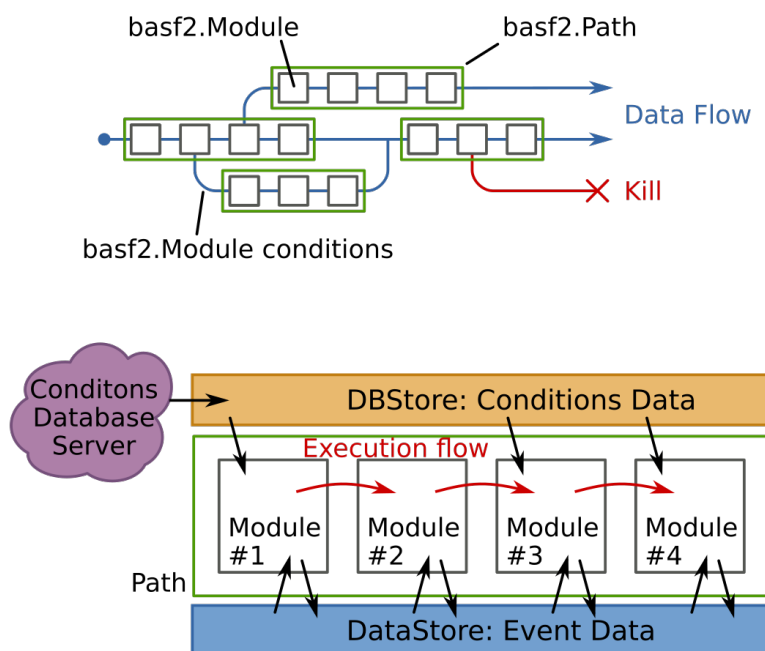
Figure 13.1: Schema of basf2 paths with modules and conditional branches (top) and interaction of modules with the common `DataStore` and conditions database (bottom).

In addition to event data, the processing also needs many other inputs, which may depend on time and conditions during the experiment operation. This condition data is stored in a database. The database directly stores the C++ objects with the conditions, like calibration or alignment constants, called **payloads**. Each such payload is assigned an Interval of Validity (IOV), which defines the range of experiment and run numbers where it is supposed to be valid. The set of IOVs and payloads are accumulated for all possible condition types in a **global tag**. Each global tag (GT) has a unique identifier. Global tags can be chained together such that if no valid payload is found in the GT with the highest priority, other GTs are searched for a valid set of calibration constants.

A standard reconstruction path in the current production has over 100 modules. In the next section, we will discuss only the most relevant part for the alignment: hit and track reconstruction in the CDC and VXD. Naturally, the reconstruction continues to the outer detectors and is also performed for other particles, which do not leave any signal in the tracking detectors, like photons in the ECL or K_L^0 mesons in the KLM.

13.2 Belle II Track-Finding and Tracking

To even start any alignment studies, one needs to find the particle trajectories in the reconstructed hit patterns. The process of track finding follows after initial unpacking (or simulation of particle propagation and detector response on MC) and preliminary hit level processing, like masking of hot channels and clustering, all provided by the respective modules from the sub-detector packages. The tracking chain [88] is highly modular. Its simplified schema is shown in Fig. 13.2.

13.2.1 Tracking Inputs

The PXD hits, which enter the full tracking chain, are determined by Regions of Interest (ROI), provided by the HLT with preliminary tracking without the PXD. The neighboring individual pixel hits with a charge above a given threshold are merged into clusters. The charge of the cluster and the best estimate of the particle intersection with the sensors are provided to the tracking.

In the SVD hit reconstruction, a similar clustering procedure is employed at the strip level, where the additional time and signal wave-form information can be utilized to improve signal and off-time beam background discrimination.

In the CDC hit reconstruction, the front-end electronics providing analog signal digitization with a time resolution of ~ 1 ns is used to measure the delay from the trigger signal to the time of arrival of the drift electrons to the sense wires and the read-out. Assuming the particle propagates at the speed of light, the drift time of the electrons to the wire can be estimated and converted to drift length using a calibrated $x-t$ relation. Additional amplitude information related to the unit energy loss of the particle is used for dE/dx particle identification.
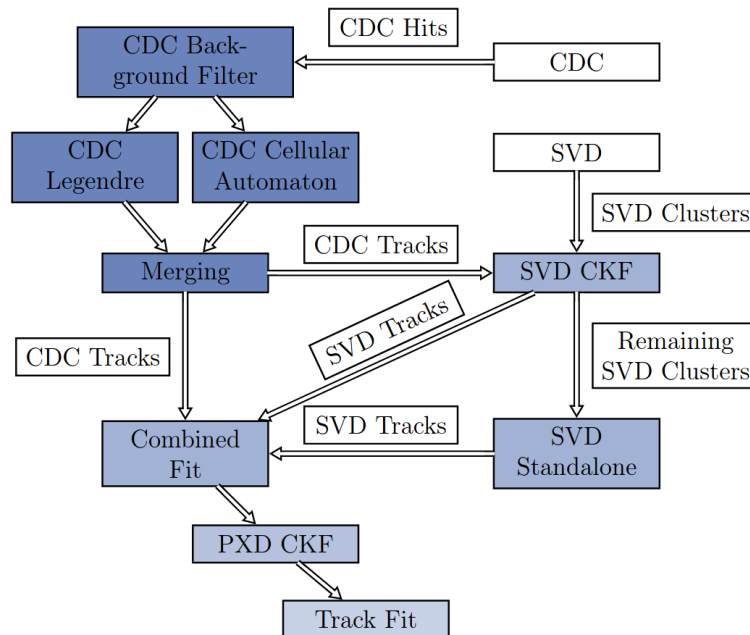
Figure 13.2: Simplified schema of the Belle II tracking chain with modules in blue and intermediate `DataStore` objects in white boxes. [88]

13.2.2 CDC Track Finding

After initial CDC hit filtering to suppress background hits, CDC track finding follows. This step is performed in two different approaches whose results are later combined (merged). The global CDC tracking-finding is based on the Legendre transformation and is mainly aimed at the reconstruction of tracks from nearby the interaction point. The algorithm starts in the $R - \phi$ plane with axial wires, where every CDC hit is represented by a drift circle. A transformation into the conformal plane of plausible circles crossing the origin and being a tangent to each of these drift circles is performed. The result is a plot with many sine-like functions. In this conformal mapping, tracks are found as common intersections of these functions (whose coordinates correspond to the track parameters). This is iterated with the previously found and assigned hits being removed. A fast circle fit is performed, and additional merging of compatible tracks is done, with possible hit reassignment or attachment of unassigned hits. Finally, the stereo hits are added to the tracks by similar procedures, only now the tracks are straight lines in the appropriate conformal space, where the intersections correspond to the helix parameters z_0 and $\tan \lambda$ [88].

The local CDC track finding approach complements the global approach by searching for tracks without assumptions about their origin. It is based on a weighted cellular automaton that searches for connected hits in CDC super-layers, called segments. Compatible segments are combined into tracks, with weighting based on χ^2 of a circle fit or a linear fit when stereo hits are combined.

At the merging stage, tracks found in the global approach are used as a baseline, and segments from the local approach can be added using a multivariate method trained on simulations to discriminate between wrong and good matches.

13.2.3 Tracking for SVD and PXD

The resulting tracks from the CDC are extrapolated to the SVD by Combinatorial Kalman Filter (CKF) [88], which starts from an initial seed and iteratively extrapolates inwards, looking for compatible SVD hits. For tracks that do not reach the CDC or leave only a few CDC hits, a standalone SVD track finding algorithm is used. The inputs are 3D reconstructed global positions of the SVD hits. Connected graphs are created with compatible hits, determined using a map of compatible sectors, and trained on simulations. This significantly reduces the initial combinatorics. A cellular automaton is applied to the graphs, which yields plausible paths, forming track candidates. As a final step, the final collection of tracks is retrieved by selecting the best non-overlapping candidates.

Once the SVD hits are attached to the CDC tracks, these tracks are merged with CDC-only and SVD-only tracks and a combined fit by a Deterministic Annealing Filter (DAF) is performed [88]. DAF is used instead of a standard Kalman Filter (on which it is based), as it allows additional down-weighting of bad measurements. This is very important for left/right passage disambiguation for the CDC hits as the wrong side hits are down-weighted. In fact, DAF is used at multiple stages in the tracking chain to yield the best estimate of the track parameters. These estimates are provided by GENFIT2, already discussed in Sec. 11.2, which is interfaced with the tracking. All track finding and fitting algorithms operate on a common `DataStore` object representing a track candidate and its fitted version,

called `RecoTrack`.

In the final stage of the track finding, the PXD hits are attached to the `RecoTracks` using CKF, and the final track fit is performed.

13.2.4   Tracking Output

Each `RecoTrack` can be an input to the alignment framework. The resulting weights from DAF are used to choose the correct left/right side for each CDC hit, such that only a single measurement is added at any arc length of the GBL trajectory. This is a general design requirement of GBL. For SVD hits, which usually come as two individual 1D measurements at the same point of the `RecoTrack`, these must be first combined into a 2D measurement. For slanted SVD sensors, this also requires a non-diagonal covariance matrix, which encodes the correlation of the two non-orthogonal strip hits.

The chain of the modules can also be reorganized for the reconstruction of cosmic ray tracks. Each `RecoTrack` is extrapolated to the origin in the standard reconstruction chain to determine its helix parameters and construct a compact high-level `Track` object, which is the input of further reconstruction and physics analyses. For most use-cases, the hit-level information is pruned to save disk space. Alignment, however, needs complete hit-level information.

It should also be noted that a particle hypothesis is required for track extrapolation and fitting. The default particle hypothesis is a pion, the most commonly produced particle. In alignment, if the inputs are directly the `RecoTracks`, we instead assume the muon hypothesis. This is correct for cosmic ray tracks and di-muons. For hadron and off-IP samples, a mix of particle species is present, but the effect of using incorrect mass can be mostly neglected on average.

13.2.5   Vertexing

Finally, a very common procedure after the tracks are found is the decay reconstruction and vertex fitting. Several options are available in basf2. The default is `TreeFitter` [102], providing a global decay chain fit that allows to include some additional constraints, not only on particle masses but also on lifetimes, allowing to, e.g., include K_S^0 decay into the vertex fit. Naturally, the vertex fit, which requires two or more tracks to meet at a single point (the vertex), is the most common. In addition, one may introduce more constraints stemming from the underlying physics process knowledge, e.g., an invariant-mass constraint or an IP profile constraint, which includes information about the probability distribution of the vertices coming from the primary interaction region. For alignment, the RAVE [103] vertex fitter is used for di-muon decays. After the vertex fit, the helix parameters of the daughter particles can be updated. This updated set of parameters is used as a seed for constructing the GBL reference trajectory instead of the original track seeds provided by the track finding. These are used only in case plain `RecoTracks` are the input of the alignment.

13.3 Belle II Calibration Framework and Automation

A dedicated basf2 calibration package with a subset of tools known as Calibration and Alignment Framework (CAF) was introduced to automate common calibration-related tasks and allow developers (usually physicists) to concentrate on algorithm development. The building blocks for each particular calibration included in the CAF are

- **Collector module**, aggregating calibration-specific data and executed in the event loop, and

- **Algorithm class**, which contains the logic of the calibration and is executed over the collected and merged data.

Both are implemented in C++ with Python bindings exposed, such that most of the CAF code devoted to workflow management can be written in Python.

The CAF takes care of the parallelization of the data collection step. Different back-ends for batch submission provided by computing clusters are supported, as well as a local back-end for testing and development. The calibration developer should provide the collector and algorithm code and a basf2 path with necessary pre-processing before data can be retrieved by the collector, typically with standard reconstruction modules.

The individual calibrations are typically dependent on each other, and the CAF takes care of the correct execution order. In addition, the databases and payloads are managed by the CAF, as well as possible iterations, which can be requested by the algorithm.
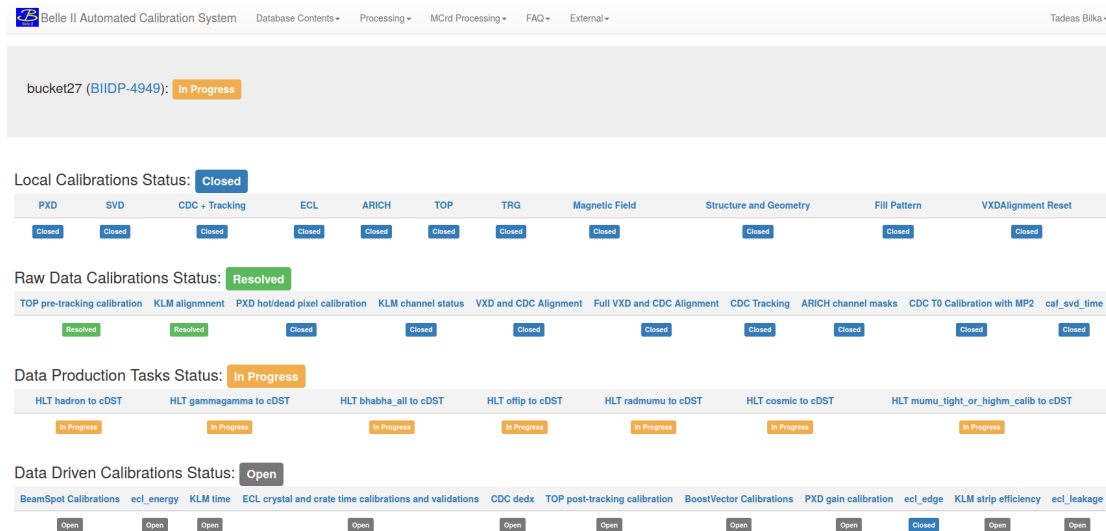
Figure 13.3: A snapshot of the Belle II Airflow-based calibration website shows the current automated calibration status in the prompt processing of one calibration block (`bucket 27`) after the alignment task has finished.

The high-level automation of all calibration tasks is managed by the Python-based Apache Airflow [104] platform for workflow scheduling and monitoring.

Calibration is an Airflow task that creates new conditions by producing payloads and their corresponding IOVs. The calibration workflow is separated into four stages:

1. Local calibrations, which are independently provided by the sub-detector groups. These contain mainly operating conditions but also initial calibrations, like masking of hot or dead channels.

2. Raw data calibrations, which require full access to the data at the hit level. Before alignment, the CDC tracking calibration, described below, is executed. Only then the alignment of the CDC and VXD can run, followed by the alignment of the KLM, which is also done using the alignment framework and Millepede II. This, like the following calibrations, rely on properly calibrated tracking detectors.

3. Data production for high-level calibrations, which run over all data and reconstruct it again for storage in a final format, known as calibration DST (cDST). In cDST, the usually available information for high-level physics analysis is accompanied by additional data objects useful for further calibrations and validations.

4. Data-driven high-level calibrations use the cDST information without the need for any additional data reconstruction and thus are usually very fast. Examples are calibration of the IP size and position, the boost vector, or the CDC dE/dx calibration.

The calibrations are submitted by experts from a convenient web interface, which allows selecting input data and configuring the prepared calibration tasks. Each task executes a CAF instance, which can be composed of multiple dependent calibrations. The list of all the calibration tasks currently present is illustrated in Fig. 13.3.

13.3.1   CDC Calibration

Alignment relies upon a reasonable calibration of the CDC unless only a stand-alone VXD tracking and alignment is performed. In the CDC hit reconstruction, initially, the drift time T_{drift} of the electrons needs to be derived, which is related to the TDC count when the signal arrives in the readout electronics as

$$T_{\mathrm{drift}} = T_0 - T_{\mathrm{event}} - T_{\mathrm{TOF}} - T_{\mathrm{prop}} - T_{\mathrm{tw}} - a \cdot \mathrm{TDC}, \qquad (13.1)$$

where T_0 is a constant offset per channel (wire), which needs to be calibrated first. T_{event} is the event time of the trigger, synchronized with the TDC clock, T_{TOF} is the estimated track flight time from the IP (or a reference plane for cosmic rays) to the wire, T_{prop} is the propagation time of the induced signal along the sense wire to the backward end-plate [21], T_{tw} is correction for the time-walk effect, which is a function of the hit charge in values of ADC[1], and $a = 0.98$ ns/count is TDC clock resolution. CDC calibration proceeds in iterations with multiple reconstructions of the tracks. After initial T_0 corrections are converged, the $x - t$

---

[1]Usual parametrization is $T_{\mathrm{tw}} \propto 1/\sqrt{\mathrm{ADC}}$.

relation, necessary to translate T_{drift} into drift length and position residual with respect to a fitted track, is calibrated. This is followed by calibration of position resolution and the time-walk corrections. Multiple passes may be needed for some steps or possibly for the full chain.

CDC tracking calibration is very time-consuming—it usually takes about two days on the computing cluster. This is naturally caused by the large number of constants to be derived. There are about 14×10^3 T_0 constants, one per wire. The time-walk corrections are parametrized per each of 299 front-end readout boards. The $x - t$ relation is parametrized by a 5th order Chebyshev polynomial. Near the cell boundary, the parametrization is replaced by a linear function. The $x - t$ relations are determined separately for each layer, side (left or right passage of track w.r.t. wire), incident angle α, and polar angle θ, see also Sec. 18.2.4. Also, the position resolution is parameterized separately for each layer, side, and different incident angles [21].

13.3.2 Buckets and Prompt Calibration Automation

The initial purpose of the fully automated calibration in Airflow was to provide preliminary constants for the experiment data in a timely manner. This so-called prompt calibration loop is executed over data blocks, known as **buckets**. A bucket should ideally span several consecutive days with stable operating conditions. In practice, the buckets are currently longer, usually corresponding to two weeks of data-taking between each two maintenance days. The integrated luminosity target is 9 fb^{-1} per bucket, but occasionally smaller buckets are processed. Much more often, more data is collected, and the bucket is pre-scaled to 9 fb^{-1} equivalent of data.

The automated prompt calibration has been used since the end of 2019, and over 30 buckets have been processed so far. In addition, the Airflow infrastructure is also partially used for large data **reprocessings**, which usually happen once per year.

13.4 The Belle II Alignment Package

The alignment package, developed and maintained by the author, provides the complete implementation of the presented method, additional tools, and the infrastructure to efficiently and reliably apply the method to simulated or real data. One of the most considerable benefits is the very tight integration. For example, alignment uses exactly the same inhomogeneous magnetic field, energy loss estimation, or extrapolation methods as the full official reconstruction. This is possible because we previously integrated GBL into GENFIT2, which in turn is tightly integrated into basf2 tracking. This integration is partially experiment-independent. One can replace the geometry and interfacing detectors and run the alignment (properly configured). This can be used, for example, for beam tests or studies of future vertex detector upgrades.

13.4.1 Core Components

The core of the alignment package is the interface, which calculates the global derivatives with respect to the alignment parameters for every measurement used to construct the GBL trajectory. The classes that translate the local hit coordinates into measurements on virtual GENFIT2 planes are called `RecoHits`—for example, `PXDRecoHits` or `CDCRecoHits`. In the alignment package, special classes derive from those, as well as from abstract base class, which allows passing global and local derivatives to the GBL trajectory construction.

Another key component hidden from the user is a number of generic C++ classes to handle different types of payloads and standardize the access to the various alignment and calibration constants, calculation of global derivatives, and management of possible time-dependence. In fact, in developing this infrastructure, the major challenge was to coordinate work across several mostly independent groups and the ways their software packages and interfaces were organized. Beyond the alignment constants discussed in this thesis, we have implemented a number of other calibrations and alignments. In particular, alignment for EKLM and BKLM, as well as calibration of a number of CDC parameters, are supported. While KLM alignment is used in production, it is performed after the VXD and CDC alignment, not simultaneously. All CDC calibrations are currently experimental in different stages of testing (see also Chapter 18), and currently, Millepede II is not used for any of the CDC calibrations introduced above.

From the CAF point of view, the method is integrated into the `MillepedeCollector` module and the `MillepedeAlgorithm` class. Prior to the collector, a standard (or cosmic) reconstruction must be executed without hit pruning. The collector loops over `RecoTracks` in arrays specified by the user or `RecoTracks` related to particles from particle lists or decays, which must be provided by the user by putting appropriate modules to the pre-collector path. For each such track or decay, a simple or combined GBL trajectory is constructed and fitted to obtain the $p-$value, on which the user can impose a cut. Accepted trajectories are written to binary files to be used by the algorithm. The `MillepedeAlgorithm` prepares the configuration files for Milepede II, executes the external tool over the collected binary files, and processes the output into payloads by updating alignment parameters determined previously.

13.4.2 Code example

From the CAF user perspective, much of the details are hidden, and a relatively simple but highly configurable Python interface is provided. For easier configurations, fixing parameters, or adding constraints, several utility functions and classes are provided. A much larger configuration space opens with unlimited categories of input data samples. Every data sample is usually provided by a separate collection. A collection is a set of jobs with the same pre-collector path and collector configuration. The collector supports single trajectories (also from cosmic rays, including those without magnetic field) as well as vertex-constrained decays. Further supported is the IP constraint, which adds the knowledge of the primary IP to the alignment. In addition, experimental support was added for two-body decays to include an invariant mass constraint. This can be combined with the

IP constraint. A full four-momentum constraint, which utilizes the knowledge of machine boost and invariant mass and could be used for $e^+e^- \to \mu^+\mu^-$ events, was also developed.

Let us demonstrate the features of the alignment package by a realistic example of a user configuration script (adjusted from an actual production script), which could be executed by the CAF and performs full-scale alignment with over 60×10^3 free global parameters. The configured calibration to be executed (also demonstrated) is created by a single function call with a number of parameters. The example code follows in Listing 13.1:

```
1  import millepede_calibration as mpc
2
3  di-muon_path = create_std_path()
4  fillParticleList("mu+:good", "p > 1 and muonID > 0.9", path = di-muon_path)
5  reconstructDecay("Upsilon(4S):mumu -> mu+:good mu-:good", "9 < M and M < 11", path = di-
       muon_path)
6  raveFit("Upsilon(4S):mumu", 0.0, daughtersUpdate=True, path=di-muon_path)
7
8
9
10 cal = mpc.create(
11   name='VXDCDCalignment',
12
13   dbobjects=['VXDAlignment', 'CDCAlignment','BeamSpot'],
14
15   collections=[
16     mpc.make_collection("cosmic", path=create_cosmics_path(), tracks=["RecoTracks"]),
17     mpc.make_collection("hadron", path=create_std_path(), tracks=["RecoTracks"]),
18     mpc.make_collection("mumu", path=di-muon_path, primaryVetices=["Upsilon(4S):mumu"])
19   ],
20
21   tags=["data_reprocessing_prompt"],
22
23   files=dict(hadron=[], cosmic=[], mumu=[]),
24
25   timedep=[
26     ((alignment.parameters.beamspot(),
27       [(0, run, 0) for run in range(0, 100)]),
28     ((alignment.parameters.vxd_halfshells(),
29       [(0, run, 0) for run in range(0, 100, 10)]))
30     ],
31
32   constraints=[
33           alignment.constraints.VXDHierarchyConstraints(type=2, pxd=True, svd=True),
34           alignment.constraints.CDCLayerConstraints(z_offset=False, twist=False),
35           alignment.constraints.CDCWireConstraints(layer_rigid=True, layer_radius=[53],
       cdc_radius=True, hemisphere=[55])
36   ],
37
38   fixed=
39     alignment.parameters.vxd_sensors(rigid=False, surface2=False,
40     surface3=False, surface4=True)
41
42   commands=[
43     "method decomposition 6 0.001",
44     "scaleerrors 1."],
45
46   params=dict(
47     minPValue=0.00001),
48   )
49
50
51 cal_fw = CAF()
52 cal_fw.add_calibration(cal)
53 cal_fw.backend = backends.HTCondor()
54 cal_fw.run()
```

Listing 13.1: Alignment script example.

Let us explain it by individual parts, identified by the line numbers:

- Line **(1)** only includes the basic import. Other needed imports are not shown to keep the example shorter.

- Lines **(3–6)** demonstrate how a special track sample, in this case, IP-constrained di-muon events, can be created. The definition of `create_std_path` is not shown, but only the basic data unpacking modules and standard (or cosmic for `create_cosmic_path` below) reconstruction modules are added, followed by one final DAF fit. The user then can fill a particle list and perform decay reconstruction and vertex fit using standard framework utility

214

functions from the analysis package. The final particle list of the mother particle can be passed as alignment input, as well as the individual particle lists, identified by unique names. Charge conjugation is always imposed in the analysis package, so the reconstruction of μ^+ means that also corresponding μ^- particle list is created.

- Lines **(10–11)** create a fully configured Python `Calibration` class with a unique name using a single function call configured by several arguments, discussed below.

- Line **(13)** defines the payloads whose constants should be calibrated. All constants in the given payload are subject to alignment by default. The user can limit the number of parameters via the `fixed` argument, described below. The possible options which correspond to payload names are:

    – `VXDAlignment`, which enables alignment of all sensor, ladder, and half-shell parameters for VXD,

    – `CDCAlignment`, which enables alignment of the CDC layers and CDC wires

    – `BeamSpot` for alignment of the IP position in case some decays with primary beam constraint are used as input (see below),

    – `CDCTimeZeros` for calibration of CDC T_0 per each wire,

    – `CDCTimeWalks` for calibration of the time-walk corrections per each front-end board,

    – `CDCXtRelations`, which allows still fully experimental calibration of the $x - t$ relations,

    – `BKLMAlignment` and/or `EKLMAlignment`, which enables rigid body alignment of the BKLM and EKLM modules.

- Lines **(15–19)** define the input data for alignment, in this case in the form of three collections with standard data samples introduced in Sec. 11.7. An alignment collection is defined by a unique name, the pre-collector path, and the name of input arrays for alignment[2]. These inputs can be

    – `tracks` for single GBL trajectories created from the `RecoTracks`,

    – `particles` for single GBL trajectories created from the `RecoTracks` assigned to user-defined reconstructed particles,

    – `vertices` for composed GBL trajectories with a vertex constraint created from the `RecoTracks` of all daughters for each mother particle in the list,

    – `primaryVertices`, which in addition adds the primary IP constraint and allows for alignment of the IP position,

---

[2]And optionally specific configuration options for the given collector

– `twoBodyDecays` and their several variations[3], which work only for decays into two same-mass particles of type $V^0 \rightarrow f^+ f^-$ and automatically add an invariant mass constraint based on the mass of the mother particle from the `EvtGen` table.

- Line **(21)** defines the input global tags, which contain the set of initial calibration and alignment constants.

- Line **(23)** is to provide the list of input file names per each collection defined above.

- Lines **(25–30)** demonstrate an advanced feature to define time-dependent parameters. This is achieved by increasing the number of global parameters by the corresponding number of blocks. The boundaries of these blocks can be defined by 3-tuples with experiment, run, and the event number. In this particular example, the three parameters for `BeamSpot` (IP position) alignment (x, y, z) will be determined for every run from 0 to 100, while the alignment of the 24 parameters for rigid body alignment of the four VXD half-shells is done every ten runs. For each half-shell alignment block, a new set of 24 constraints is needed to fix the average movement of sensors with respect to half-shells in each given interval. The time-dependent constraints are generated automatically. All parameters, including the time-dependent ones, are determined in the simultaneous global fit. This feature is very powerful and potentially useful but requires prior explicit configuration with a good knowledge of the input data. Automating such a time-dependent configuration will require additional development in the automated Airflow workflow if used in the future.

- Lines **(32–36)** define the sets of constraints to be included in the solution. Several Python classes implement the linear equality constraints defined in Sec. 12.6. The `type=2` for VXD constraints means sensors are placed into half-shells (ladders are not used). Special constraints are used for layer 53 and CDC radius, and for layer 55, the hemisphere constraints are active.

- Lines **(38–40)** define fixed parameters. In this particular example, only the highest-level surface deformations for all VXD sensors are fixed. This is mostly used to exclude entire sets of parameters but can also be utilized to fix individual parameters, e.g., to define alignment reference (instead of constraints).

- Lines **(42–44)** allow to pass (or override default) configuration commands to Millepede II. In this case, the decomposition method with a minimum of six iterations for outlier rejection and down-weighting and convergence limit of 0.001 is used for the solution. The parameter `scaleerrors` has no effect when set to 1 but can be utilized for preliminary alignment in case of very large misalignments to scale the measurement errors.

---

[3]For example `primaryTwoBodyDecays`, which implements the full four-momentum kinematic constraint together with the IP-constraint

- Lines **(46–48)** allow to set or override default parameters common to all collector modules. Here the p-value of the GBL fit performed on each trajectory (single or combined) is required to be larger than `minPValue = 0.00001`. Collection-specific configurations can be set during collection creation.

- Lines **(51–54)** add the created Calibration object to the CAF and execute it, using the `HTCondor` back-end.

With proper data as input, this script will span many collector jobs, typically one per input file in each collection. After all collector jobs finish, their output is passed to the algorithm, which runs Millepede II and produces the updated payloads from the line (13). In case the solution method also provides parameter error estimates $\boldsymbol{\sigma_a}$, we also compute the normalized sum of parameter pulls

$$S = \frac{1}{N} \sum_i^N \Delta a_i^2 / \sigma_{a_i}^2. \tag{13.2}$$

For $S \geq 1$ or $\max_i |a_i/\sigma_{a_i}| > 10$, iterations with full re-running of the collectors using the just determined constants are requested. The CAF will iterate until the convergence criterion is fulfilled or the maximum number of iterations (by default 5) is reached.

13.5 Alignment Validation

In this section, we shall briefly discuss various alignment validation methods. For MC tests, we can directly validate the alignment quality by looking at the final payloads. For data, the only way to estimate the goodness of the alignment is to utilize multiple methods that cross-check the real performance of the tracking. A variety of methods exist, but we will only mention those really developed and applied to the Belle II data. More details can be found in Chapter 14 and 17, or in the corresponding references.

Validation with Residuals

Validation methods based directly on the track-to-hit residuals can be divided into two categories: those that do not need track information (like its incidence angle) and those that utilize the track parameters.

The former methods only use the average residuals collected in each sensor (or integrated over groups of sensors). Depending on how long the range of the data is accumulated in each such distribution, the method becomes sensitive to effects happening at different time scales. From a physics point of view, the total accumulated dataset is usually used, and thus most stochastic effects will be averaged, only contributing to a worse resolution and with some probability to some small systematic deformation[4]. By evaluating average residuals from the sensors over such a large dataset, one can estimate the average statistical precision of the alignment. An example for data of one bucket is shown in Fig. 13.4,

---

[4]With N sensors randomly displaced with a fixed misalignment of statistical precision σ, the average size of the expected coherent movement should be roughly proportional to $\sigma/\sqrt{N}$
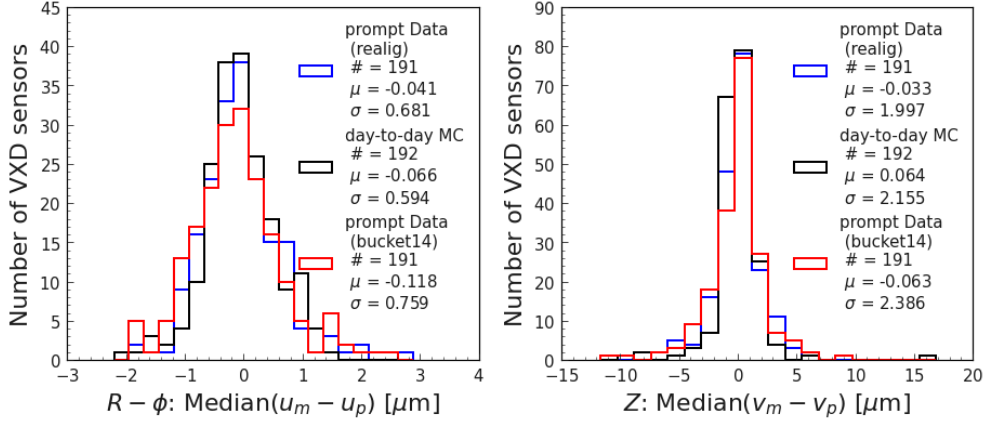
Figure 13.4: Histograms of the medians of distributions of track-to-hit residuals in u (left) and v (right) from all VXD sensors for real data of `bucket 14` with default prompt alignment (red), the same data after run-dependent alignment of the VXD half-shells (blue, denoted as 'realig') and MC simulation (with one more PXD sensor, no background) misaligned using day-to-day alignment difference from data (black). The number of entries (sensors), the mean μ, and the RMS denoted as σ of the medians are shown (in μm).

compared to a case with an improved alignment and to an MC simulation with misalignment, derived from a day-to-day alignment difference. The statistical alignment precision estimated with this method is around 1 μm and 2 μm for the $R - \phi$ and z directions, respectively. However, such a method does not allow full decoupling of the alignment degrees of freedom contributing to the sub-optimal residuals.

Increasing the time resolution of the method by using smaller blocks of data allows for observing the evolution of the residuals. We devote Sec. 17.3 to these residual-based methods and their results.

The second class of methods utilizes the information about the coordinates of the track intersection with the sensor and its incidence angles. This allows us to extract additional information and perform validation of almost all VXD sensor alignment parameters [105] (also as a function of run), including sensor deformations, which can be directly monitored. This method and some of its results are briefly discussed in Chapter 14.

One additional method explored at Belle II is based on overlap residuals [106], created by tracks crossing a VXD layer twice, where the sensors overlap. This method can detect and distinguish some weak mode misalignments, and its results for cosmic data are briefly demonstrated in Chapter 14.

Validation with Cosmic Rays

Muons from cosmic rays constitute a unique validation data sample. In the standard method employed for the vertex performance evaluation and alignment checks, used at both BaBar [70] and Belle [16], each cosmic track is split at the POCA to the origin. Each of the two arms is then reconstructed separately. The differences in the two arms' helix parameters can be used to measure detector resolutions. Their correlation to the helix parameters can be used to study
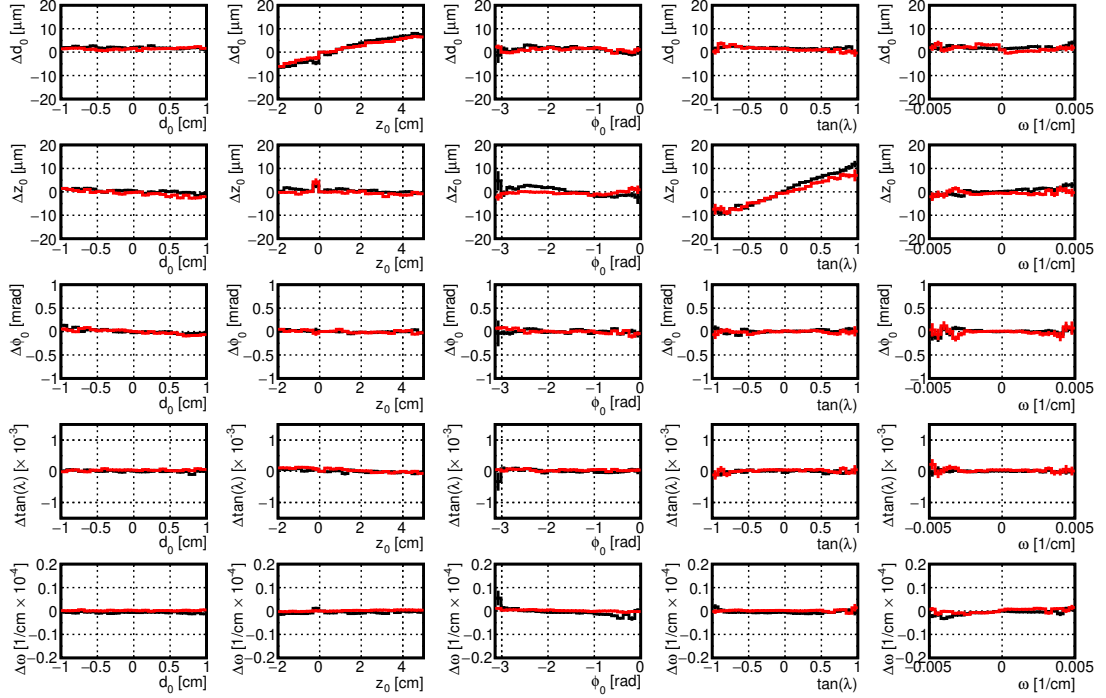
Figure 13.5: Difference of the reconstructed helix parameters of the upper and lower cosmic arm in the prompt alignment data of exp 12 (black) and MC with a model of prompt misalignment (red) as a function of helix parameters.

systematic misalignment.

In Fig. 13.5, we show the correlations of cosmic helix parameter differences as a function of the helix parameters themselves. In this particular example, one set of points comes from real data (prompt alignment) and the second from MC simulation with misalignment, extracted from the difference between prompt and reprocessing alignment. We discuss this method and the improvement due to the reprocessing in Chapter 17. Notable are the correlations in the upper two rows with vertex parameter differences, which are nicely reproduced by the misaligned MC.

Validation with Di-muon Events

Muons from $e^+e^- \to \mu^+\mu^-$ events, as well as, e.g., electrons from Bhabha scattering events $e^+e^- \to e^+e^-$ constitute relatively high momentum, experimentally clean, low-background samples. The tracks from these events are back-to-back in the CM frame and originate from a common interaction vertex. This knowledge can be exploited to compare the expected helix parameters of the two tracks, similarly to the cosmic validation. We devote Sec. 17.2 to this method.

Validation with Physics Analysis

The alignment shall also be validated at a much higher level by checking physics performance. These checks currently have the disadvantage that a long time range must be accumulated for sufficient statistics and that they usually depend also on the quality of other calibrations or on the level of data/MC correspondence.

Various analysis groups have performed multiple such cross-checks. Namely, studies of the estimated particle lifetime dependence (currently at least for D^0, D^+, Λ_c^+) on the data-taking period, azimuthal or polar angle, etc., constitute an important check of possible alignment systematics. None of these checks revealed any statistically significant deviation from expectations with the available dataset, except for problems with early data, which have been fixed, see Chapter 14. With much more data in the future, more subtle details could be resolved, and such lifetime analyses could serve as a standard validation tool for alignment performance evaluation.

13.6 Misalignment

In this section, we first discuss the implementation of the detector misalignment at the simulation and reconstruction level and its evaluation and visualization using a couple of typical examples. Afterward, the actual misalignment configurations are discussed, with emphasis on residual misalignment scenarios used for the estimation of the systematic uncertainty due to alignment in physics analyses.

13.6.1 Simulation of Misalignment

The geometrical configuration of sensitive elements of a real detector can never be known exactly. In other words, the detector is misaligned. If the misalignment is sufficiently large, physics performance can be degraded. In real data, the misalignment is *for free*, and we aim to reduce it. However, in the basf2 GEANT4 [99] simulations, the detector is assumed ideal for technical reasons. If misalignment is applied to the simulation, the sensitive volumes could start to overlap with the support structures, which is not allowed for a reliable simulation.

Despite that, the concept of **displacement** was introduced. Displacement means a difference from the design geometry directly at the simulation level. If such displacement is implemented, the geometry needs to be checked for overlaps, and these need to be (somehow) resolved. Currently, displacements are only used for CDC wires, using information from a mechanical survey [21]. The VXD is always simulated in the design position.

A correct simulation of misalignment on top of a design geometry should take into account, for example, edge effects, where some hits get lost (or should be added) to a misaligned sensor. This would require modification of the simulation result at the level of individual digits, even before track-finding. We have not implemented this approach. Instead, the misalignment is simulated by using *incorrect* non-zero alignment payloads in the reconstruction (both track finding and track fitting are using alignment-corrected sensor positions). Thus the correct final alignment result in MC tests should be exactly the **nominal geometry** (design + displacements), i.e., zeros for all alignment parameters. The difference to the ideal result is called **residual misalignment**. Such an approach makes the check of the alignment result rather trivial.

The task of the alignment method on MC is then to bring the alignment constants as close to zeros as possible[5]. Exactly the same approach is used to

---

[5]Ideally, the level of residual misalignment on simulation should also correspond to that in

correct for misalignment with real data. But with real data, the expected result is, of course, non-zero, as the source geometry is not nominal.

13.6.2 Evaluation and Visualization of Misalignment
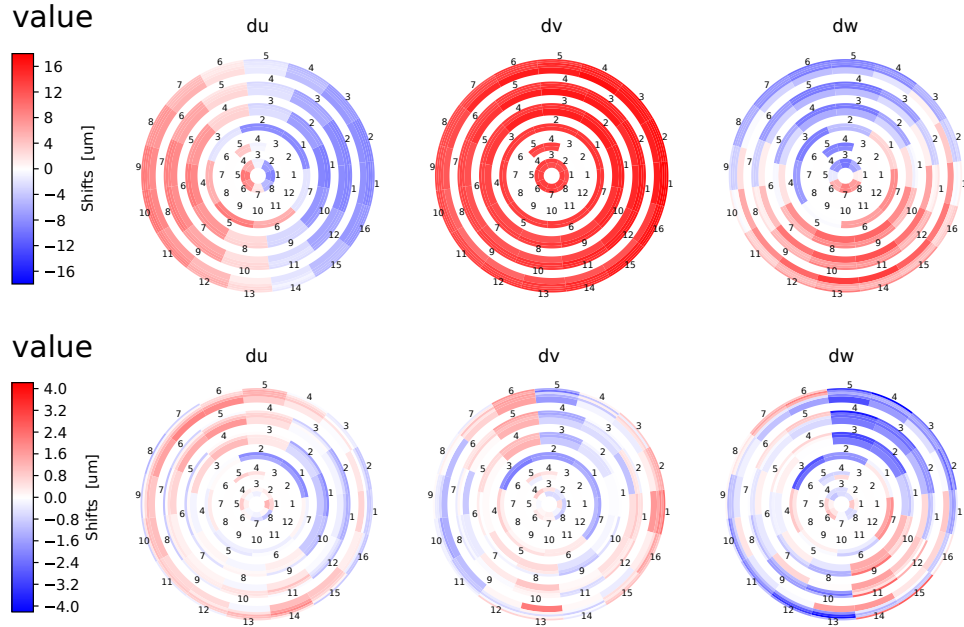
Figure 13.6: Ring-plot visualization of the residual misalignment for VXD sensor shifts from an MC test of the baseline alignment without constraints (top) and after transformation into constraints' reference frame (bottom), where sensors and half-shells are aligned together and the coherent movements are absorbed into half-shell parameters, shown as histograms in Fig. 13.7.

One drawback for the evaluation of the resulting residual misalignment occurs in the (typical) case when linear equality constraints are used. With constraints, the alignment parameters are expressed in a different reference frame. As an example, consider a situation where the misalignment resulted in an average total shift of the sensors in the local v coordinate (which points along the global z direction). For a hierarchy where sensors are placed in half-shells, the alignment is restricted, such that this average over sensors cannot change. Thus it has to move the half-shell along z in the opposite direction. After (successful) alignment, the absolute position of the sensors will be close to the nominal, but the actual parameters in the payloads will be non-zero. We can we discuss the residual misalignment only after we correctly interpret them and combine the corrections for sensors and half-shells.

Therefore, one needs to do some additional calculations to interpret the residual misalignment, or one can already do this transformation before the misalignment is applied. In the latter case, we say the misalignment respects the constraints. We are mainly using this latter approach as we find it more convenient

---

data. This is, however, pretty much impossible to guarantee, and the goodness-of-alignment must be evaluated using performance studies. Relying entirely on MC simulation is not sufficient.

to aim for numbers as close to zeros as possible in the payloads after MC alignment tests. Both approaches are equivalent, and the necessary transformations are achieved by iterative application of Eq. 12.48 and Eq. 12.47, which allows solving for parameter values, which satisfy the constraints in Eq. 12.49.
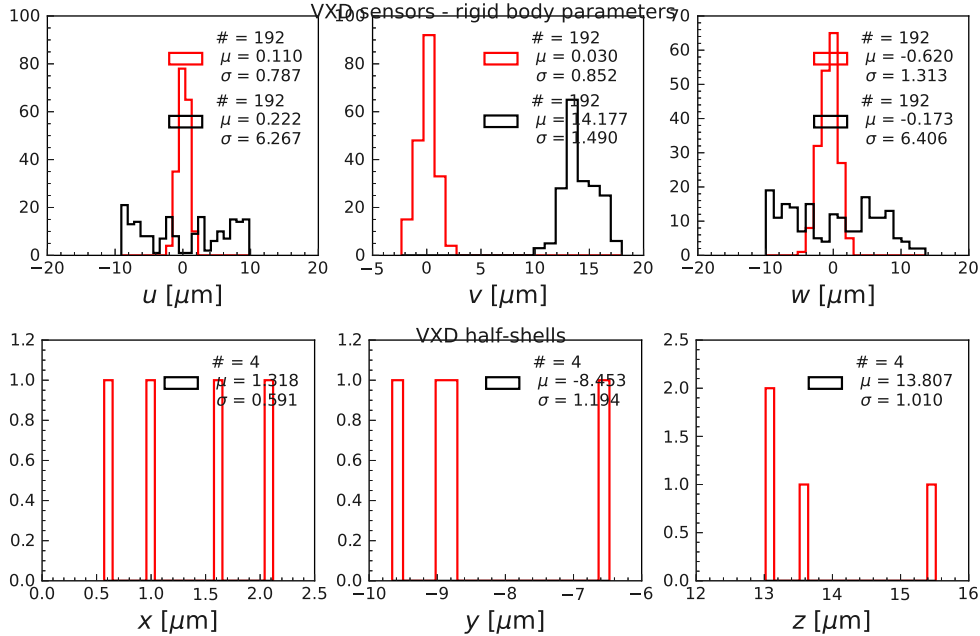
Figure 13.7: Histograms of residual misalignment for shifts of the VXD sensors (top) and half-shells (bottom) from an MC test of the baseline alignment without using constraints (black) and after transformation into constraints' reference frame (red), where sensors and half-shells are aligned together, and the coherent movements are absorbed into half-shell parameters. Note that the black distributions in the bottom plots would be identically zeros. The number of entries (sensors or half-shells), the mean μ, and the RMS denoted as σ of the residual misalignment distributions are shown (in μm).

As an example, residual misalignment from an alignment MC test (baseline alignment with cosmic rays and single tracks from $B\overline{B}$ events), where the detector was intentionally misaligned as a telescope (250 μm of initial amplitude), is shown in Fig. 13.6. This is an example of so-called ring plots, in which values of alignment parameters for all VXD sensors can be visualized. Sensors are organized into six concentric layers, where the ladder numbers are shown. The sensors in each ladder are shown from the most forward (inner-most) to the most backward. The slanted SVD sensors have a larger area, which might be slightly visually misleading. One also clearly sees that only four PXD sensors are installed in the second layer, in ladders 4 and 5.

After the alignment (with multiple iterations), where the sensors were aligned individually with respect to the CDC (and each other) without constraints, we observed residual misalignment with a scale up to 16 μm for sensor shifts. One can directly see in Fig. 13.6 that there is a global shift of all sensors towards $+v$ ($+z$ in the global system). After transformation to the system of half-shell constraints, we can see the dominant pattern is gone, and the actual scale of the misalignment at the sensor level (relative to half-shells) is only in the range of

± 4 μm.
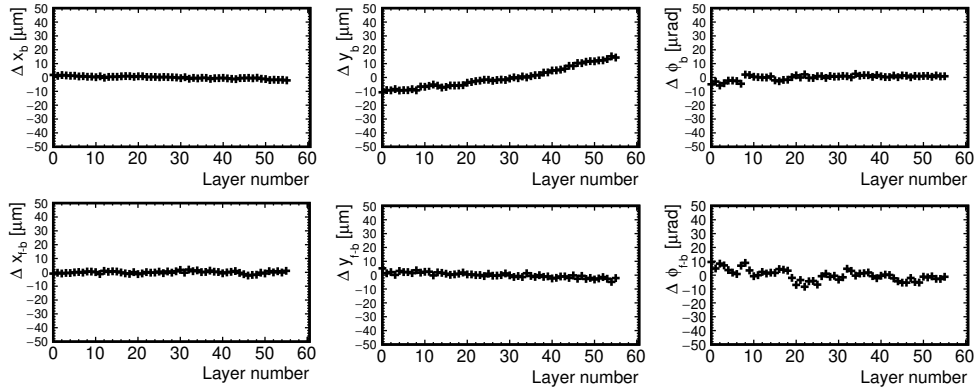
Figure 13.8: Residual misalignment for six CDC layer parameters plotted as a function of the layer number from an MC test of the baseline alignment.

The histograms of the rigid body parameters corresponding to the shifts[6] of the sensors and half-shells are shown in Fig. 13.7. From these figures, we can read the global shifts absorbed into the half-shell parameters. Their mean μ is the systematic global offset caused by a weak mode in the CDC layers, causing the vertical deformation responsible for the global residual shift in y of about 8 μm. This can also be clearly seen in our typical representation of the CDC layer alignment parameters. The corresponding residual misalignment of the CDC layers is shown in Fig. 13.8. The parameter controlling the vertical movement at the backward end-plate Δy_b (forward end-plate moves coherently as Δy_{f-b} is close to zero) shows a clear slope, and layer zero is lower by ~ 10 μm. The VXD half-shells directly follow this deformation at the inner part.

Similarly, the global shift in z by about 14 μm is a typical weak mode. Such global offsets are irrelevant for physics at this scale. These only represent the absolute systematic error in the determination of the center of the detector from tracks with respect to the reference system defined by the CDC constraints[7]. We will return to the systematic biases and weak modes in Sec. 15.2.1 and Sec. 15.4.2.

The systematic precision of the relative half-shell position determination can be estimated from the RMS in Fig. 13.7 (denoted as σ) to be an order of magnitude better, about 1 μm.

13.6.3 Misalignment for MC Studies

We have developed utilities for the generation of misalignment for two basic types of MC studies:

- to test the alignment procedure and its ability to recover from misalignment, and

---

[6]Rotations are not very interesting in this case, they change only marginally after the transformation.

[7]With a CDC radius of about 1 m, these ~ 10 μm systematic alignment errors translate to a relative precision around 10 ppm in the determination of the CDC center from (cosmic and collision) tracks.

- to study effects of various types and levels of misalignment on the tracking and physics performance.

Since the first data arrived, we have produced several special misalignment scenarios for estimating alignment systematics, discussed in the next section. In addition, we could also use the total derived misalignment from data as a starting point for the alignment tests on MC. This misalignment or its variants represent the most realistic estimates of the starting geometry after detector construction.
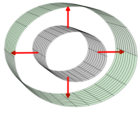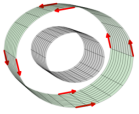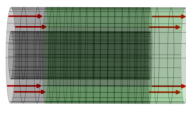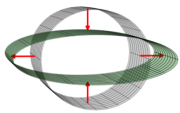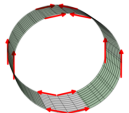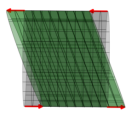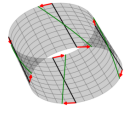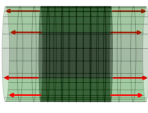
	Δr	$r\Delta\phi$	Δz
r	**Radial expansion** $\Delta r = c_{scale} \cdot r$	**Curl** $r\Delta\phi = c_{scale} \cdot r + c_0$	**Telescope** $\Delta z = c_{scale} \cdot r$
ϕ	**Elliptical expansion** $\Delta r = c_{scale} \cdot \cos(2\phi) \cdot r$	**Clamshell** $\Delta\phi = c_{scale} \cdot \cos(\phi)$	**Skew** $\Delta z = c_{scale} \cdot \cos(\phi)$
z	**Bowing** $\Delta r = c_{scale} \cdot |z|$	**Twist** $r\Delta\phi = c_{scale} \cdot z$	**Z expansion** $\Delta z = c_{scale} \cdot z$

Figure 13.9: The name, generating expression and schematic visualization for nine classical weak modes, where sensors are misaligned in radial, $R - \phi$ or z direction, depending on their cylindrical coordinates.

Two main categories of misalignment can be distinguished and generated in the alignment package:

- Random misalignment, where misalignment for each parameter and, e.g., a sensor is drawn from a normal or uniform distribution. Random misalignments can be resolved by a single Millepede II iteration even at a level of 200 μm or larger, which was verified on MC.

- Systematic or weak mode misalignment, where long-range coherent deformations are introduced. Following the work of BaBar, [70] and CMS [81, 74, 107], we introduce nine classical categories of systematic misalignment for the VXD, typical for a detector with cylindrical symmetry in a solenoid magnetic field and tracks coming from IP, illustrated in Fig. 13.9.

In addition, one can generate global misalignments of the VXD half-shells or CDC layers, or any parameter at will by modifying the payloads.

Systematic misalignment is the most challenging[8]. While from sensor to sensor, the residual misalignment is small, the total effect over all sensors can lead to a systematic deformation causing biases in estimated track parameters. The classical weak modes correspond to the invariant modes of the χ^2 for tracks coming only from the origin. For example, radial expansion can cause bias in estimated particle momenta or particle lifetimes, as the detector appears smaller/larger. Similarly, z-expansion modifies the total z scale, crucial for time-dependent *CP*-violation measurements.

In the past, the effects of these systematic misalignments at larger scales (around 200 μm usually) were studied for their physics impact, as well for the possibility of identifying them with some validation methods, like the overlap residuals. In Chapter 15, it will also be demonstrated that our alignment method can recover from such misalignments, but sometimes iterations are needed, even for systematic misalignments small in scale with respect to random misalignments, recoverable easily in a single alignment step.

13.6.4 Misalignment for Physics Analyses

One important task of the author was to provide methods for the evaluation of alignment-related systematic uncertainties for physics analyses. Naturally, the alignment cannot be absolutely perfect, and for example, for lifetime measurements, it can be a major source of systematic uncertainty.

The recommended procedure is to generate and simulate analysis-specific signal decays and reconstruct them with different misalignments applied on MC. The parameters of interest shall be extracted from the complete (signal-only) analysis, typically using a maximum likelihood estimate. The difference between the result with nominal geometry (without misalignment) is then taken as the systematic error corresponding to a given misalignment scenario.

It should be noted that maximum likelihood fits might be potentially sensitive to variations in the event content in the likelihood function. With different misalignments, some particular requirements, e.g., on the track quality, could lead to different events entering the likelihood calculation for different misalignments. This could lead to over-estimation of the alignment systematics when it is not statistically significant. The simplest solution to this problem is a simulation of a significantly larger MC sample than what is expected for real data. In this approach, the statistical error on MC can be brought down such that systematic deviations due to misalignment become statistically significant.

We introduced three complementary types of misalignment scenarios recommended for alignment systematics estimation on MC:

1. Residual misalignments from MC studies, which are results of exercising the alignment algorithm on a misaligned detector. These (we provide several options) misalignment scenarios should represent the precision of the alignment method itself.

2. Day-to-day misalignment, which was extracted from real data, when a difference of alignment constants determined from two consecutive days in

---

[8]Apart from distortion stemming from systematic effects due tracking or detector mismodeling, which are not weak modes.
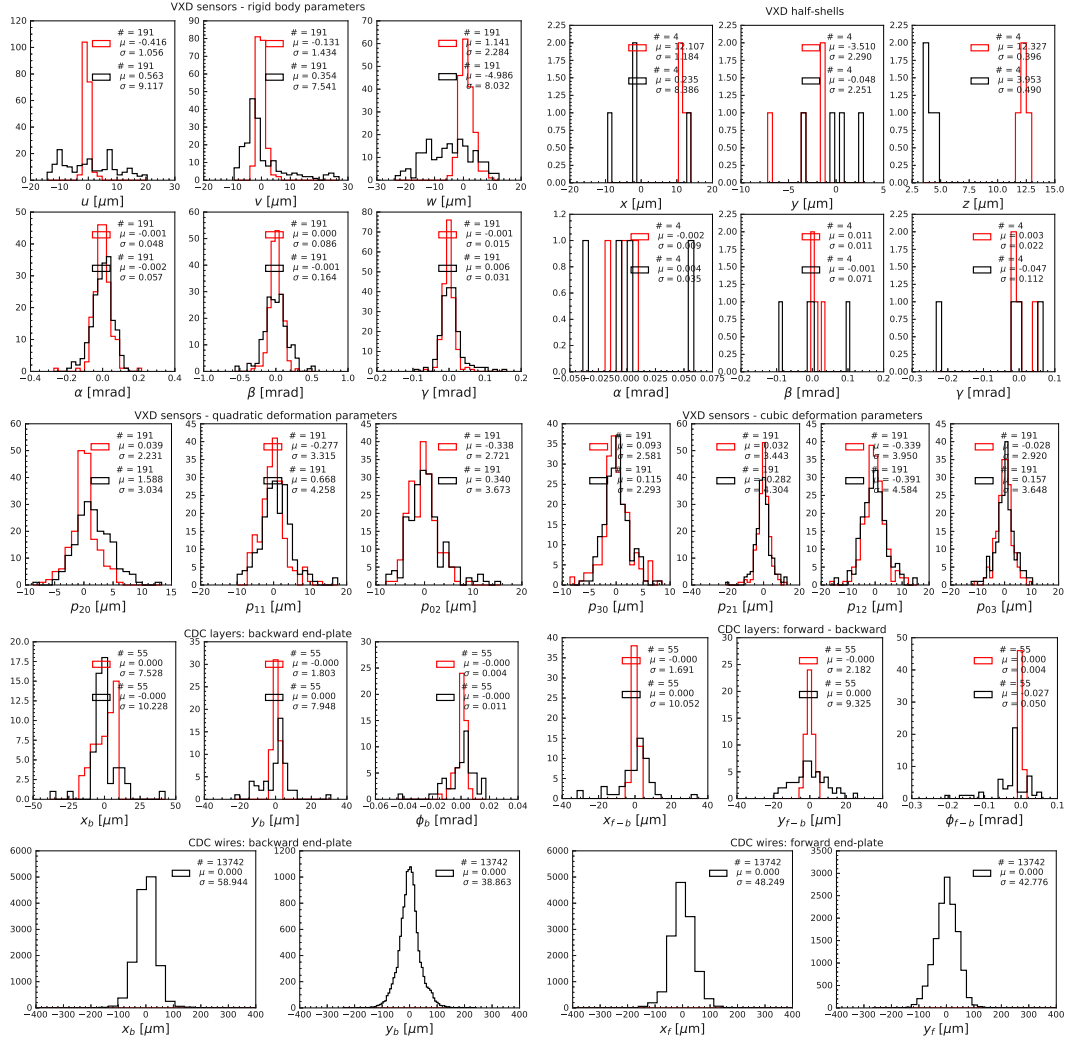
Figure 13.10: Histograms of the residuals misalignment for all alignment parameters in the model of prompt misalignment (black) and for the day-to-day alignment difference extracted from data (red). Wires (last row) are only misaligned in the prompt misalignment model.

2020 was taken. This misalignment, partially visualized in Fig.13.11, represents the typical scale and distribution of effects, which can be easily left uncorrected even after the reprocessing, especially at lower luminosities. Besides a large bowing in the second PXD layer and misaligned slanted SVD sensors, one can also identify a typical (real) CDC deformation in the Δx_b parameter. We have validated this misalignment on MC by comparing distributions of medians of residuals from all VXD sensors to those observed on average bucket data, see Fig. 13.4. This validation suggests that the statistical alignment precision is around one and two micrometers in the $R-\phi$ and z direction, respectively- It is also well reproduced by the day-to-day misalignment.

3. Model of the prompt misalignment from data, partially visualized in Fig.13.12, which was derived during 2021 data reprocessing. As we have been able to improve the alignment performance significantly, we investigated what the

difference in the alignment constants is. This difference was then (with an opposite sign) used to produce a special misalignment payload, which reproduces the previously observed alignment problems on validation with MC, as demonstrated in Fig. 13.5. This misalignment scenario represents a realistic estimation of the misalignment in the prompt calibration and a very conservative upper estimate for the reprocessing.

The latter two misalignment scenarios are compared in Fig. 13.10, also showing half-shell parameters not shown in the visualizations.

Such misalignments are the baseline for alignment systematic studies in physics analyses. In the future, more variants of the day-to-day alignment or residual misalignments can be officially introduced. However, simulation of the misaligned signal samples is computationally expensive, as the full event reconstruction must be repeated. In addition, many analyses are using the IP information, which needs to be re-calibrated for each particular misalignment. The reason is the small movement of the apparent center of the detector due to a slightly deformed CDC (by a weak mode or a real misalignment). While such an effect is not observable in data, for MC simulation, the default beam spot calibration is determined for a perfectly aligned detector, and this difference becomes observable as an incorrect IP position. Thus for every misalignment, the MC sample with di-muon events needs to be generated in addition, on which the IP position calibration is repeated. Otherwise, unrealistically large systematic effects can be estimated.
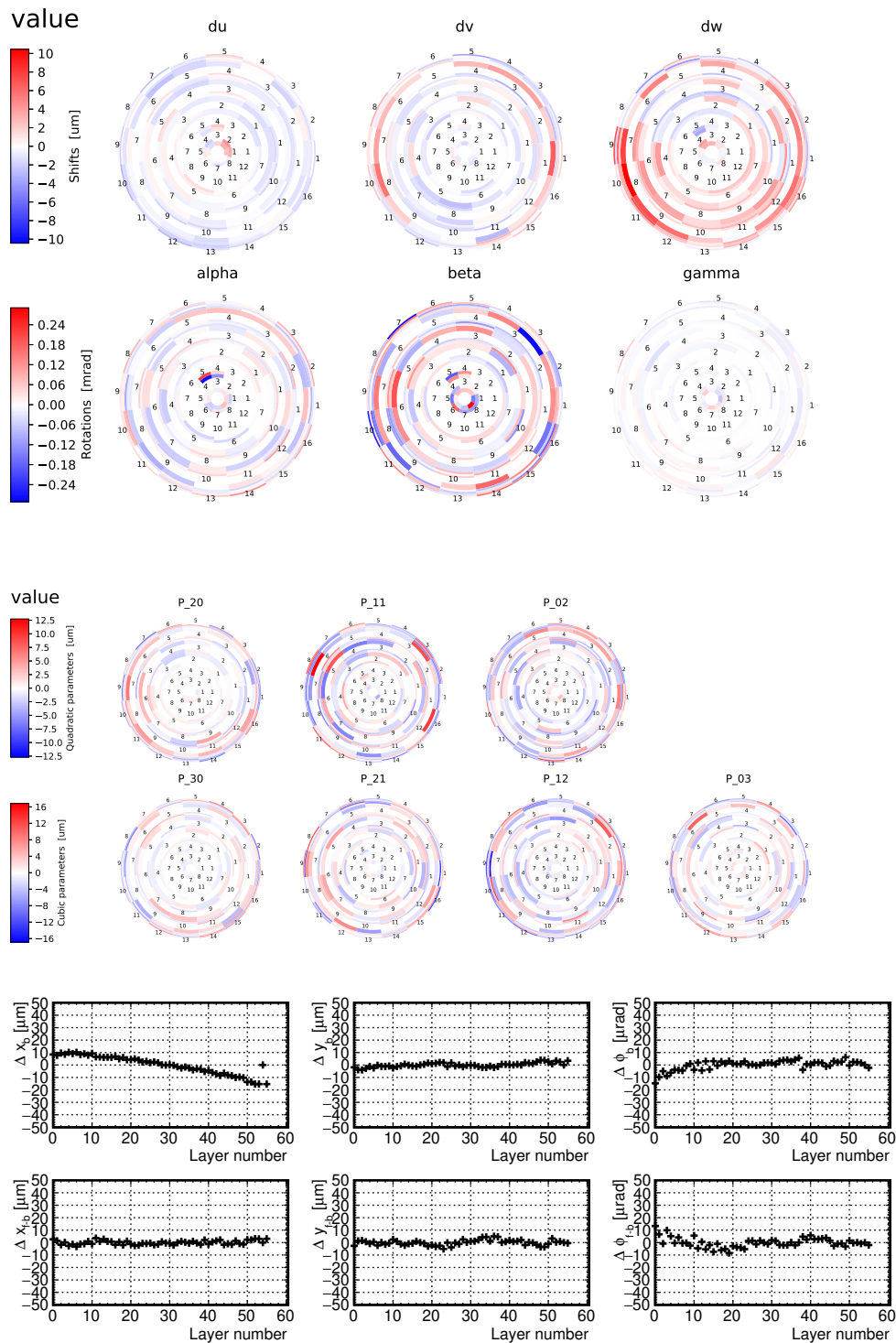
Figure 13.11: Visualization of the day-to-day misalignment extracted from data for VXD sensors' rigid body (top two rows) and surface (middle two rows) parameters and CDC layers (bottom two rows). The CDC layer alignment parameters are shown as a function of the layer number.

228
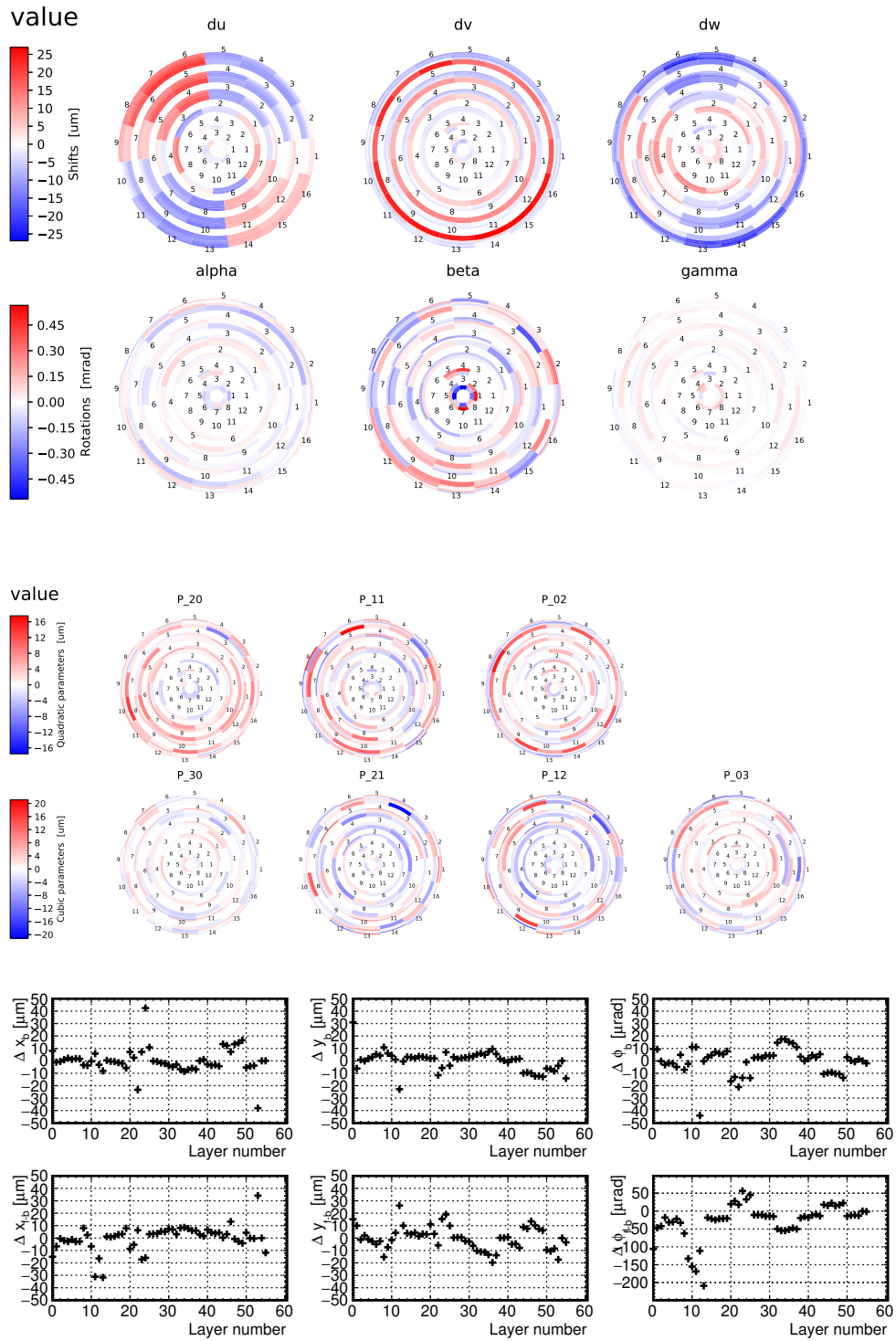
Figure 13.12: Visualization of the prompt misalignment model extracted from data for VXD sensors' rigid body (top two rows) and surface (middle two rows) parameters and CDC layers (bottom two rows). The CDC layer alignment parameters are shown as a function of the layer number.

CHAPTER 14

Belle II Alignment History

The evolution of the Belle II alignment from the first tests to the complete automated infrastructure for the determination of about sixty thousand parameters has a long history, which started with the authors' project conducted at DESY Summer Student Program in 2013 [108]. This resulted in the authors' master thesis [77], and the alignment procedure successfully applied to the 2014 beam test [109]. At the end of the master thesis, the very first test of a complete VXD alignment on simulations was performed.

Since that first implementation, the code was completely rewritten for better automation and a more straightforward addition of further sub-detectors and alignment or calibration constants, such that support for alignment of the CDC or BKLM and EKLM could be added. Features like the VXD alignment hierarchy or vertex and IP-constrained decays were added continuously. The main development was basically finished at the end of Summer 2019 when full support for time dependence was introduced. Most of the latter effort was devoted to understanding and improving the performance of the real data, as well to the finalization of procedure automation. The introduction of several features in real data, like the use of specialized data samples with cosmic rays recorded during collisions or application of IP constraint in di-muon events, had to wait for sufficient maturity of the experiment.

The following sections will briefly review the main periods where experimental results were achieved and often also published. These gradually led to the increased complexity of the alignment problem and improved performance, up to its current status, presented in this thesis.

14.1 Beam Tests

At all stages, we wanted to test the method on any available real data next to extensive simulation studies. Such data was initially only available in dedicated beam tests of subsections of the vertex detector, where the alignment method was successfully applied.

There were two major beam test campaigns at DESY in 2014 [109], and 2016 [110], which followed after a number of tests with smaller matrices and preliminary designs of the DEFPET sensors at the CERN's SPS accelerator.

In 2014, the first combined beam test of the PXD (only one layer) and four layers of the SVD was organized at the DESY beam test facility. The setup, illustrated in Fig. 14.1 was placed in a 1 T solenoid magnetic field and illuminated by a 2–6 GeV/c electron beam.

This was also the first time we used basf2 and the newly developed alignment
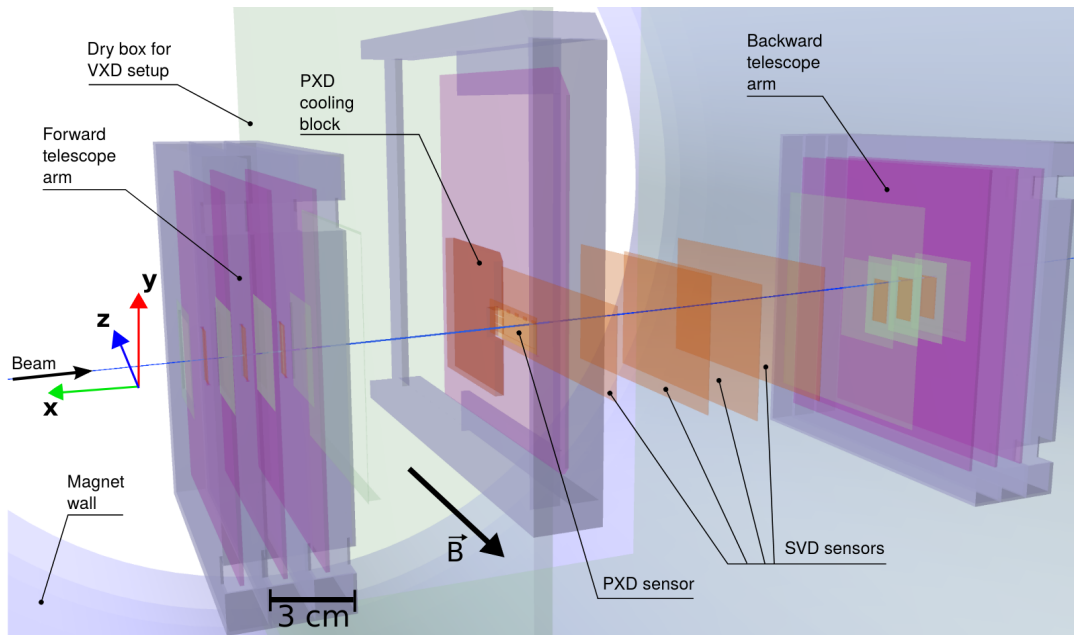
Figure 14.1: The geometry in basf2 used for the 2014 beam test simulation and reconstruction with a single PXD matrix and four SVD sensors, with three EUDET telescope modules in front and behind the VXD setup. Different colors denote different material types: silicon (light orange), copper (orange), aluminum (blue), electronic boards (magenta) and plastic (green).

procedure. Only a limited number of parameters can be left floating for alignment in such a setup. The reference system was defined by fixing one of the sensors (we usually used the first SVD layer). In addition, three sensors had to be fixed in the vertical direction, corresponding to the direction of beam bending by the magnetic field, to remove the weak bending mode. While this weak mode could be removed by also using primary positrons in the beam or by combining with data without the magnetic field, this could not be practiced in the beam tests. We did not switch to positron beam, although that is in principle possible at the DESY beam test facility. Several runs without the magnetic field were taken, but the data could not be combined due to large ($> 40~\mu$m) movements. The last sensor along the beam was also fixed in the horizontal coordinate to fix the otherwise unconstrained global rotation along the vertical axis.

In 2016, a full section of the vertex detector (to be installed during the Belle II commissioning phase) was the subject of beam tests at DESY. An example of track-to-hit residuals and improvement in momentum estimation is illustrated in Fig. 14.2. While full SVD ladders were installed, only the central sensors were sufficiently illuminated, and thus the alignment setup was almost identical to the 2014 configuration.

The beam tests were essential in many aspects. Next to gaining operational experience with the complex setup, we solved many issues. Some of the challenges for alignment were data clean-up, particularly masking of hot pixels/strips and proper configuration of track finders and outliers down-weighting. Also, problems with the correct matching of hardware and software channel counting directions were identified and fixed, as well as issues in the GBL and GENFIT2 interface. The recorded data allowed for the characterizations of the sensors, verified cooling
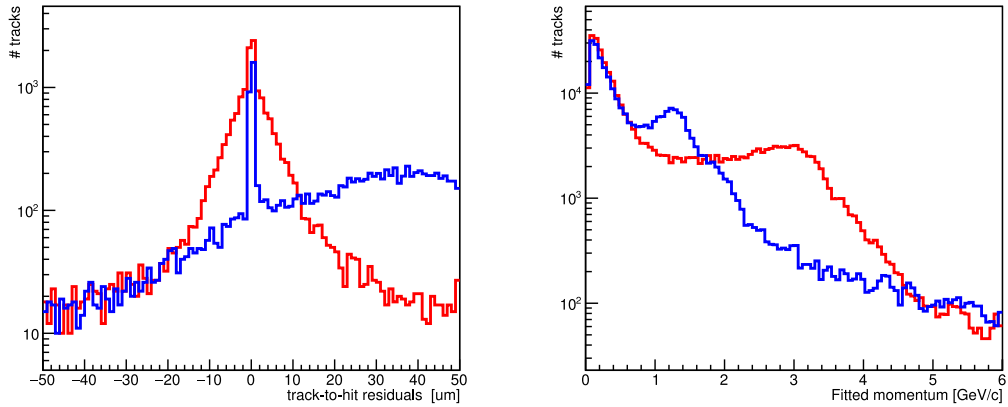
Figure 14.2: Track-to-hit residuals (left) in the horizontal (local u) coordinate for the second SVD sensor in the 2016 beam test before (blue) and after alignment (red). Improvement of the track momentum estimation for a 3 GeV/c electron beam after alignment (right).

procedure, proved the concept of PXD data reduction, validated trigger and data acquisition chain, track finding and fitting, alignment, as well, as access to the condition database.

14.2 First Collisions and Phase 2

The start-up of the SuperKEKB accelerator and the Belle II experiment was divided into three phases:

- Phase 1, where the SuperKEKB accelerator was commissioned, still without beam collisions and the Belle II detector at the IP position at the beginning of 2016. After detector roll-in into the beam-line position in 2017, a global cosmic ray test was performed, which resulted in the first full calibration and alignment of the CDC [21]. This was handled by stand-alone procedures developed by the CDC group and served as a baseline for the next data-taking period.

- Phase 2, taking place in Spring 2018, was devoted to SuperKEKB and Belle II detector and background commissioning. The vertex detector was reduced to one section of 18 sensors and accompanied by specialized background radiation detectors.

- Phase 3, starting in March 2019, with the full Belle II detector.

The first collisions were recorded on 26 April 2018. The first VXD alignment was provided just a couple of hours later [111]. The existing CDC alignment from the global cosmic ray test was used as a reference. Full tracks, including CDC hits, were provided to Millepede II, but all CDC alignment parameters were fixed. We aligned all 108 rigid body parameters (18 sensors $\times$ six parameters) simultaneously using only the collision data. The residual distributions integrated over all VXD sensors are compared before and after alignment in Fig. 14.3.
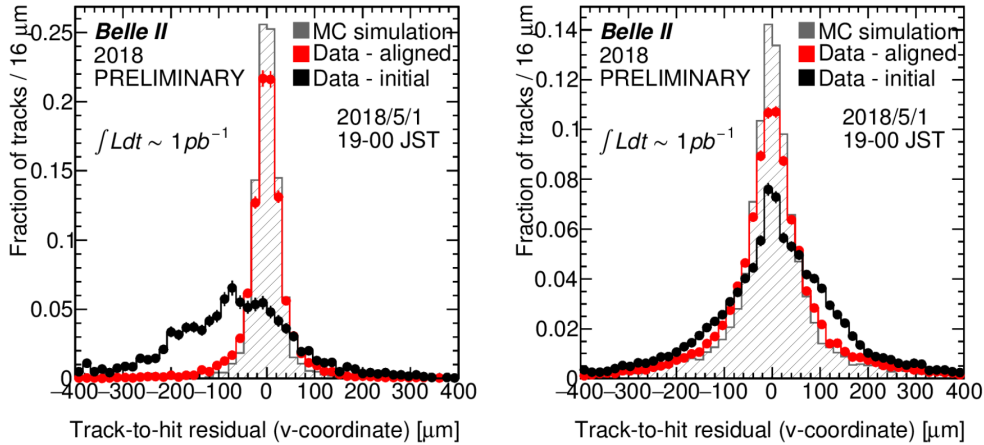
Figure 14.3: Track-to-hit residuals in the v (left) and u (right) local coordinate integrated over all VXD sensors in Phase 2 before (black) and after (red) alignment. A simple particle gun (electrons, positrons, protons) is used to generate 20×10^3 tracks originating near the IP for comparison to MC without any misalignment. All histograms are normalized.

The large angular spread of the incident tracks allowed for the first time to study sensor deformations, where the necessary validation work was conducted by J. Kandra [105]. Using the same geometrical considerations as in Sec. 12.3, the reconstructed w-coordinate residual (for inclined tracks) is defined as

$$r_w = \frac{r_u}{\tan \psi} \quad \text{or} \quad r_w = \frac{r_v}{\tan \zeta}, \tag{14.1}$$

where r_u, r_v are the residuals in u and v direction, and ψ and ζ are the respective track incidence angles. The average w-residuals from many tracks binned over the sensor area, weighted by their respective precision, proportional to $\tan^2 \psi$ ($\tan^2 \zeta$), can be used to determine deviations from the flat sensor surface. Such a distribution is shown in Fig. 14.4 (left) for one particular SVD sensor with a simple rigid-body alignment.

As we have observed significant deformations of the sensors, in particular in the SVD, where the total deformation amplitude can reach almost 0.5 mm, the alignment and reconstruction were extended to correct for this effect, as described in Sec. 12.3. This led to further significant improvements in the residual distributions. The effect of correcting up to second and third-order surface deformations is demonstrated in Fig. 14.4 (middle and right).

14.3 SVD and VXD Commissioning and First Cosmics in Completed Belle II Detector

After Phase 2 data-taking finished on 17 July 2018, the VXD assembly and installation preparations started. During this period, only cosmic ray tracks triggered by scintillator plates were used to independently align the two SVD half-shells. After the installation of PXD into the SVD, we aligned the setup again. A snapshot from the event displays of the two configurations is shown in Fig. 14.5.
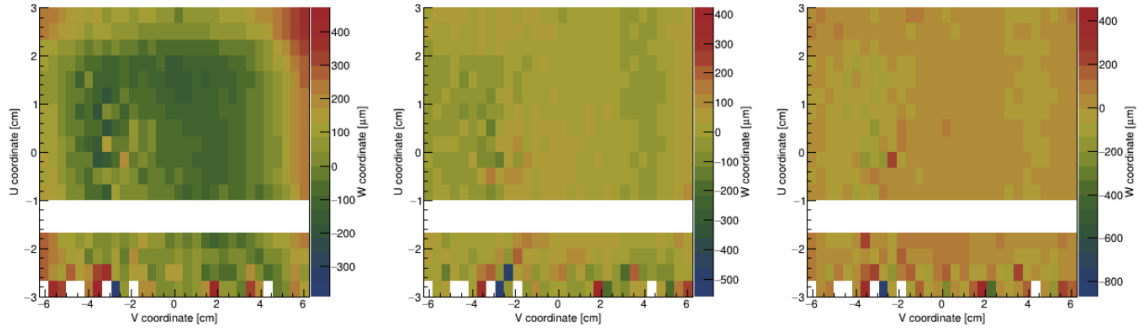
Figure 14.4: SVD sensor 4.1.2 in Phase 2 and the distribution of the average w-residuals before surface alignment (left), after using 2nd order deformations (middle) and including also 3rd order deformations (right). The blank space corresponds to a broken APV chip resulting in an inactive part of this particular sensor.

Due to challenges in the PXD ladder assembly, the full two-layer PXD could not be completed in time, and so it was decided to proceed with only the first PXD layer fully equipped. Only four DEPFET sensors were available for the second layer, which was intentionally oriented to cover a problematic sensor in the first layer. This sensor (backward sensor in the third ladder with ID 1.3.2) was later fully disabled.
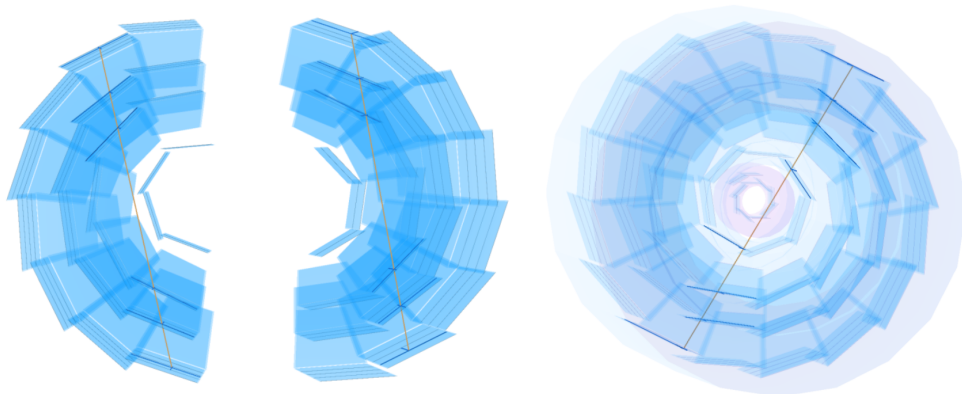
Figure 14.5: Visualization of the two SVD half-shells (left) and the completed VXD (right) in the basf2 cosmic reconstruction.

This was the first time we had to use alignment constraints, typically fixing the total average shift and rotation of all SVD sensors. For the first time, we got an almost complete picture of the total VXD misalignment after assembly, including sensor deformations.

After VXD installation into the Belle II detector, two periods with a global cosmic run followed, first without magnetic field and second after solenoid excitation, when the end-yokes hosting the forward EKLM had been closed.

Already in this data, the most serious historical alignment problem could be observed. Unfortunately, the related validation arrived only much later, when we fixed the issue soon after the first collisions in 2019. The problem can be seen on residuals collected from sensor overlaps [106]. With the cosmic sample, a
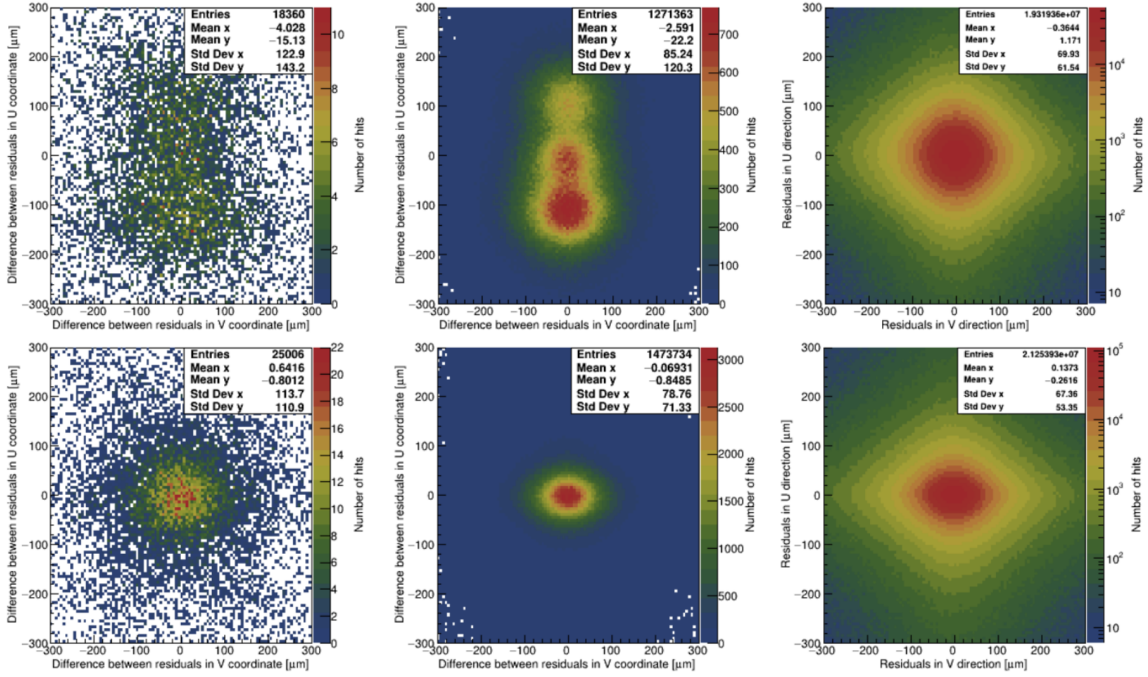
Figure 14.6: Overlap residuals for different (left) and same sensor number (middle) and residuals without an overlap (right) from Spring 2019 cosmic ray data before the fix of SVD barrel sensor pitch (top) and after (bottom). [106]

strange *snowman* structure was identified in Fig. 14.6. We show overlap residual distributions in the same figure after a fix, discussed at the beginning of the next section. Since this fix, no further problems were identified in this overlaps residual validation, which requires a large amount of data and is only sensitive to coherent VXD deformations with amplitudes larger than several tens of micrometers.

14.4 Phase 3 and Early Physics Data

In March 2019, the first collisions in the full Belle II detector were recorded. Soon after, the aforementioned issue was identified in the dependence of track transverse impact parameter d_0 on its azimuthal angle ϕ_0. As at Belle II, the nominal IP position is about 0.5 mm from the origin of the alignment reference system; this induces a correlation of the helix parameters extrapolated to the POCA to the origin. This distribution with just a nominal geometry before any alignment, with the initial alignment with collision data, and after a fix is shown in Fig. 14.7. After a number of investigations, we could not find any problem in the alignment itself. By careful inspection of the inputs, the SVD group found a mistake in the XML configuration file used for geometry construction. The problem was a slightly wrong pitch of the rectangular SVD sensors in the u-coordinate. This led to effects resembling radial expansion (about 100 μm for the barrel SVD sensors), which can also be identified with the overlap residuals in Fig. 14.6 by comparing to MC studies with various misalignments [106].

We have entirely recomputed the alignment with the correct SVD sensor pitch and obtained rigid body parameters shown in Fig. 14.8. At that point, the VXD was aligned sensor by sensor with respect to the CDC (still with its stand-alone
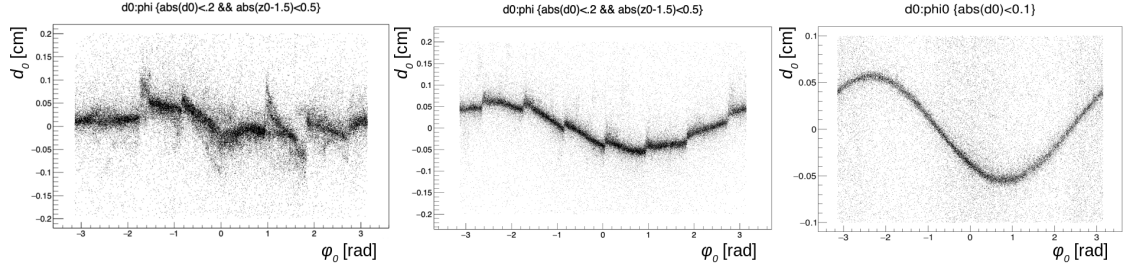
Figure 14.7: Transverse impact parameter d_0 for collision tracks from early 2019 data as a function of track azimuthal angle ϕ_0 without any alignment (right), with alignment before the SVD pitch was fixed (middle) and after (right).

alignment derived before collisions), without alignment hierarchy enabled. First, one can identify a large global shift with respect to the CDC reference system origin, about 1 mm towards $+z$ (v). Notable is the scale of the initial misalignment and that this was recovered without any dedicated pre-alignment procedures completely using Millepede II[1]. Many sensors are displaced more than 300 μm from their nominal position, which was considered a limit of VXD mechanical construction precision.
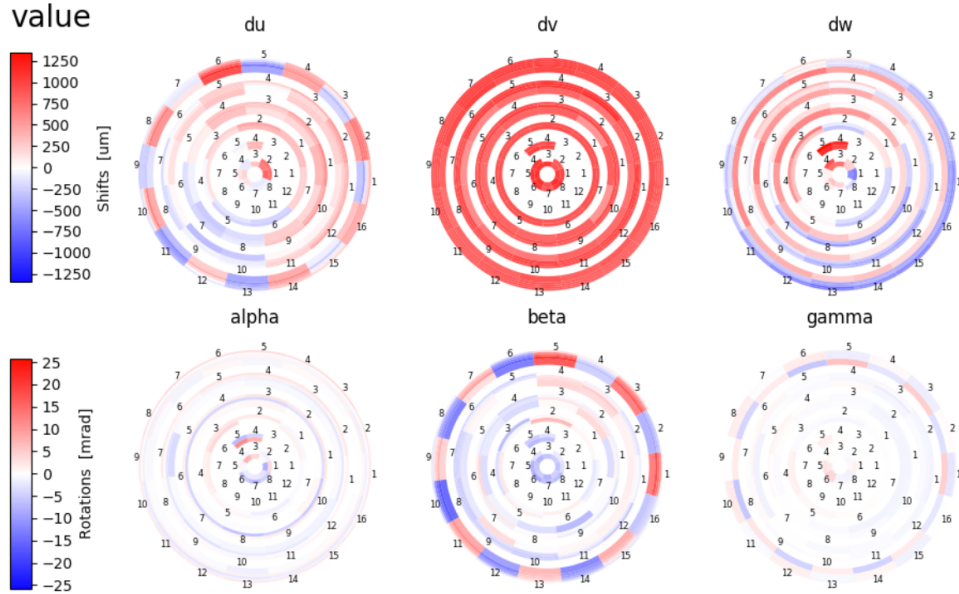
Figure 14.8: Rigid body parameters of the VXD sensors showing the total VXD misalignment with respect to design geometry, determined sensor-by-sensor.

The second feature is the alternating pattern for the SVD sensors in the outer layer in the u ($R - \phi$) shift and angle β. This is known as a *zebra effect*, and it is caused by a slightly wrong diameter of bending of the cooling pipes connecting doublets of outer SVD ladders at the forward side. This pattern can also be observed in some of the deformation parameters, namely P_{11}. The determined surface alignment parameters up to third order are shown in Fig. 14.9.

---

[1]About five iterations are needed, where initial measurement errors need to be scaled and requirements on track quality removed.
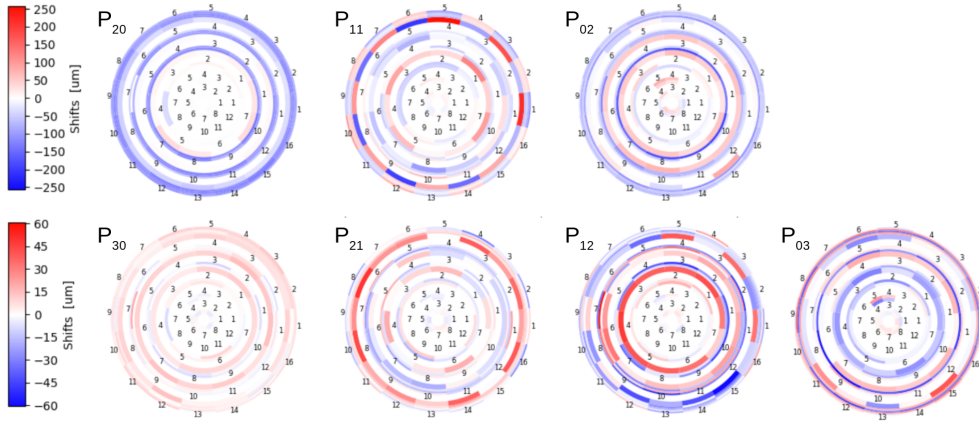
Figure 14.9: Surface deformation parameters of 2nd (top) and 3rd (bottom) order for the VXD sensors showing the total VXD misalignment with respect to design geometry, determined sensor-by-sensor.

The amplitude of bowing in the u-direction (P_{20}) is apparent in SVD layers starting from layer four (L4), reaching over 100 μm. On the other hand, even larger bowing is present for L4 in the v-direction (P_{02}). Note that the total deformation amplitude is the sum of all the contributions, which is larger than 300 μm for some sensors. While the scale of second-order deformations is up to 250 μm, we only observe amplitudes up to 60 μm for third-order deformations. At that time, higher-level deformations were not considered; their scale was only about 10 μm. We started to use also this level of deformations since the 2021 data reprocessing.
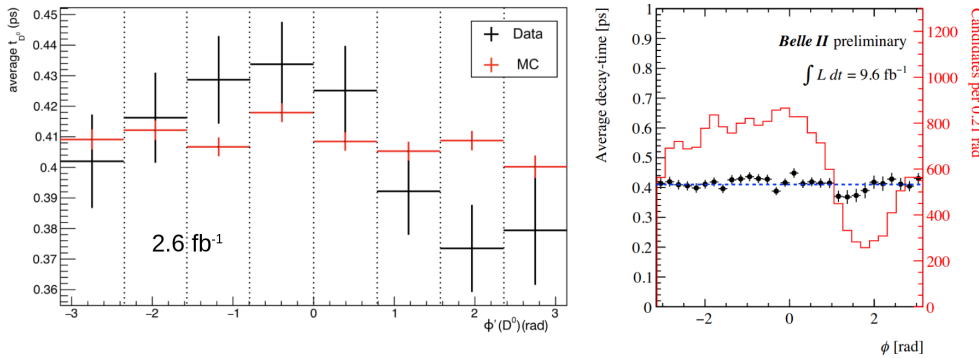
Figure 14.10: Preliminary analysis of D^0 lifetime dependence in $D^0 \to K^-\pi^+$ on the azimuthal angle of reconstructed D^0 momentum in early 2019 data before CDC alignment with Millepede II (left) [112]. The absence of any significant variations after we introduced the new alignment is demonstrated using a larger dataset, including data from Autumn 2019 (right).

Already with the early physics data, the Belle II Charm analysis group started with their first look at D^0 lifetime measurements [113], as it was recognized that Belle II should be able to improve the current most precise measurements already with less than 100 fb^{-1}. These early studies identified [112] a bias in the lifetime as a function of D^0 flight direction, shown in Fig. 14.10 (left). Other indications for similar problems in analyses that utilize the IP position information lead to the

first global VXD and CDC alignment. At that point, the wire-by-wire alignment was taken from the previous older stand-alone CDC alignment, and only the layers were aligned together with VXD. At that time, it was not known that large deformations CDC could easily happen—this came as a large surprise later, after the first data-taking period in 2019 finished in the Summer. Nowadays, we can understand better what happened. Since its last alignment, CDC has been deformed. Even with VXD realignment, an inconsistency is introduced in the IP position seen by the VXD and the (misaligned) CDC. All these problems disappeared after we switched to the global VXD and CDC layer alignment. One example from the D^0 lifetime analysis with much more data from a later period is also shown in Fig. 14.10 (right).
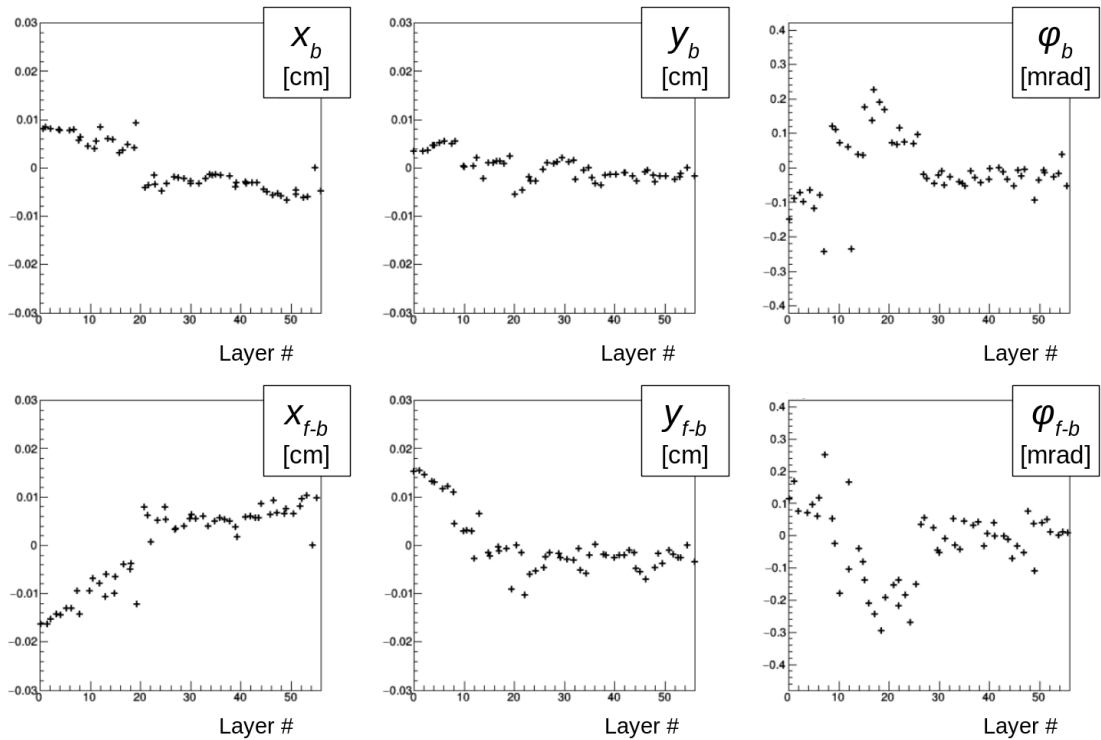
Figure 14.11: Initial alignment parameters of the CDC layers, showing the total misalignment with respect to surveys (displacements). The horizontal axis shows the layer number. The range of vertical axis is $\pm 300 \mu$m for shifts and $\sim \pm 0.4$mrad for rotations.

The initially determined CDC layer alignment constants are shown in Fig. 14.11. Note that these are determined on top of displacements already implemented in the geometry. The corrections to shifts are under 200 μm, and we can see that a larger misalignment is generally present in the inner layers.

Most of the features from these plots, as well as the previous for VXD, are still present, only much smaller corrections are induced from alignment to alignment, and in addition, the alignment has a tendency to oscillate, with occasional *re-alignment* of the detector after a misaligned period. This can be nicely demonstrated on alignment validation plots for the PXD sensors as, for example, in Fig. 14.12, showing data just before the Summer 2019 shutdown. Here an initial constant alignment is used to see time-dependent deviations. This method [105,
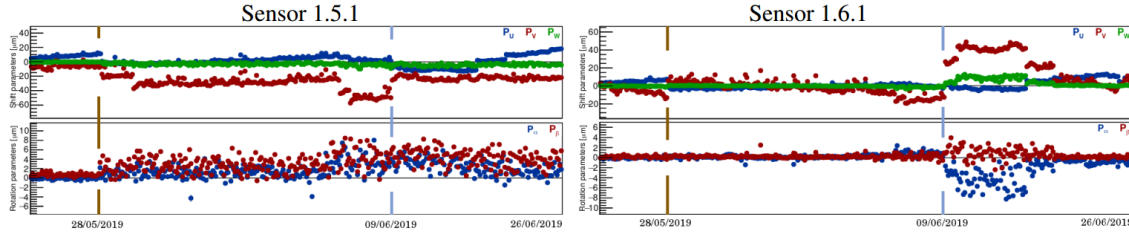
Figure 14.12: Alignment validation for rigid body parameters u (blue), v (red), w (green) (top) and surface parameters P_{20} (blue) and P_{02} (red) (bottom) showing time-dependence with two major QCS events denoted by vertical lines. On 28 May 2019, a QCS power supply issue occurred, while on 9 June 2019, a quench happened in the QCS magnets. [114]

114] is based on the determination of the local w-residual like for validation of sensor deformations. This allows extracting estimates for rigid-body (except angle γ) and surface misalignment. However, as the shifts in u and v are subtracted first using means of the residual distributions, the results can be misleading. This method cannot fully decouple the responsible alignment degrees of freedom and, for example, shows apparently too large misalignments in v, which are, in fact, combinations of sensor rotation and shift along w. However, this method is sufficient for qualitative insight into the time-dependence.

Two major events could be correlated to the large misalignment changes, one due to a power supply issue and the second due to a quench in the QCS magnet system. However, besides those, one can see many more steps and sometimes continuous time evolution, for which no explanation exists yet. It is also interesting that the parameters eventually return very close to the initial (aligned) positions, at least for some time. This highly complex time dependence was analyzed for a long time to find a reasonable alignment strategy. Until then, alignment was computed for all VXD sensors and CDC layers only once per data block, spanning even a month. One of the main limitations preventing more frequent alignment was that cosmic ray tracks, essential for a good alignment, were recorded only in dedicated runs without collisions and beams. Thus we could not be sure if the alignment is compatible with the physics runs taken at a different time.

We generally observed a good performance for stable alignment periods or when a dedicated alignment was performed, even with ignored time dependence. In the initial collision data from 2019, after the above-mentioned alignment issues were fixed, a vertex resolution study was performed, particularly for the transverse direction. The results are shown in Fig. 14.13, where two independent methods are utilized for the estimate and compared to MC simulation [115].

The first standard method compares the helix parameters of two tracks, typically from di-muon or Bhabha events. These tracks, extrapolated to the average position of the IP, should have the same origin. The comparison of their transverse impact parameters d_0 is shown on the left. On the right, we utilize the tiny IP size in the nano-beam scheme, where for the vertical direction $\sigma_y \sim 1.5 \ \mu$m for the early physics runs. This is much smaller than the detector resolution, and for horizontal tracks ($\phi_0 \sim 0 \pm \pi$), the spread of individual tracks coming from the IP should be entirely dominated by the VXD resolution in d_0. The width of this distribution should follow $\sigma(d_0) = \sqrt{\sigma_i^2 + (\sigma_x \sin \phi_0)^2 + (\sigma_y \cos \phi_0)^2}$,
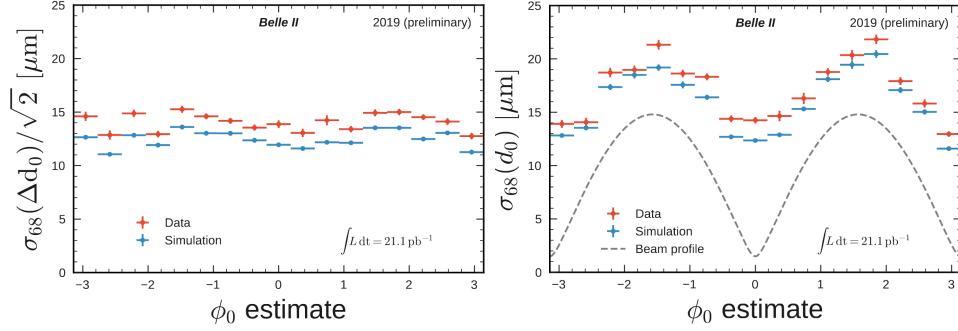
Figure 14.13: Spread of d_0 difference (left) between two tracks from IP and the spread of the d_0 distribution for the tracks (right) as a function of track azimuthal angle ϕ_0 for early 2019 data (red) and MC simulation (blue). The dashed line shows an ideal distribution for a detector with infinite resolution, assuming beam size expected from machine parameters and validation measurements [115].

where σ_i is the detector resolution and $\sigma_x \sim 15$ μm is the IP size in the horizontal direction. Both methods agree on the detector resolution (averaged over the momentum range of mostly Bhabha events) in the transverse impact parameter $14 \pm 0.1_{\text{stat}}$ μm. Both methods also agree on the slight discrepancy with respect to the MC simulation. This is now understood as a too optimistic simulation of the SVD sensor resolution.

Towards the end of the first data-taking period in 2019, we also experimented for the first time with the CDC wire-by-wire alignment using Millepede II, together with VXD and CDC layers. A large amount of cosmic data without a magnetic field was taken with a closed Belle II detector after a fire accident at the LINAC equipment. This accident resulted in a break of a couple of weeks when cosmic data without and later with magnetic field were taken (on a special request of the alignment group). A dedicated full CDC calibration was also available, and thus, we combined this data with collision tracks to determine an average wire alignment from scratch (after resetting the existing old wire-by-wire alignment). As we verified that this new alignment improved several validation plots with average tracking biases, we started to use it officially, despite no corresponding MC studies being available at that time. This wire alignment was, in fact, left unchanged for production until the reprocessing in 2021.

14.5 Towards Automation and Higher Luminosity

The second data-taking period in 2019 following the Summer shutdown marked several important steps toward alignment improvements and automation, fully commissioned for the prompt calibration loop starting in 2020. These were mainly:

- New cosmic skim, where cosmic ray tracks recorded during collisions started to be stored in a dedicated skim. This allowed us to combine cosmic and collision data consistently for the first time.

241

- Automation with Airflow [104], which allowed us to drastically simplify the regular calculation of the calibration constants.

- Finalized development in the alignment package with newly supported time-dependence in the global alignment.

In fact, the new time-dependent feature was used for the reprocessing of Autumn 2019 data, which followed shortly after the data-dating finished in mid-December. The data set was split into several blocks manually, using only a guess and observations from preliminary results and validations. Thanks to an increased accelerator luminosity, this was also the first time for IP-constrained di-muon events to be used, and also the IP alignment was performed simultaneously with the VXD and CDC layer alignment. However, it was realized that the IP is significantly intra-run dependent, so the IP constraint's full and reliable utilization had to wait for a more advanced IP calibration. While preliminary IP calibration was available, this was only performed once per run and significantly overestimated IP size in y.

It was soon realized that the Autumn CDC layer alignment significantly differs from values before the Summer shutdown and that this change actually occurred soon after the data-taking restarted, but almost no collision data was taken yet.
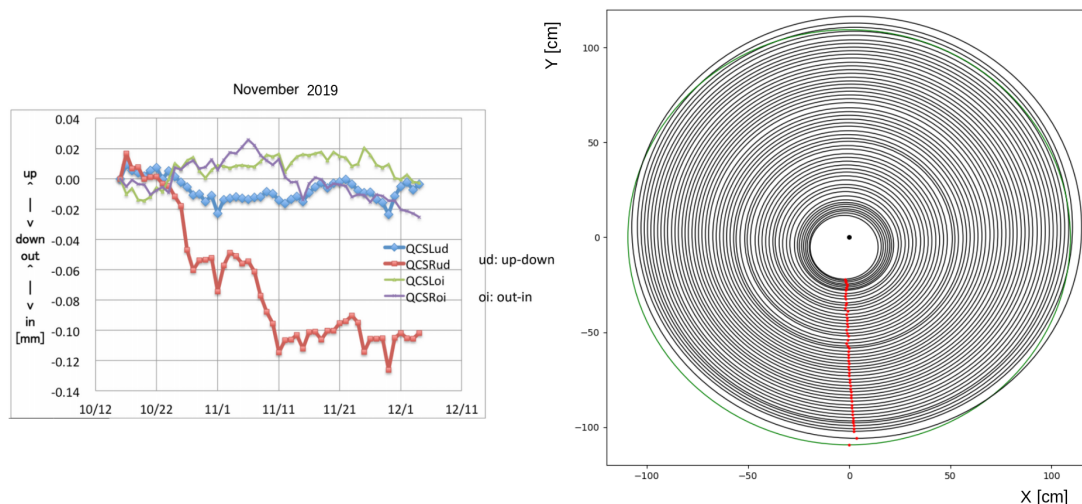
Figure 14.14: The data from a monitoring system of gaps between the QCS magnet and the inner CDC cylinder from Autumn 2019 (left) shows a large vertical drop by about 100 μm from late October for one of the gap sensors (red). Right: exaggerated visualization of CDC layer alignment difference in late November data with respect to alignment from the beginning of Summer 2019, showing the mostly vertical deformation with the sinking of the inner part. The green layer is the inactive layer 54, for which no corrections are determined.

In Fig. 14.14, we show an exaggerated illustration of the layer alignment difference between Summer and (most of the) Autumn data. The scale and timing of the deformation have been correlated to readings from a hydrostatic leveling system and sensors monitoring the gap between the CDC forward outer support cylinder and the QCS magnet. As the experimental hall and the accelerator tunnel have different depths of their foundations (the hall is founded deeper), they

242

react differently to changing underground water levels. Relative movements of the tunnel and experimental hall then lead to the accelerator's movements relative to the detector. Due to issues with space for VXD cables during installation, it is very likely that the accelerator can get into unexpected mechanical contact with the CDC, and its movements cause deformation of the chamber, mostly in the vertical direction with a scale up to 100–150 μm. Later, with the more frequent and automated alignment, we learned that changes in the QCS magnetic field also cause CDC deformations, typically when the magnets are turned off/on during a maintenance day.

During 2020, the automated baseline alignment was performed regularly for each bucket. A number of studies were performed to understand the time-dependence better, how to suppress it in the alignment, and in particular, what is the possible physics impact. Also, several other issues were revealed in the validations, and as we started to suspect large CDC deformations can easily happen, it was also desirable to recompute the wire alignment regularly. The baseline alignment and its extension to the full alignment with the CDC wires are discussed in the following two chapters, followed by a chapter that compares the performance of the two approaches on data.

CHAPTER 15

Baseline Alignment Studies

This chapter is devoted almost entirely to MC studies of the first fully deployed simultaneous alignment of the VXD and the CDC with about 3×10^3 parameters. However, for the CDC, only layer alignment is considered. We fix the wires to previously determined positions (zeros for simulations). We refer to any such alignment without wires as baseline alignment. At the time of writing, most of the Belle II data was still calibrated by the baseline alignment running in the prompt calibration loop, see Sec. 13.3.2. Baseline alignment was not using IP-constrained di-muons, as the fine time-dependent IP calibration was regularly performed only after the alignment for real data. Naturally, this baseline alignment started for the VXD only. Quite soon, we added support for a relative VXD and CDC alignment, and soon after the arrival of the first collision data, the CDC layer alignment was officially introduced, as also briefly discussed in Chapter 14.

We have decided not to expose the reader to the full path, which led to the baseline alignment configuration. We even do not aim for an optimal alignment at this stage, as the final precision is achieved by the advanced full alignment described in Chapter 16, where on the contrary, we attempt to explain the reasoning step by step. Following a conservative philosophy, we look for the worst case to set upper limits on alignment precision in MC studies. The focus will be on cross-checking the final procedure configuration, subject to reasonable variations or additional studies to demonstrate some of our arguments explicitly.

Although the reprocessing alignment will supersede this result, the baseline alignment is essential to the entire workflow. The final alignment depends on a well-aligned detector to provide a good linearization point, which is the baseline alignment's primary purpose.

15.1 Alignment Configurations

Many possible variations can be easily derived from the baseline alignment. First, in reality, the alignment will likely start with cosmic-ray tracks without a magnetic field, which we have almost completely skipped in this thesis. While this sample and possibly also cosmic ray tracks in the magnetic field can be used for a very reasonable starting alignment, especially for the VXD internal alignment (with sufficient statistics), the aim is to present properties of setups used to derive alignment for physics analysis. Thus we also do not explicitly discuss trivial variations with limited degrees of freedom (e.g., standalone VXD alignment). In practice, this is never recommended, as CDC deformations are frequent, and neglecting them can lead to systematic biases.

The basic rule for a good alignment is to combine cosmic ray and collision

tracks, and this is mostly common to all configurations presented. Collision-only single tracks should never be used for the global alignment or standalone VXD alignment. The resulting systematic displacement of some sensors is larger than 30 μm [71], which is not acceptable. In most cases, only the hadron sample is used for collision tracks. In some cases, we also include single muon tracks from di-muon events and occasionally demonstrate the effect of IP-constrained di-muons. Typically roughly 0.5×10^6 to 1×10^6 cosmic ray tracks are combined with $0.5 - 2 \times 10^6$ generic $B\overline{B}$ events and/or $0.5 - 1 \times 10^6$ events from the di-muon MC sample.

We do not simulate backgrounds for the misalignment studies. The effect of adding backgrounds is only an increased fraction of rejected data (outliers), while the alignment result is almost unaffected. On the other hand, we have observed some dependence on the track quality cut (p-value requirement in the `MillepedeCollector` module, see Sec. 11.8 and Sec. 13.4.2), which is since then always imposed.

Regarding the alignment parameters, in addition to the fixed CDC wires, we also do not consider quartic deformations of silicon sensors. In addition, the VXD ladders are not aligned in any of the presented configurations: either no hierarchy is used, or the sensors are placed into half-shells. This later became the default for real data. Regarding CDC layer constraints, we used all six constraints fixing CDC as a rigid body and the twist constraint (seven constraints in total), see Sec. 12.6.2.

15.2 Alignment Precision in Misalignment Studies

We have performed several simulation campaigns to cross-check the ability of the alignment to recover from a wide range of possible misaligned configurations. Before the first data arrived, these tests were limited to random misalignment and individual weak mode distortions for the VXD. At that time, we only probed misalignments at levels up to 300 μm. With the first cosmic ray tracks, it became apparent the actual misalignment was more prominent than assumed initially. We identified about 1 mm shift of the VXD relative to the CDC in the z-coordinate, significant deformations of the SVD sensors or the *zebra effect* (see Chapter 14 and Figs. 14.8 and 14.9) in the outer SVD layers. Naturally, we wanted to verify that we can also align such a configuration from scratch in simulations.

This section will use several variants of the baseline alignment to assess the alignment performance, evaluated from the residual misalignment. We probed those in different campaigns, with varying versions of software, relative fractions of input data samples, or configurations for constraints for the CDC or VXD hierarchy. All such variations are also deemed plausible in real data and should yield a very good alignment. This section is devoted to the demonstration of this last statement.

We have selected four MC campaigns for this evaluation (the color code is used for most histograms in the various figures in this section):

1. **Test with a completely realistic misalignment** imported from data, performed soon after first collisions. At that time, we did not use the VXD
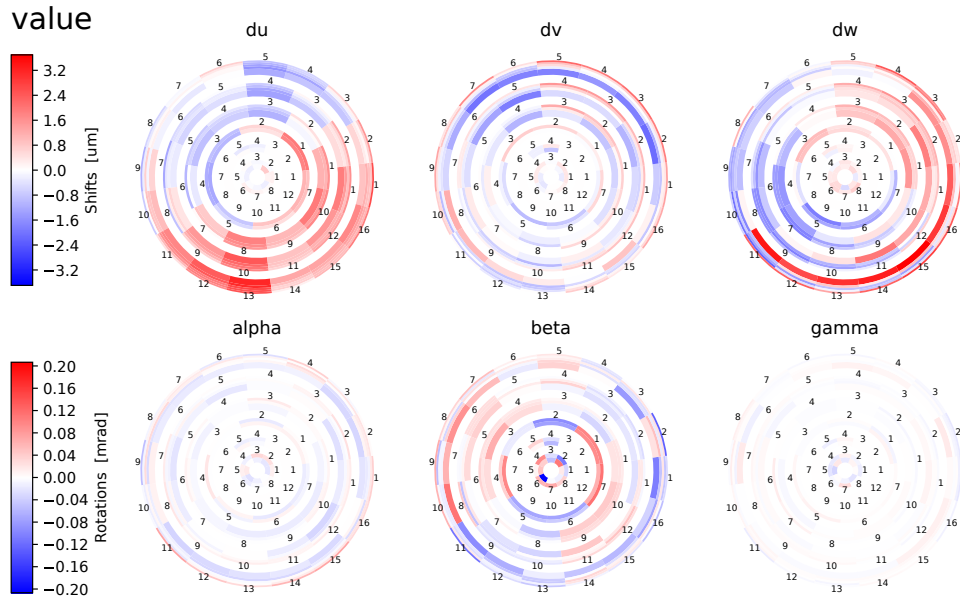
Figure 15.1: Residual misalignment for rigid body parameters of the VXD sensors after MC test of a realistic misalignment imported from data.
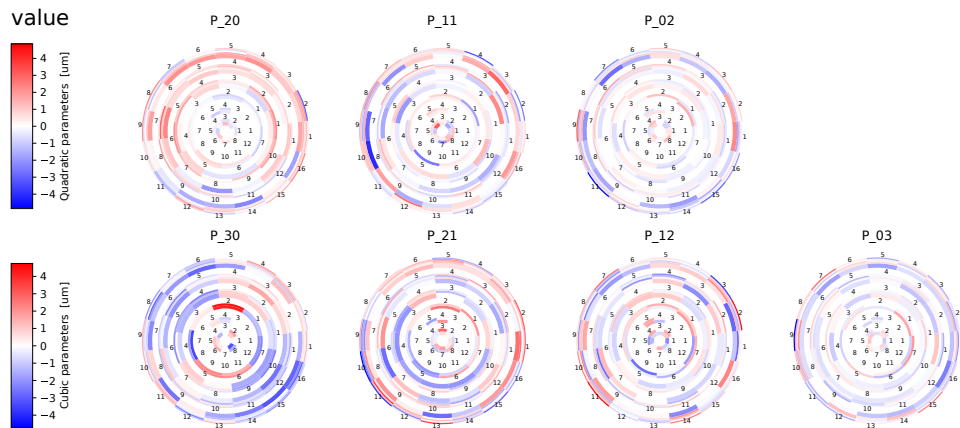
Figure 15.2: Residual misalignment for surface deformation parameters of the VXD sensors after MC test of a realistic misalignment imported from data.

hierarchy and aligned the VXD sensors with respect to the CDC (and each other). We noted that Millepede II handles the related correlations very well, and the result is almost identical (after transformation to the half-shell system) to that when hierarchy is used. However, the alignment hierarchy is useful in practice, as the real movements are truly mostly correlated and separation of these degrees of freedom allows for a frequent realignment of the half-shell positions with much less data. We show residual misalignment for (most of) the parameters in the traditional approach in Fig. 15.1 for rigid body sensor parameters, Fig. 15.2 for surface deformations, and Fig. 15.3
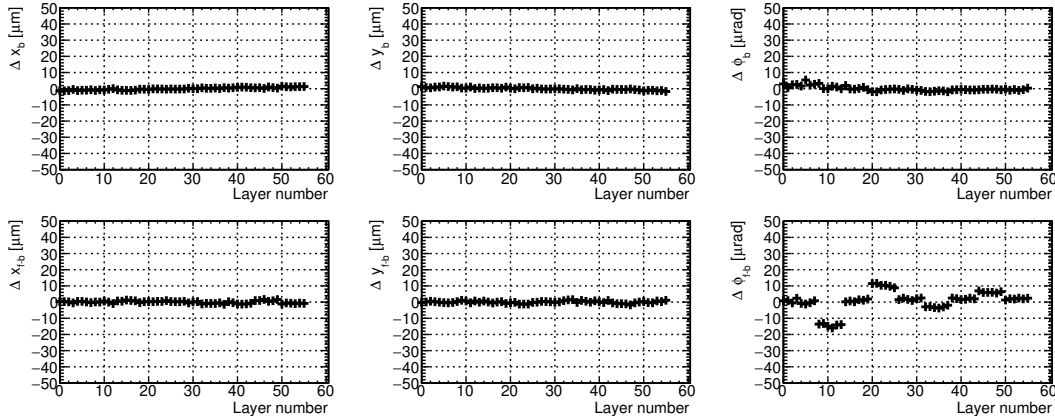
Figure 15.3: Residual misalignment for CDC layer parameters after MC test of a realistic misalignment imported from data.

for CDC layers. Parameters of the VXD half-shells will be discussed in detail in the next subsection.

2. **Test of weak mode misalignment as a telescope** (with scale $\sim 250\ \mu$m), performed around the same time and also not using VXD hierarchy. We transformed the payloads from (1) and (2) into the half-shell constraint system for the following comparisons. We have already used this residual misalignment for the discussion in Sec. 13.6 and shown some of its parameters in Figs. 13.6, 13.7, and 13.8.

3. **Test of the most recent prompt alignment**, starting from the realistic data misalignment, which is slightly artificially modified (ladder level alignment is removed, and the total z shift of the VXD half-shells is reduced). This and the following campaign already used the VXD (half-shell) hierarchy, as all data currently used for physics analyses. In this specific case, the alignment algorithm is not fully converged, as the final S, defined in Eq. 13.2 is close to 1.3. This means the automated algorithm would perform one additional iteration. However, this iteration is supposed to be not necessary; we only utilize it as a solid convergence cross-check. Thus we include this residual misalignment into our evaluation to check that this iteration is, in fact, not needed.

4. **Explicit test of a wide range of weak mode distortions** for VXD, discussed in detail in the next section. As these tests use exactly the same data as input, the variations of the residual misalignment are very small. Thus we select one particular example after alignment of the 'curl' deformation with an initial scale of $\sim 700\ \mu$m after the fourth alignment iteration.

These residual misalignments will be the basis for evaluating the baseline alignment precision. The size of the simulated data samples is roughly the same as that used for data. Therefore, we should resemble both statistical and systematic precision. The residual misalignment is a combination of statistical and systematic effects. However, systematic effects dominate, as can be seen, e.g., from the colorful patterns in the ring plots of the VXD sensors. Nevertheless,

the distributions of the parameters are often close to Gaussian, which permits a straightforward interpretation of the residual misalignment as an overall alignment precision. Naturally, the scale of any systematic effect cannot be larger than the displacement of any sensor. Thus one could simply use any maximal displacement as an upper bound on the scale of any systematic distortion. A more stable estimate would take, e.g., 3σ of the Gaussian distribution as this upper limit. Keeping this in mind, we will evaluate this σ for various parameter categories to give quantitative estimates of the alignment precision. In later sections, we will look at the systematic distortions directly. Here, we will simply produce histograms of the residual misalignment parameters after medium to very large misalignment on MC. Many systematic effects can be hidden in these histograms, but a straightforward quantification is possible.

15.2.1   Global Degrees of Freedom

Let us first concentrate on what we call systematic global biases. These are specific to our simultaneous alignment of the CDC and VXD, and we will see that these represent the dominant systematic effect of the entire problem. We briefly mentioned those in Sec. 13.6.

The major systematic bias in z is also a weak mode in the baseline alignment. It results in the final positions of the VXD half-shells being different from $z = 0$. We can see this effect in Fig. 15.4 (top right). Here we can see the true nature of a weak mode with a reasonable eigenvalue but at the low end of the spectrum. The control over the z-positions is definitely solid, as the initial misalignment was about 1 mm (0.5 mm) in z in the black (green) case. As we will see in Sec. 15.4.1, the range of the eigenvalues is quite large, spanning seven to eight orders of magnitude. This implies a possible loss of precision and a large sensitivity to input data. The CDC precision in z is poor (1–2 mm), and many tracks are necessary to determine the VXD z-position with good statistical accuracy. However, the actual minimum with the optimal VXD position with respect to the CDC[1] is shallow along the direction of this eigenvector in the high-dimensional objective function. This results in a large sensitivity to, e.g., data outliers or subtle mis-modeling effects. Indeed, for the green case, we realized that introducing a p-value requirement can reduce the bias in z from ~ 30 μm to the ~ 10 μm visible in the histogram. Since then, this requirement has been strictly imposed, see also Sec. 11.8.

The histograms in Fig. 15.4 should be interpreted with care. The estimates of the means μ represent the absolute precision of the positions (and rotations) of the VXD in space. It would be wrong to interpret them as a precision relative to the CDC. In fact, the VXD position follows the inner CDC layers. This is demonstrated in Fig. 15.6, where we have used studies without any initial misalignment but with different sample compositions. This figure shows the VXD and the CDC layers together, but only for the x and y parameters of the VXD half-shells and the x_b and y_b parameters for the CDC layers (see Fig. 12.8). The residual CDC misalignment shows quite a clear functional dependence on the

---

[1]This is coupled with a systematic pattern in the CDC layer alignment. Therefore it is rather a distortion of the CDC that causes the VXD to settle at non-zero absolute global z position.
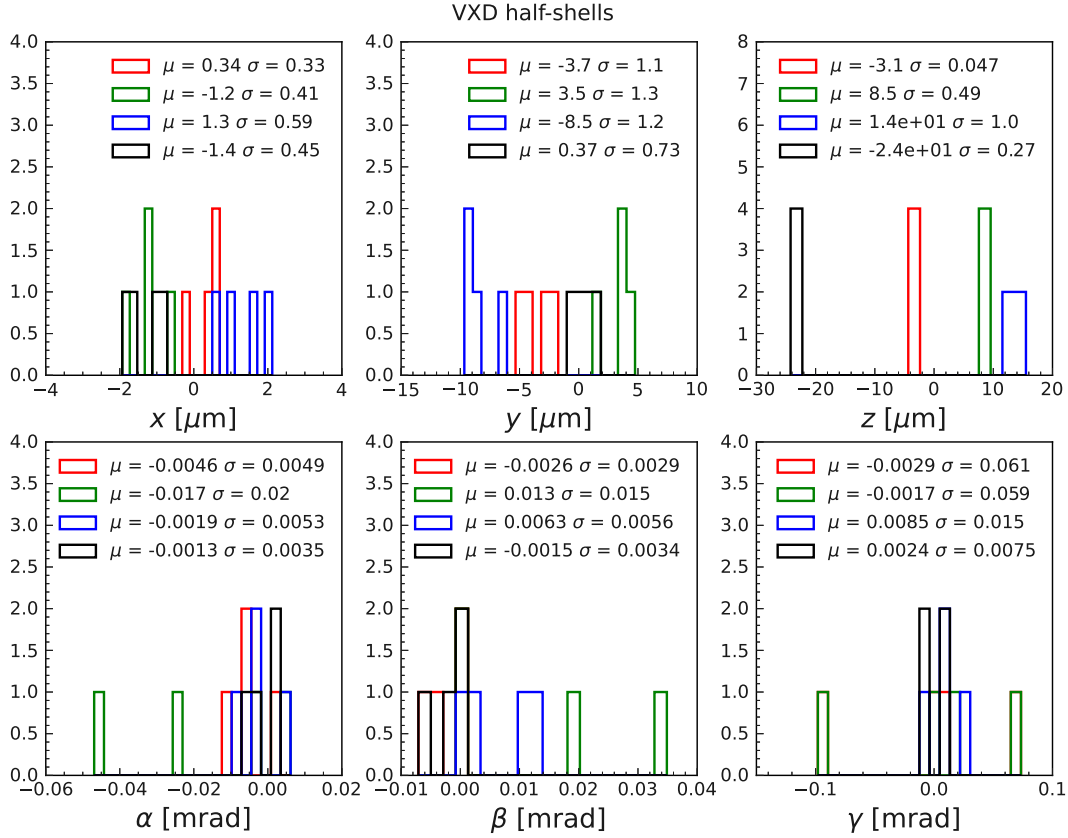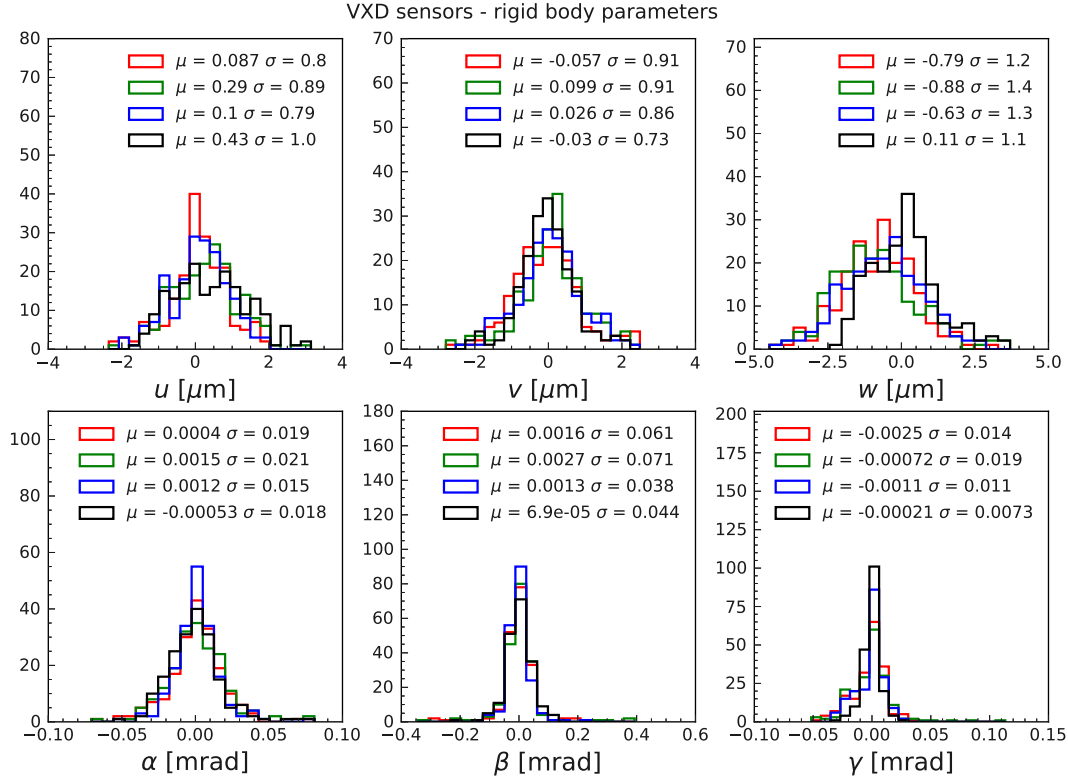
Figure 15.4: Histograms of residual misalignment for the VXD half-shells after different misalignments. Scenarios: (1) **realistic misalignment**, (2) **telescope misalignment**, (3) **realistic prompt misalignment**, (4) **weak mode misalignment**.

layer radius, and the VXD position is compatible with the residual misalignment induced to the inner layer—such that tracks are smooth at the interface of the sub-detectors. The worse precision in y is due to the absence of horizontal cosmic ray tracks, leaving room for a small sagitta distortion of the tracking system with off-centering of the inner CDC layers and the VXD as illustrated in Fig 15.5. Notable is the reduction of the systematic effects from the nominal case with cosmic and hadron samples (blue) achieved by the addition of high momentum single muons tracks (red), di-muons with an IP constraint and floating IP (green), and finally with fixed IP position, which almost removes the effect (black) in Fig. 15.6.

On the other hand, the spread of residual misalignment of the individual VXD half-shells is much smaller. In Fig. 15.4, this can be connected with the RMS, denoted σ, which can be interpreted as the alignment precision of the VXD half-shells relative to each other. This precision is roughly 1 μm.

The global CDC deformations thus cause the global offsets of VXD half-shells. As we can see from histograms in Fig. 15.7 with the residual misalignment of the CDC layer parameters, in the worst case, a layer can be misaligned by ~ 15 μm, which is remarkable over the entire CDC radius of 1.2 m. As the CDC layer constraints are used to define the reference system, the means of the distributions are trivially zeros within numerical precision. In general, the RMS σ, which indicates the alignment precision for the CDC layers, is significantly

Figure 15.5: Illustration of the off-centering in the y-direction (right) to which cosmic ray tracks (mostly vertical) are poorly sensitive. For the off-centering in the x-direction (left), this distortion creates a kink in the (merged) cosmic trajectory, which makes it resolvable by the alignment to much better precision. [81]
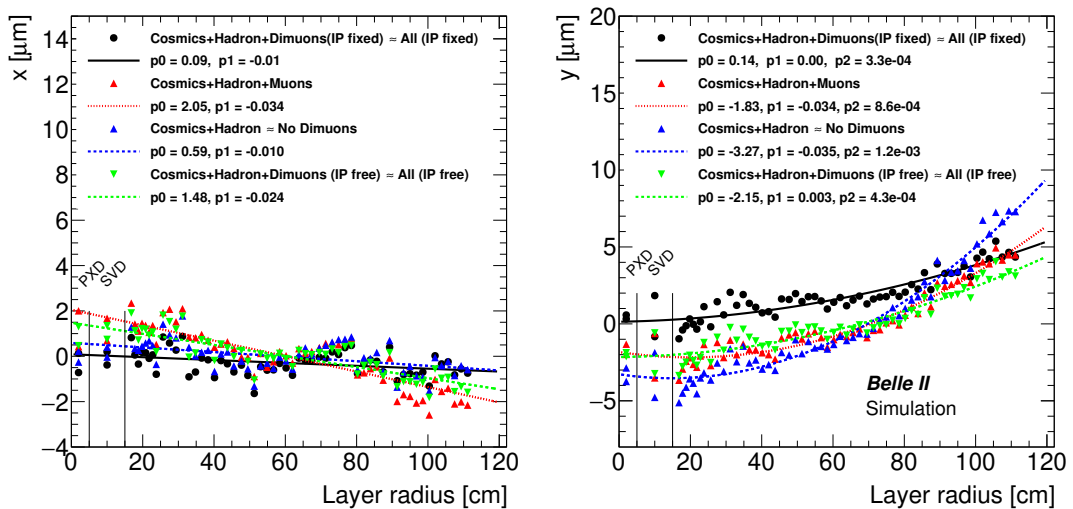
Figure 15.6: Residual misalignment after starting from the nominal geometry using different combinations of data samples in the x (left) and y (right) coordinate for the VXD half-shells and x_b and y_b parameters of the CDC layers. The VXD half-shell points are not used for the linear or quadratic fit. The color code is not related to other plots (histograms) in this section.

better than 10 μm. Also, the rotation parameters are well under control, with the largest deviations in the inner layers at smaller radii. When the rotation parameters are translated into the corresponding shifts of the wires, these are significantly smaller than 10 μm.

Let us now summarize the alignment precision for the global degrees of freedom. We follow a conservative philosophy and take the maximal observed deviation as estimates for μ and σ. Furthermore, we separate the PXD and SVD
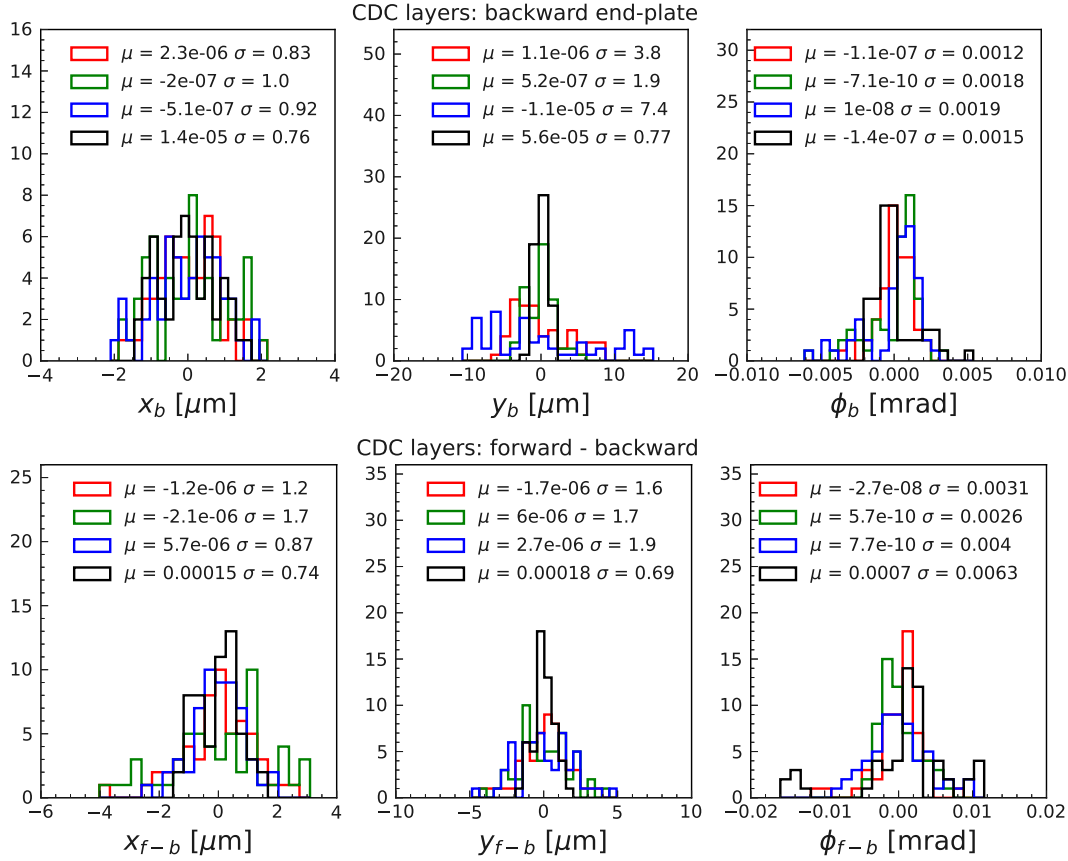
Figure 15.7: Histograms of residual misalignment for the CDC layers after different misalignments. Scenarios: (1) **realistic misalignment**, (2) **telescope misalignment**, (3) **realistic prompt misalignment**, (4) **weak mode misalignment**.

parameters. The summary of global offsets and relative precisions of all the parameters is shown in Table 15.1. Generally, the local relative precision of the half-shell alignment is 1 to 2 μm, naturally slightly better for the PXD. The global offsets are shared by the PXD and the SVD and also roughly correspond to the local precision for the CDC layers in the horizontal and vertical shifts. The vertical shift is less under control due to the absence of horizontal cosmic ray tracks. The most noticeable effect is the global z offset (bias), with the most significant observed value of about $(-)24$ μm identified in our MC studies. Such residual misalignments have no (or negligible) effect on physics performance.

15.2.2 Local Degrees of Freedom

Let us now turn our focus to the parameters of the individual VXD sensors. The histograms of residual misalignments in our MC studies are shown in Fig. 15.8 for rigid body parameters. Figs. 15.9 and 15.10 show the quadratic and cubic deformation parameters.

We can see that no single sensor is misaligned by more than 4.5 μm, and this maximum occurs for the w-coordinate, perpendicular to the sensor surface. Most problematic in this sense are the slanted outer SVD sensors. As all the

Sub-detector	parameter [unit]	global precision (max. bias μ)	local precision (max. σ)
PXD	x [μm]	1.3	0.77
	y [μm]	8.8	0.61
	z [μm]	24	0.37
	α [mrad]	0.035	0.012
	β [mrad]	0.027	0.0078
	γ [mrad]	0.013	0.086
SVD	x [μm]	1.8	0.31
	y [μm]	8.1	1.7
	z [μm]	24	1.2
	α [mrad]	0.0022	0.044
	β [mrad]	0.0012	0.00073
	γ [mrad]	0.0095	0.0081
CDC	x_b [μm]		1.0
	y_b [μm]		7.4
	ϕ_b [mrad]	——	0.0019
	x_{f-b} [μm]	(reference)	1.7
	y_{f-b} [μm]		1.9
	ϕ_{f-b} [mrad]		0.004

Table 15.1: Global and local alignment precision for the global degrees of freedom of PXD, SVD (half-shells), and CDC (layers). No numbers are shown for the global precision for CDC layers as CDC (as a rigid body) is fixed as an alignment reference by the constraints.

Figure 15.8: Histograms of residual misalignment for the VXD sensors' rigid body parameters after different misalignments. Scenarios: (1) **realistic misalignment**, (2) **telescope misalignment**, (3) **realistic prompt misalignment**, (4) **weak mode misalignment**.

distributions are close to Gaussian with a mean at zero, we evaluate the total precision as a sum of the average bias μ and RMS σ in quadrature.

In general, we can see the alignment precision is at the level of $1-2$ μm for all the parameters. For the rotations, the size of the sensors is important for precision, which can be seen in the summary Table 15.2, where we separate the parameters to PXD and SVD sensors. The larger sensors result in higher precision for the rotations α, β, and γ (and partially for surface deformations) for the SVD than the PXD, while the precision for most other parameters is comparable between the two sub-detectors.

From these histograms, we can barely extract some information about systematic distortions. As mentioned earlier, the scale of any such distortion cannot be larger than the maximum sensor misalignment[2]. One can roughly estimate the maximum as three times the alignment precision in Table 15.2. We will study these systematic effects in the next section and then connect them to proper weak modes of the solution.

---

[2]Technically, the scale can be twice as large, depending on the form of the deformation and the definition of its amplitude.

parameter [unit]	PXD precision	SVD precision
u [μm]	1.2	1.1
v [μm]	1.0	0.9
w [μm]	1.4	1.7
α [mrad]	0.047	0.018
β [mrad]	0.28	0.04
γ [mrad]	0.048	0.013
P_{20} [μm]	1.6	1.5
P_{11} [μm]	1.9	1.8
P_{02} [μm]	1.3	1.2
P_{30} [μm]	1.8	1.5
P_{21} [μm]	1.8	2.0
P_{12} [μm]	2.2	2.0
P_{03} [μm]	1.5	1.2

Table 15.2: Local alignment precision $\sqrt{\max(\mu)^2 + \max(\sigma)^2}$ for each type of local degree of freedom, separately for the PXD and the SVD.
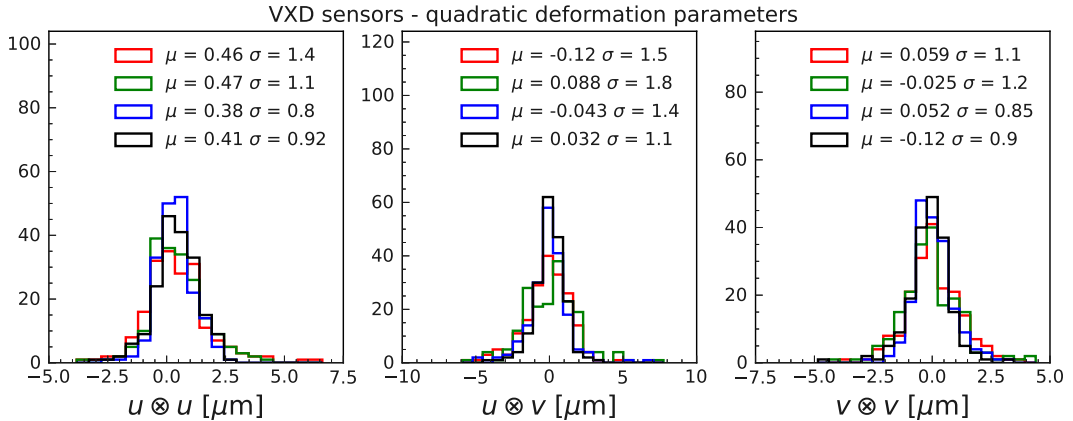
Figure 15.9: Histograms of residual misalignment for the VXD sensors' quadratic surface deformation parameters after different misalignments. Scenarios: (1) **realistic misalignment**, (2) **telescope misalignment**, (3) **realistic prompt misalignment**, (4) **weak mode misalignment**.

15.3 Sensitivity to Systematic Distortions

For the VXD, we explicitly test the alignment of the classical nine weak-mode distortions introduced in Sec. 13.6.3. We generally follow the parametrization of BaBar [70], but for example, 'bowing' can have a linear, absolute value-like, or quadratic shape. Previously we have used parametrization with an absolute value. Also, for example, the elliptical deformation scale was proportional to the sensor radius. Although we have also tested these variants, simplified parameterizations
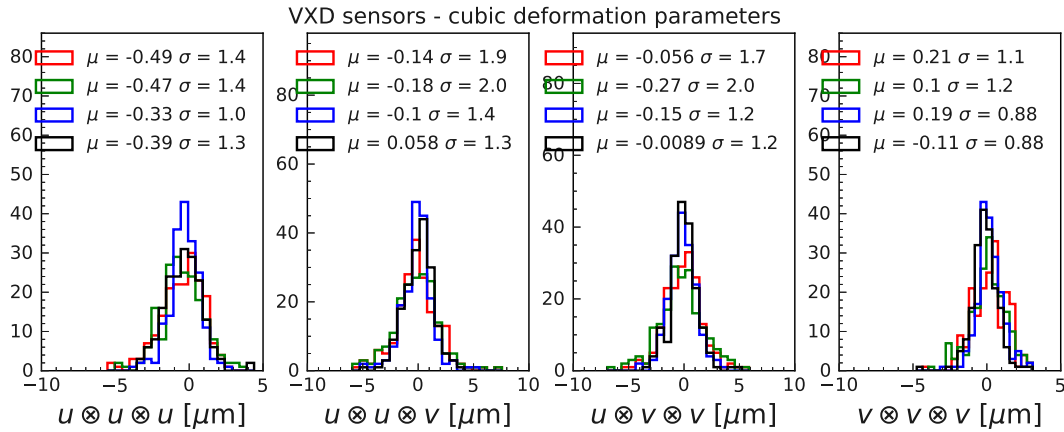
Figure 15.10: Histograms of residual misalignment for the VXD sensors' cubic surface deformation parameters after different misalignments. Scenarios: (1) **realistic misalignment**, (2) **telescope misalignment**, (3) **realistic prompt misalignment**, (4) **weak mode misalignment**.

are sometimes used for the exposition in this section to allow for a simple scale fit. For most of the distortions, we have also tried large scales around 500 μm. There is generally no difference in the result apart from a higher number of iterations needed. Following the BaBar vertex detector alignment paper [70], we present here results from initial deformations usually starting with a scale of ~ 50 μm, except for the 'curl' deformation (as an example of a large initial misalignment).

The result of the alignment of these nine distortions is shown in Fig. 15.11. This new way of visualizing the residual misalignment allows us to recognize and categorize the distortions. On the horizontal axis of each sub-plot, the sensor radius, azimuthal angle, or the z-coordinate is plotted. The residual misalignment is shown on the vertical axes, separately for each of the three translational degrees of freedom. Global movements are absorbed into half-shell parameters and are not visible in these projections. For example, the global z bias would look like a significant offset in all plots showing v (Δz). In each subfigure, the name of the weak mode, the fit results to the amplitude of the deformations of the initial misalignment $A(0)$, and in the n-th alignment iteration $A(n)$ are shown, together with an error estimate from a simple χ^2 fit. The fit functions are chosen (sometimes together with the misalignment parametrization) such that the initial deformation is well described and a scaling factor is introduced, such that its amplitude encodes roughly the displacement of the most misaligned sensor. We use linear, first-order polynomial, or trigonometric functions to fit the initial and remaining deformation. As the final deformations generally do not correspond to these simple functions, the fit of the final amplitude may only estimate the systematic local effects to first order. The initial misalignment is, in all cases, completely removed, although from 2 to 3 iterations may be needed (4 for the large 'curl' misalignment). The remaining patterns have very small amplitudes, with the most noticeable distortion being the z-expansion (z-contraction in this case) with a scale of 1.4 μm. The second most prominent effects are related to deformations which are functions of the azimuthal angle, presumably due to the asymmetry of the distribution of cosmic ray tracks.
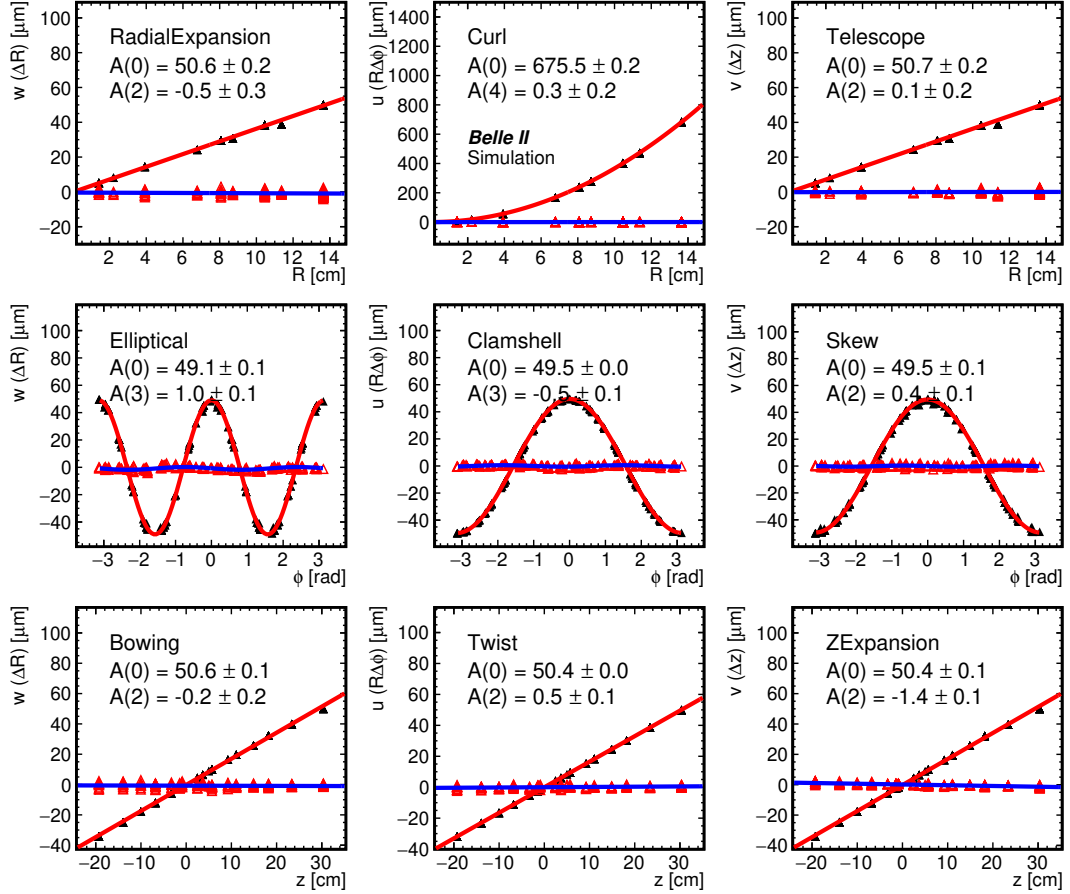
Figure 15.11: MC tests of baseline alignment of the nine classical weak mode distortions, showing initial and residual misalignment projected into cylindrical coordinates. The initial misalignment (black points) is fitted with a function (red curve) to extract the scale of the deformation, corresponding to the most misaligned sensor. The final misalignment after the n-th iteration is shown (red points) together with the corresponding fit result to the scale (blue curve).

In Fig. 15.11, each subfigure shows the residual misalignment of that particular misalignment study. However, we have observed that the difference among the final misalignments is almost entirely independent of the initial misalignment. We have selected cases (1) and (3) from the misalignment studies to show more detailed examples of the residual misalignment in this view. This time, each subfigure shows the same misalignment, only in a different projection. The cases are shown in Fig. 15.12 and Fig. 15.13. In this case, we do not show the (very large) initial misalignment, only the final fit to the amplitude. We have chosen quadratic instead of linear function for 'bowing' as suggested by the residual pattern.

The small z-expansion is clearly visible in both cases. Also, some oscillatory patterns are present in the second row, and elliptical distortion is usually the most prominent. In all cases, the extracted amplitude is 2 μm or less, but, for example, quite a large 4 μm error for the elliptical deformation amplitude remains visible in Fig. 15.13. As 4 μm is also roughly the maximal sensor displacement, we prefer to refer to this number as a maximal scale of any deformation—even
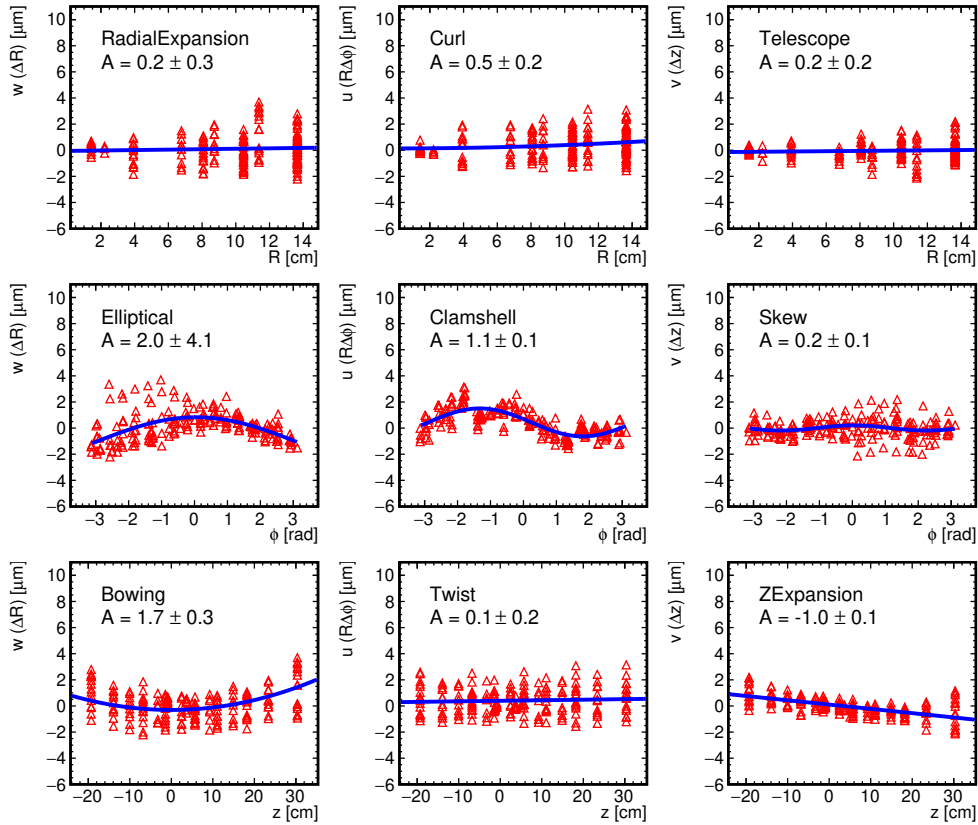
Figure 15.12: Residual misalignment from a study starting from **realistic misalignment** (1) extracted from the data, projected into cylindrical coordinates. This alignment configuration was not using the VXD hierarchy. We transformed the residual misalignment into the half-shell coordinate system for these plots.

those which average over the detector. For example, ϕ-dependent telescope or radial expansion would average to zero in the corresponding subfigures in the top row.

15.4 Weak Modes

In other parts of this thesis, we use the term weak mode quite freely. Also, in the previous section's explicit test of weak mode distortions for the VXD, we used the classical set of nine canonical weak modes. In reality, these do not really correspond to the true eigenvectors of our solution, which are significantly more complicated. This section will explicitly study the eigenvectors and associated weak modes and attempt to partially categorize them and connect them to the already observed residual systematic effects, like the systematic global offsets.

15.4.1 Spectra of Eigenvalues

The about three thousand parameters of the baseline alignment allow for a fast solution with the diagonalization method. By diagonalization, the symmetric global matrix $\boldsymbol{C}$ can be expressed as $\boldsymbol{C} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^{\mathsf{T}}$, where $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots \lambda_n)$ is a diagonal matrix and $\boldsymbol{U} = (\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n)$ is formed from columns of vectors
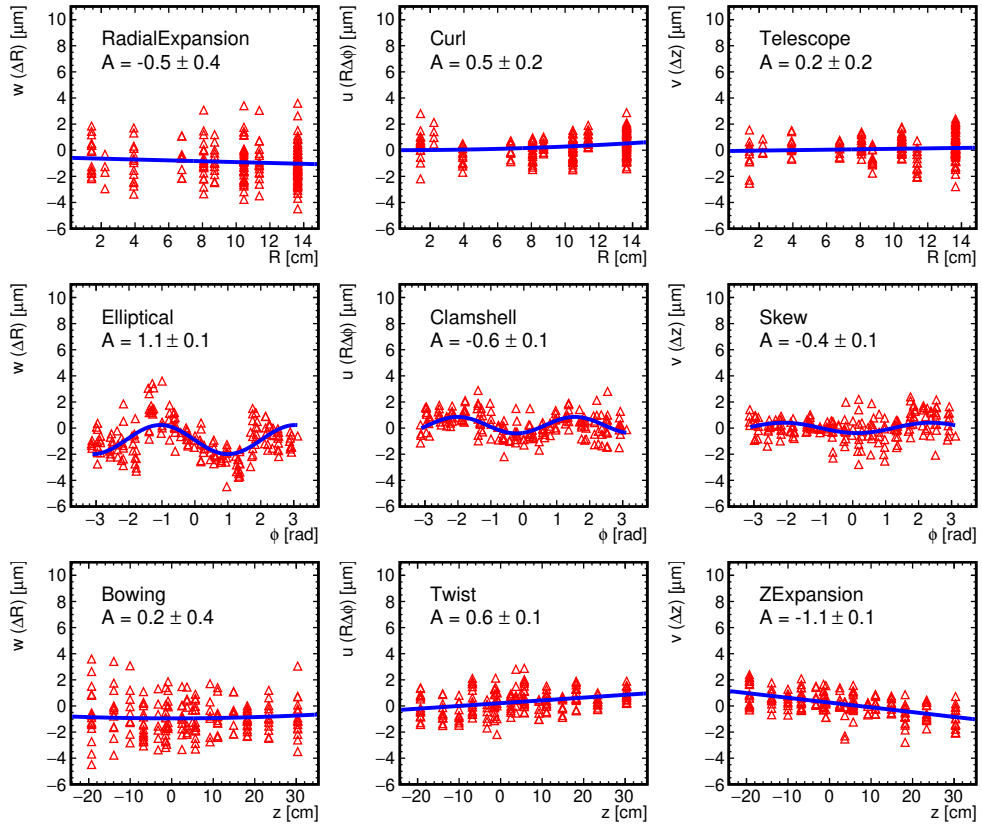
Figure 15.13: Residual misalignment from after a study starting from **realistic prompt misalignment** (3) extracted from the data, projected into cylindrical coordinates.

satisfying $\boldsymbol{U}\boldsymbol{U}^{\mathsf{T}} = \boldsymbol{U}^{\mathsf{T}}\boldsymbol{U} = \mathbf{1}$. The numbers λ_i and the associated orthogonal vectors $\boldsymbol{u}_i$ correspond to the eigenvalues and eigenvectors of the matrix $\boldsymbol{C}$, respectively.

Using the eigenvalues and eigenvectors, the solution of the alignment problem $\boldsymbol{C}\Delta\boldsymbol{a} = \boldsymbol{g}$ can be written as [116]

$$\Delta\boldsymbol{a} = \sum_{i=1}^{n} \frac{1}{\lambda_i}(\boldsymbol{u}_i \cdot \boldsymbol{g})\,\boldsymbol{u}_i, \tag{15.1}$$

where each eigenvector in the sum is multiplied by its projection onto the vector $\boldsymbol{g}$ (formed from the gradient of the objective function; after matrix reduction in Millepede II) and by the inverse of the corresponding eigenvalue. Eigenvalues close to zero prevent matrix inversion. Very low eigenvalues cause a large sensitivity to input data (residuals), as the gradient of the objective function along the corresponding direction (eigenvector) is shallow. This means that even a large change of the alignment parameters in the direction of such an eigenvector has a small impact on the total χ^2 value. Or, in the opposite direction, a small change in the scalar product $\boldsymbol{u}_i \cdot \boldsymbol{g}$ is multiplied by a large number $1/\lambda_i$ and thus can result in a large change of the estimated parameter values in $\Delta\boldsymbol{a}$.

Suppose the largest eigenvalue $\lambda_{\max}$ is significantly larger than the smallest one $\lambda_{\min}$. In that case, a loss in numerical precision is possible, where the movement along some weak-mode direction might happen only at the very last couple of
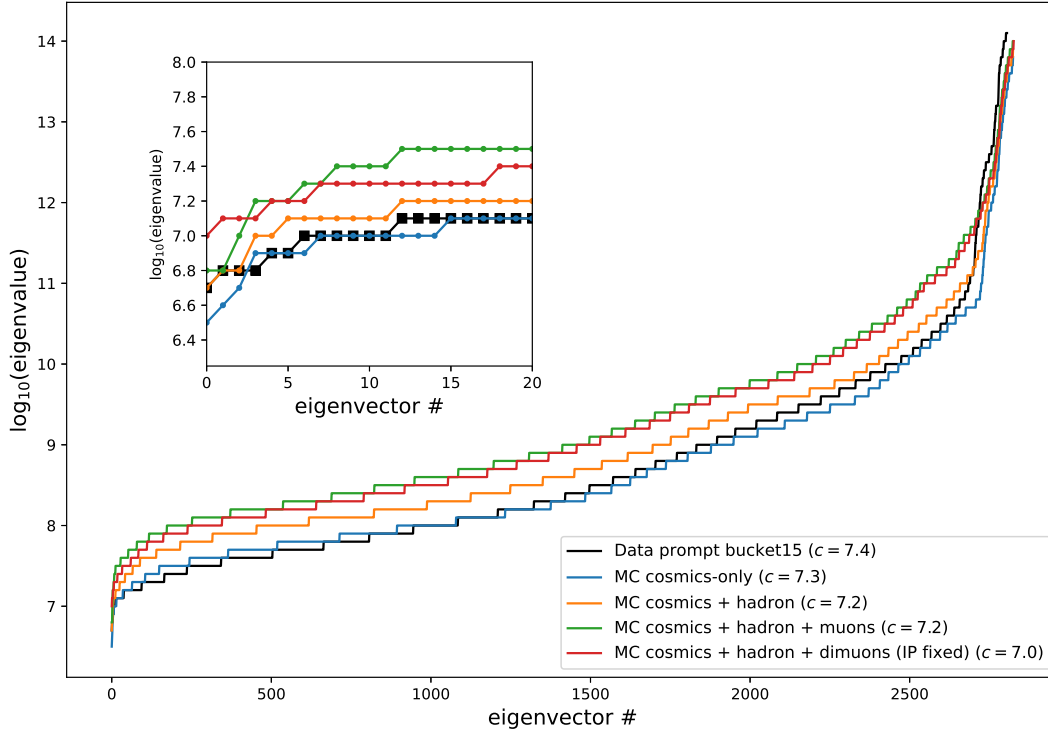
Figure 15.14: Spectra of eigenvalues, ordered from smallest to largest one for several variants of sample composition in MC and for data of prompt `bucket 15`. The inset shows in detail the low end of the spectrum. The condition number c for each of the solutions' global matrix is shown in the legend.

digits of the objective function or even below the numerical resolution. One can define a condition number for the matrix $\mathbf{C}$ as

$$c = \log_{10} \frac{\lambda_{\max}}{\lambda_{\min}}, \tag{15.2}$$

where we already introduce logarithmic scaling because, in the alignment problems, the eigenvalues typically span several orders of magnitude. We show example spectra of the eigenvalues for the baseline alignment in MC and data in Fig. 15.14. First, no zero or vanishing eigenvalues can be seen. The smallest eigenvalue is larger than 10^6. On the other hand, the condition numbers are near $c \sim 7$, i.e., the largest eigenvalue is about ten million times larger than the smallest one. We compare several MC variants with different data samples and one example from real data of `bucket 15` (see Sec. 13.3.2), taken during the last two weeks before the 2020 Summer shutdown.

In the inset of Fig. 15.14, we also show in detail the low end of the spectrum with the twenty lowest eigenvalues. As expected, the blue case with cosmic-only data has the lowest eigenvalue and worst c among the MC spectra. Addition of the hadron sample (orange) and single muons (green) further flattened the spectrum and decreased the condition number. The best condition number and the flattest spectrum are naturally achieved by fixing the global offsets with IP-constrained di-muons. However, we do not normalize the spectra in any way, so an exact 1:1 comparison is not possible. The number of parameters for data is slightly smaller (e.g., due to one inactive CDC layer). In general, the spectrum

for data seems the least flat and more similar to the worse cases in MC, with the largest condition number $c = 7.4$. On the other hand, the very low end of the spectrum is subjectively more flat for data than for most of the cases with MC simulation. Note that the data example includes off-IP tracks, single muons, and cosmic ray tracks and has a slightly different sample size and composition than the MC alignments.

We can conclude that the MC studies faithfully reproduce essential characteristics, like the condition number of the global matrix. Furthermore, no truly weak χ^2-invariant modes are present, and the lowest-lying ones should still be under reasonable control given machine numerical precision[3]. However, these eigenvectors can be suspected to be the most sensitive to variations in the data input and correspond to some of the previously observed systematic effects. The similarity of data and MC will be further supported by investigating the individual eigenvectors corresponding to the low end of the spectra in the following sub-section.

15.4.2 Eigenvectors and true Weak Modes

Visually pleasing and sufficiently compact visualization of the entire vector of global parameters is difficult to obtain (especially for CDC layer parameters and in particular after wires are added). Multiple different ways to look at the global parameters are used in this thesis with their advantages and limitations.

The most compact way is to draw the parameters ordered by a continuous index. This is used in Fig. 15.15, where seven eigenvectors with the smallest eigenvalues for one alignment on MC (cosmics + hadron corresponding to the orange spectrum in Fig. 15.14) and one on data (`bucket 15`—black spectrum in Fig. 15.14) are shown. On the vertical axis is the weight of each parameter in the eigenvector, between -1 and $+1$. The vertical dashed lines visually separate the VXD layers, and the last block of parameters on the right are the parameters of the CDC layers. The parameters are further separated by type using different markers. The (sometimes detached) points at the first and third dashed lines correspond to the PXD and SVD half-shells (i.e., the global degrees of freedom for the VXD).

Let us discuss qualitatively several of the eigenvectors in detail, in particular, the first three in MC, trying to identify them also in data. The most recurrent eigenvector (first on both data and MC) corresponds to a known global z bias, as we can see from the large weight for the z parameters of VXD half-shells. Interestingly, the sign of this eigenvector is the opposite in MC and data. For the green spectrum in Fig. 15.14 (cosmics + hadron + muons), this eigenvector is, in fact, only the second weakest, switched with the second one on the left in Fig. 15.15. This z bias weak mode is, however, the most interesting one. Also, it is the most often observed lowest-lying eigenvector. Moreover, it seems to almost correspond to a misalignment deformation of the real detector, later removed by the reprocessing.

From the figure of the global vector, we can only hardly see the pattern in the weights of parameters other than the z positions of the VXD half-shells.

---

[3]Note, however, that single precision (float) would be at resolution limits and instead double precision should be used for all calculations.
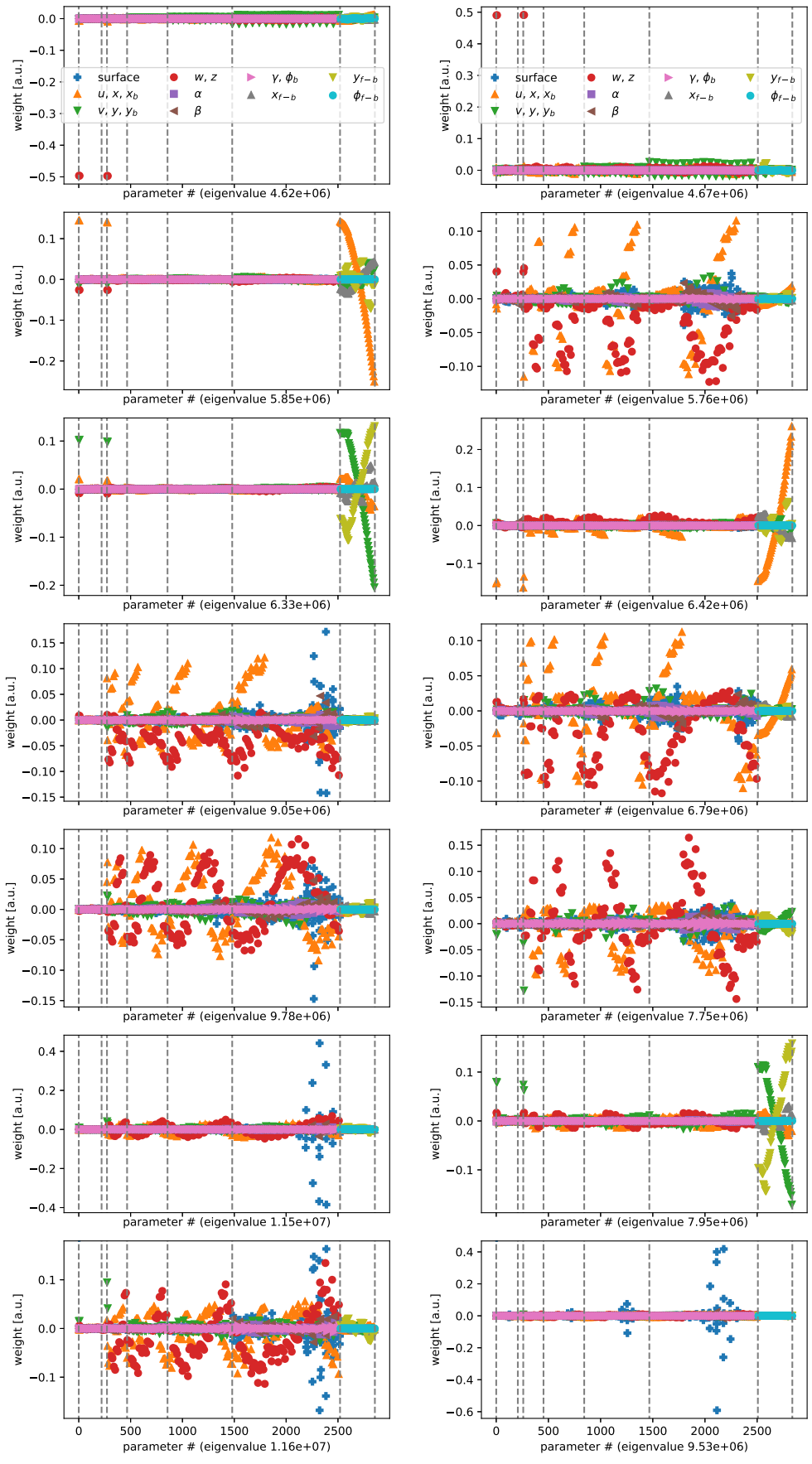
Figure 15.15: Parameter weights of seven eigenvectors with lowest eigenvalues for MC simulation (cosmics + hadron) (left) and real data of `bucket 15` (right).
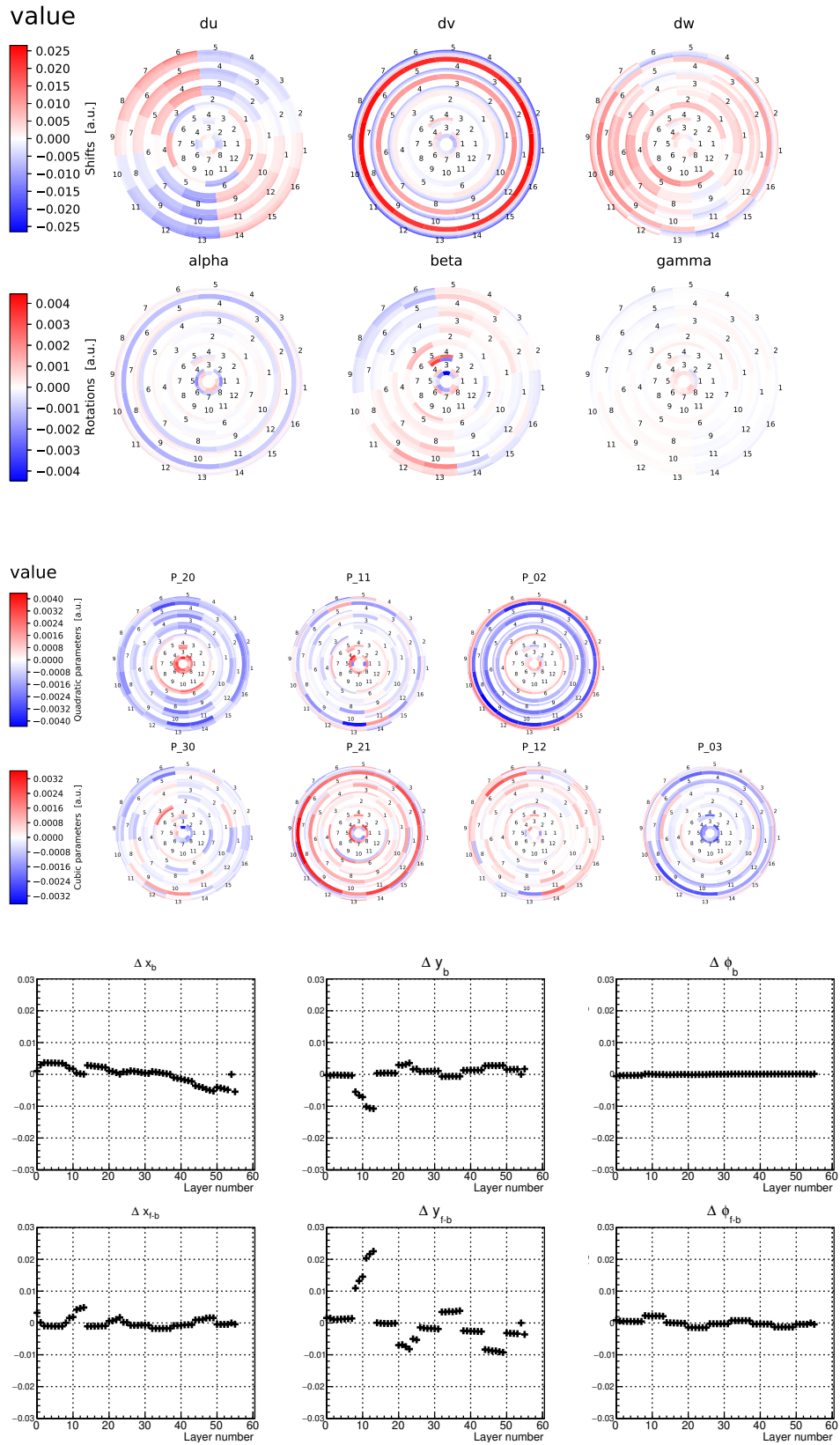
Figure 15.16: Eigenvector of the global z offset with the smallest eigenvalue on data for VXD sensors' and CDC layers' parameters.
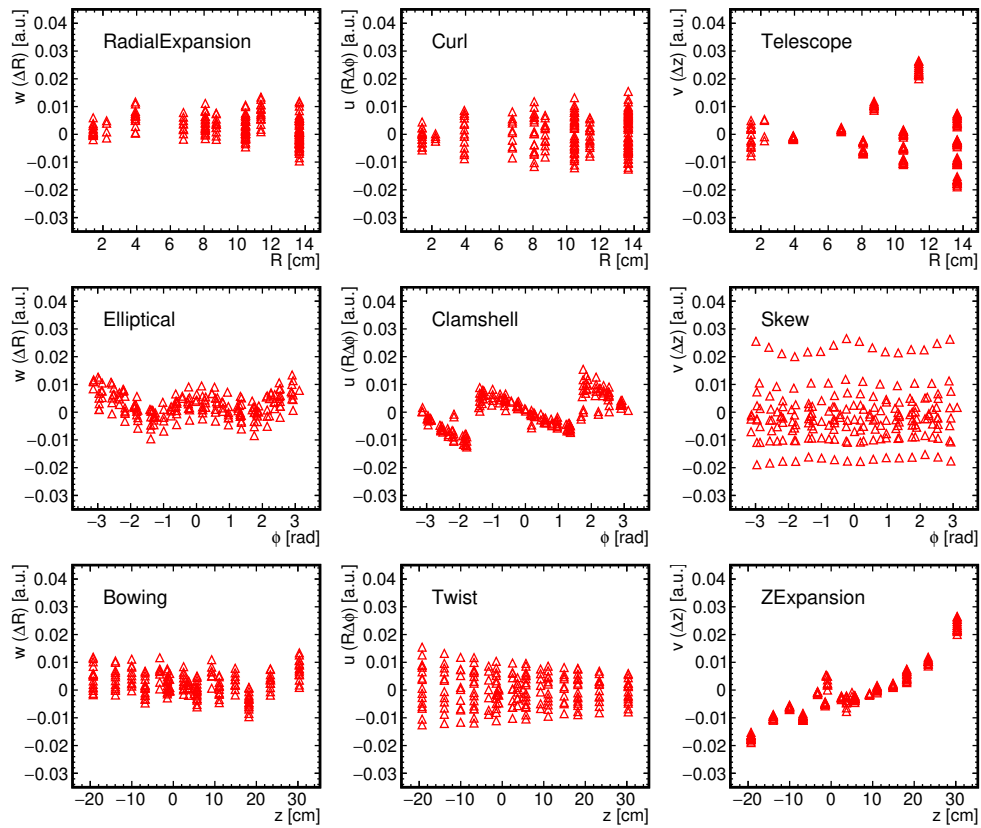
Figure 15.17: The smallest eigenvector of the global z-offset (in data) for the VXD sensor shifts projected into cylindrical coordinates. The vertical axis shows parameter weight.

Thus we also visualize the VXD sensor and CDC layer parameters traditionally in Fig. 15.16. The most prominent pattern is the z-expansion in the outer SVD layers (the v coordinate). In addition, a clear pattern is also visible in the u-coordinate. These patterns can possibly be identified in the prompt misalignment model (see Fig. 13.12). Also, the angles and surface parameters show some patterns, but their weight is generally one order of magnitude smaller than for the z-expansion. Thus we see that the global z offset can be coupled with z-scaling effects. In Fig. 15.17, this weakest eigenvector is also shown for the VXD shifts as a function of the cylindrical coordinates of the sensors. This time the vertical axis represents the weight of the parameters in the eigenvector. We can see clear patterns in the elliptical, clamshell, and z-expansion subfigures. The pattern in the skew subfigure is connected with the z-expansion, where the most significant weight is on the most forward outer slanted SVD sensors, roughly constant as a function of ϕ.

For what concerns the CDC layer parameters in the weakest mode, the patterns shown in Fig. 15.16 (bottom) are complex, and no straightforward interpretation of how the coupling among these degrees of freedom works has been found.

Regarding other weak modes, we can roughly identify the two remaining global offsets by the large weights of the CDC layer alignment parameters. Namely the second mode for MC (left in Fig. 15.15) corresponds to the global CDC deformation in x. Interestingly, this mode has a lower eigenvalue than the y-deformation, which can be at least partially identified in the third MC mode. However, in this case, the y_b parameters are also coupled with the y_{f-b} parameters, next to the y-shift of the VXD. This is slightly surprising as, from the misalignment studies, one would naturally expect the y-deformation to be more problematic. Also, in data, these global modes are only third and sixth weakest, interleaved by weak modes mostly localized in the VXD. Possibly the spectrum for data is so flat at the low end that the order of weakest modes can be interchanged.

The VXD parameters show some interesting oscillatory patterns among the eigenvectors. Mostly the u and w coordinates are affected, followed by v, and quite surprisingly also, some deformation parameters in the outer SVD layers can have a large weight and contribute to these specific distortions of the VXD. A simple interpretation by means of any of the classical weak modes is not possible. Instead, we always see that several parameter categories are coupled. Sometimes we also see modes where local sensor parameter misalignment is coupled with the global degrees of freedom, e.g., for the fourth eigenvector for the data spectrum in Fig. 15.15 (right).

For completeness, let us also visualize the weak modes from real data explicitly for the VXD sensors and their shifts (weights of rotations are much smaller in the eigenvectors) using ring plots in Fig. 15.18. The lowest-lying mode coupled with the global z bias is not shown here (see Fig. 15.16 top row). The following four modes are very different from the weakest one: practically only u and w parameters are affected. Notable is the similarity of the last two shown eigenvectors. This might look strange as the eigenvectors are expected to be orthogonal. However, by looking back into Fig. 15.15 (fifth and sixth row, right), we can see these eigenvectors affect differently local (VXD sensors) and global degrees of freedom (CDC layers and VXD half-shells).
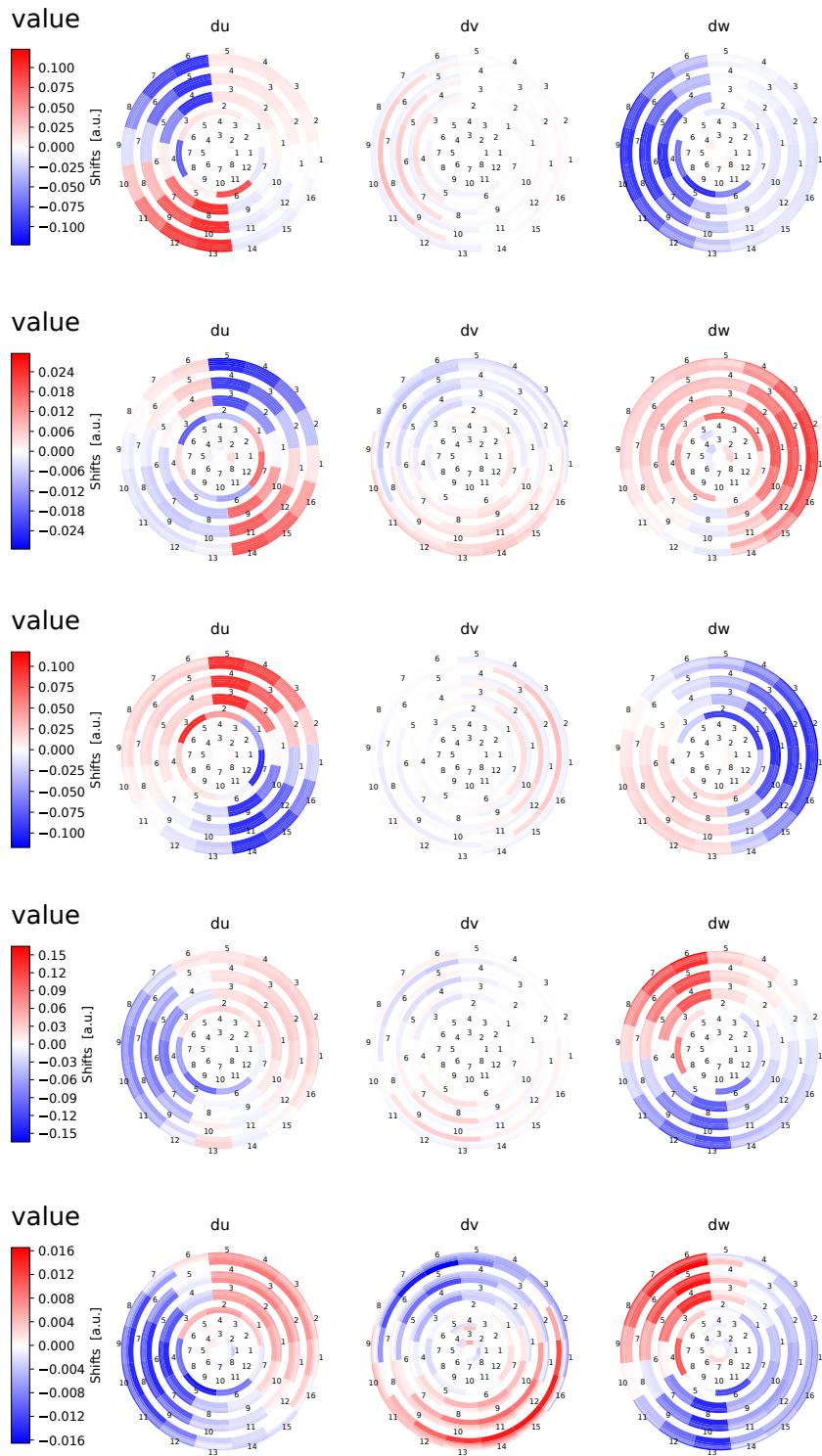
Figure 15.18: The eigenvectors from second to sixth (first is shown in Fig. 15.16) for data of prompt `bucket 15` alignment ordered (from top to bottom) starting with the lowest eigenvalue for the VXD sensor shifts. The color axis shows parameter weight.

15.5 Summary

In this chapter, we have started with directly exercising our baseline alignment procedure on the entirely realistic misalignment of the VXD and CDC. We have demonstrated that the baseline alignment can recover from a wide range of misalignments, including a number of systematic distortions for the VXD. Variations in the different studies allow the assessment of the alignment capabilities and reduce the risk of a good result achieved by accident and irreproducible.

Following a conservative approach, we quote the worst observed results as our precision estimates. Nevertheless, the results would be excellent if such precision could be really achieved on data. In general, the alignment precision is 1–2 μm for the VXD sensor parameters, including parameters of surface deformations and rotations (when translated into shifts of sensor corners). Systematic distortions in the VXD are limited to less than ~ 4 μm in amplitude, with z-expansion being the most prominent systematics at an average scale around 1.4 μm considering a simple parametrization. For the CDC layers, the precision is from $1 - 7$ μm, dominated by a residual global misalignment, which results in a limited systematic precision on the absolute position of the VXD half-shells. These are ~ 1 μm in the x-coordinate, ~ 8 μm in the y-coordinate, and ~ 24 μm in the z-coordinate (where the CDC resolution is worst).

We have also directly identified the eigenvector connected with the largest observed systematic effect of the global z bias. This usually indeed corresponds to the lowest eigenvalue. The systematic study of the weak modes and spectrum of eigenvalues is for completeness; some aspects are not yet fully understood. One cannot, for example, identify the VXD parameter patterns among MC and data eigenvectors. The CDC layer weights are roughly similar, at least in the three main global distortions, but for the VXD parameters, at least the phase of parameter oscillations seems systematically shifted. The additional tracks (like off-IP or wider phase space of available cosmic ray tracks) in data likely result in such differences. Anyway, we have built our confidence in the alignment procedure by direct tests and later by data validation.

However, the weakest mode is probably present significantly in the prompt baseline alignment of the real data. As we have no indication for its significant presence in the MC studies, this was probably caused by other inconsistencies, which distort the minimum of the objective function. One should not rely on simulations entirely. For alignment, validation with real data is absolutely essential. We devote to this validation Chapter 17, where we will identify some problems. To mitigate most of them, the reprocessing alignment was developed, which also includes IP-constrained di-muons and wire-by-wire alignment of the CDC[4]. This work is discussed in the next chapter. To save some space, the validation of the reprocessing alignment is compared to the prompt baseline alignment simultaneously. The alignment precision becomes comparable in data and MC studies only after this advanced alignment.

---

[4]In addition, run-dependent alignment is applied for a limited number of parameters.

Studies of full Alignment with Wires

In this chapter, the alignment problem will be scaled up by order of magnitude to include the geometrical alignment of CDC wires. Without any additional optimizations, this would correspond to two orders of magnitude higher memory requirements and an increase in the required raw computing time by three orders of magnitude. The study presented here only became possible with the recent developments in Millepede II, allowing us to solve this large alignment problem exactly in a reasonable time. We will start with the formulation of the problem and basic optimizations on MC. Once satisfied, we will also consider real data and optimize the alignment based on validation results.

Before Summer 2020, the only solution methods for $\sim 60 \times 10^3$ parameters with a reasonable runtime were approximate. Solution by inversion took over half a day. This could be acceptable for a one-time alignment for data but not for extended systematic studies on simulations. All wire alignment attempts previously used the approximate MINRES-QLP method; see Sec. 11.1.2. In fact, this solution method was used to derive the wire alignment used for data reconstruction until 2021 reprocessing. At that time, relatively fresh data taken without a magnetic field with a proper CDC calibration were used together with standard cosmic data taken with a magnetic field and collision tracks. Although the alignment has performed well in data validations, strong support for its validity from MC studies was missing. In addition, a possible degradation of the alignment was suspected (see Sec. 14.5), and thus, a method for a regular re-alignment of the wires was desirable.

The MC studies to evaluate the full wire alignment revealed issues with the MINRES-QLP method. This was finally confirmed after its solution was compared to a result achieved by inversion. Fortunately, in Summer 2020, a new, faster exact solution method—Cholesky decomposition—became available. With solution times decreased to about three hours, it became feasible not only to use the method for data but also for a set of systematic studies of the procedure on MC. Finally, since Spring 2021, the solution time could be decreased to about 20 minutes[1]. This was achieved by using external highly-optimized implementations [84] of linear algebra libraries LAPACK [83]. While the results presented in this chapter were derived before this feature became available, the 2021 data

---

[1]The tests were using ten cores on Intel(R) Xeon(R) CPU E5-2640 v3 @ 2.60GHz. The required memory (mainly) to store the global matrix is 20 GB for the full problem with wires. Baseline alignment needs less than 2 GB of memory and a couple of minutes to finish the algorithm step. The total alignment times are dominated by the data collection, submitted as batch jobs to a computing cluster.

reprocessing was already using this fast solution method.

Some results of this chapter were presented at the 25th International Conference on Computing in High-Energy and Nuclear Physics (CHEP2021) and published as proceedings [117].

16.1 Alignment Setup

One of the motivations for studies in this chapter is the need to realign the CDC without cosmic ray tracks recorded when the magnetic field is turned off. This data can only be collected when the Belle II solenoid is not excited at the start and end of the data-taking period. Considering how large deformations (up to about 100 μm) can be induced when the superconducting optics' magnetic fields and compensating magnets are ramping up and down, one expects significant deformations to occur also when the main 1.5 T solenoid is (de)excited. Thus data without and with magnetic field are likely not compatible, and it is not known if the associated deformations could be entirely absorbed in larger structures like the CDC layers. In addition, data without the magnetic field requires a dedicated CDC calibration that is not performed regularly.

To be able to perform the full realignment frequently, we must resort to more abundant data samples recorded during luminosity runs. Three types of such data samples are considered, together with their MC equivalents:

- Cosmic ray tracks with both arms reconstructed as a single trajectory, recorded during collisions in the dedicated cosmic skim.

- Di-muon $e^+e^- \to \mu^+\mu^-$ events from a dedicated skim with tight selection criteria.

- All charged tracks from hadron skim, which is mainly populated by multi-track hadronic events with a soft momentum spectrum. As an MC equivalent of this sample, generic $B\overline{B}$ events are used. In the real production of alignment constants, this sample is also mixed with tracks coming far from the interaction point, such as protons from beam-gas events.

The initial composition of the MC data samples is based on preliminary studies and an educated guess: 5×10^5 cosmic events, 10^6 di-muon events, and 2×10^6 tracks from generic $B\overline{B}$ events. Further optimizations based on MC studies are discussed in the next section.

The floating alignment degrees of freedom are configured as follows:

- All VXD half-shell and sensor parameters, including sensor surface deformations up to third order, are floating. Corresponding 6×4 constraints for sensor movements inside the half-shells are included automatically (see Sec. 12.6.1).

- The IP position is used as a fixed constraint by default unless mentioned explicitly otherwise.

- All CDC layer parameters are floating, with the default layer constraints: Δx_b, Δy_b, $\Delta \phi_b$, Δx_{f-b}, Δy_{f-b}, z-offset constraint and twist constraint as defined in Sec. 12.6.2 (7 constraints), unless mentioned otherwise.

- All wire parameters are floating with the default constraints fixing average shifts and rotation of wires in each layer at each end-plate (6×56 constraints), see Sec. 12.6.3 unless mentioned otherwise.

The total number of these default constraints is 367. In the next sections, additional constraints will be derived and tested. The total number of variable alignment parameters is 60,192. The resulting number of floating parameters (global matrix size) is 59,825 for this set of constraints (367). The solution is performed by Cholesky decomposition.

16.2 Studies without Misalignment

It is convenient to start the initial studies without introducing any misalignment. In the ideal case, running the procedure with a perfectly aligned detector should result in all parameter estimates consistent with zero. This is usually not the case for realistic alignment configurations, and one should study if a larger data sample, additional track topologies, more constraints, or external measurements could improve the result.

Extensive studies of the baseline alignment without wires showed that very good alignment could be achieved just by a combination of cosmic ray tracks (in a magnetic field) and simple tracks from e^+e^- collisions. Preliminary results of the full alignment with wires indicated that IP-constrained di-muons are a necessary additional ingredient to reach reasonable results. Thus we start by combining the three primary data samples. After the weak modes of the wire alignment will be under control, further optimization of the data sample composition will follow.

16.2.1 Impact of radial Constraints and Surveys

The weak modes affecting radial scale are a typical alignment problem. For a wire chamber like CDC, there is no redundancy analogous to geometrical sensor overlaps in the VXD. Thus the radius of the chamber must be fixed by external input. A good solution would constitute using mass-constrained decays of heavy unstable particles with precisely known mass to fix the absolute mass scale, resulting in constraints on the radial scale, as done at the CMS [74]. A change in the CDC radius would result in a systematic change of momenta of charged tracks, shifting reconstructed invariant masses of the decaying particles.

An obvious candidate is $J/\psi \to \mu^+\mu^-$ decays, with a narrow width (about 90 keV, although smeared by detector resolution). Its mass is well defined, and the mass peak is only slightly asymmetric (due to final state radiation). A measured J/ψ mass off by 1 MeV (which is less than detector resolution) corresponds to a relative change of $\sim 1/6 \times 10^{-3}$ in momenta of the daughter muons. With a CDC radius of 1.2 m, this corresponds to a change in the radius of $\sim 200~\mu$m. This number is, unfortunately, about four times larger than the expected mechanical precision of CDC construction and drilling (about 50 μm).

A heavier resonance or utilizing the $e^+e^- \to \mu^+\mu^-$ events directly (with peak invariant mass of 10.57 GeV/c^2) could be a solution in theory, but practically, the collision mass itself is only known to a precision of > 1 MeV, and the mass peak is very asymmetric with a low mass tail (due to ISR and FSR). No resonances

with significantly higher (and narrow) mass than J/ψ are produced in sufficient amount at SuperKEKB.
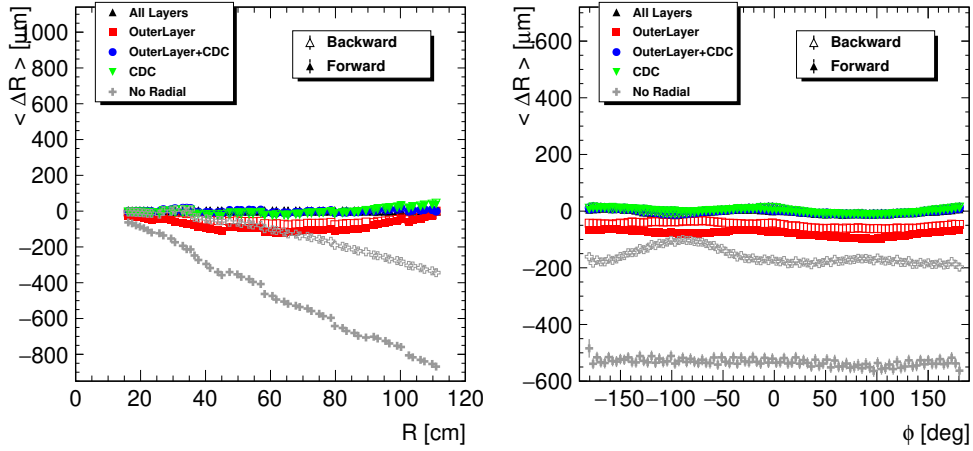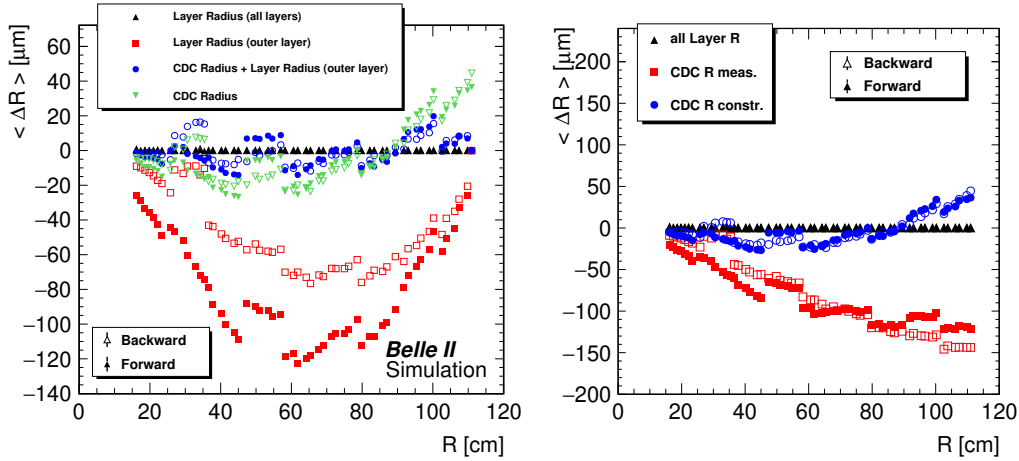

Figure 16.1: Impact of radial constraints on the average residual misalignment of wires in the radial direction for the default data sample composition as a function of wire radius (left) and azimuthal angle (right). Open (filled) markers show backward (forward) wires.

Thus we resort to external ad hoc constraints. Either the radius is completely fixed in the baseline alignment, where radial degrees of freedom are not present at all for the CDC (only layers are aligned), or we have to impose additional linear equality constraints or external survey measurements to remove or restrict certain degrees of freedom in the full alignment with wires. Several types of radial constraints defined in Sec. 12.6.3 are considered and compared:

- no radial constraints,

- radial constraints for all layers,

- radial constraints for one outer layer,

- radial constraints for the whole CDC,

- combination of radial constraints for the outer layer and the whole CDC.

The impact of the different constraint options on the radial degrees of freedom of CDC wires is shown in Fig. 16.1. These kinds of figures will be used frequently in this chapter. Note that the actual alignment parameters for CDC wires are the shifts in x and y on the backward and forward end-plate. This is a historical compatibility choice but not necessarily the best way to interpret the results. Better separation of the different systematic effects is achieved by projecting the wire alignment corrections into the radial coordinate and the perpendicular, circular $R - \phi$ direction, similarly as we do for the constraints in Sec. 12.6.3. This gives us four different types of projections. The left subfigure in Fig. 16.1 shows the mean residual misalignment averaged over wires as a function of their radius. The right subfigure shows the azimuthal angle dependence of the same quantity. Both wire-ends are shown in each plot. Two remaining projections in

Fig. 16.2 show the averaged residual misalignment in the $R - \phi$ coordinate, again as a function of wire radius and azimuthal angle. The left subfigure will not be generally shown as the points (usually) trivially sit at zero. The reason is that average wire rotation (fixed to zero by constraints) in a layer is absorbed into layer rotation parameters.
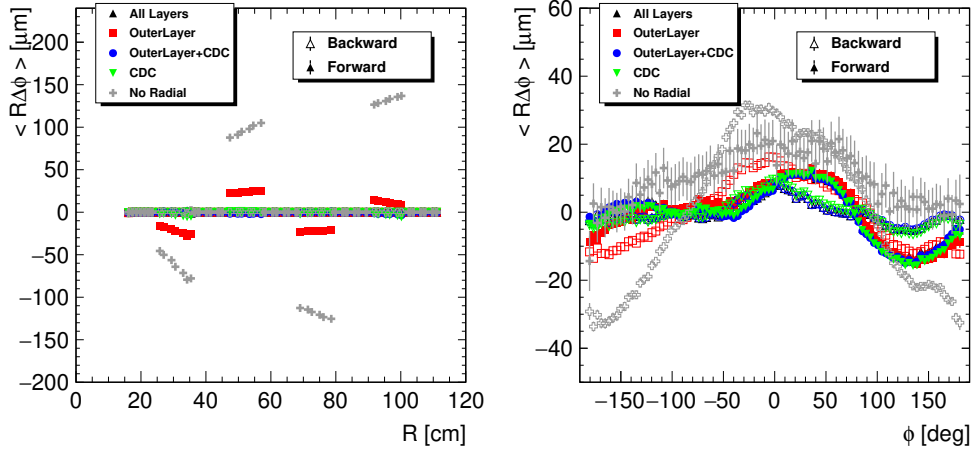
Figure 16.2: Impact of radial constraints on the average residual misalignment of wires in the $R - \phi$ direction for the default data sample composition as a function of wire radius (left) and azimuthal angle (right). Open (filled) markers show backward (forward) wires.

The gray points in Fig. 16.1 (left) show why radial constraints are needed. Without them, CDC shrinks by about 900 μm at the forward and 300 μm at the backward outer edge. The right plot shows the misalignment values averaged over radius with an almost constant offset, meaning the radial change is almost symmetric in the azimuthal angle. All the considered options with radial constraints give reasonable results. The option with only the outer layer radius being constrained results still in an average change of the CDC radius by 50 μm, although the maximal deviation occurs in the middle of the CDC, not at the outer edge. The same result is shown in detail in Fig. 16.3 (left). Based on this study, we have chosen two possible options to continue with. The combination of the whole-CDC and outer-layer radius constraints will be kept as a backup if needed for data. It was believed that the radial movement of wires has no significant physics impact [21]. Thus we will continue this chapter assuming this hypothesis and use radius constraints for each layer. It does not affect our results much, as the radial degrees of freedom seem nicely decoupled for both chosen constraint options. In the end, we can just switch to the option with CDC and outer layer radius constraint. With this option, we expect appearance of the additional small $\pm$ 20 μm systematics (blue points in Fig. 16.3 (left)). Thus, this kind of projection will be trivial as well and will not be shown in the next section.

Finally, we consider one more option. Millepede II allows the inclusion of external (survey) measurements of (linear combinations of) alignment parameters. In contrast to constraints, the equality (usually to zero) is not exact but has a defined precision. We have tested to include an external measurement for each wire position in the radial direction with zero expectation value and 100 μm

Figure 16.3: Impact of various radial constraints on the average residual misalignment of the wires in the radial direction for the default data sample composition as a function of wire radius. The left plot compares several options with constraints. The right plot compares CDC radius constraints (blue) to the case when external survey measurements for wire radii are used instead of constraints (red). Open (filled) markers show backward (forward) wires.

precision. In Fig. 16.3 (right), the result is compared to the case where two degrees of freedom are removed by using the CDC radius constraints (for backward and forward end-plate). The CDC shrinks by about 150 μm for the survey option (red points). While the used wire placement uncertainty was twice larger as expected, a naive extrapolation of the results suggests that a better precision than for the constraints cannot be achieved.

16.2.2 Impact of Data Samples and Statistics

In this section, we will use the radial constraints for each layer and investigate how the composition of the data samples affects the result. The cosmic data sample is absolutely essential to reach a good alignment already for the VXD. The IP-constrained di-muon sample is another key ingredient for achieving reasonable systematic errors for the full alignment with wires. This can be clearly seen in Fig. 16.4, when di-muons are excluded from the full alignment (blue points). A systematic trend can be observed in the residual misalignment for both ΔR and $R\Delta\phi$ components with an average amplitude of about 70 μm. The negative impact of the absence of the hadron sample on radial systematic versus azimuthal angle in Fig. 16.4 (left, red points) can probably be explained by the larger curvature of tracks in this low momentum sample, which causes enhanced sensitivity to radial movements of the wires. Making the IP position parameters floating (green points) does not significantly degrade the result with respect to the nominal case (black) with all data samples included and fixed IP position.

So all the three main data samples we consider seem to be needed for a reasonable result. One may wonder if this result can be further improved by optimizing their size. Note that in all considered cases below, the statistical errors of the parameters are expected to be smaller than the residual systematic misalignment.
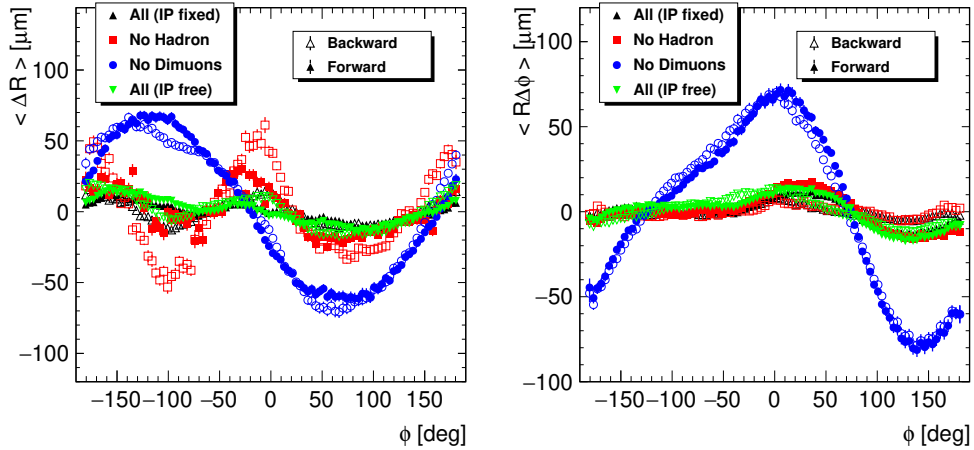
Figure 16.4: Impact of data samples on the radial (left) and $R - \phi$ residual wire misalignment as a function of wire azimuthal angle. The default data sample composition (black) is compared to cases when the di-muon sample (blue) or hadron sample (red) are not used. The option with the default sample composition and floating IP position is shown in green. Open (filled) markers show backward (forward) wires.

Fig 16.5 compares the result with the nominal size of the cosmic sample, one-half of the cosmic tracks, and $\sim$ one-fourth of the tracks. Surprisingly, lowering the number of cosmic ray tracks seems to improve the average residual misalignment. This, in fact, illustrates that the balance of data samples matters. The systematic effects for the smallest amount of cosmic tracks (blue points) are not better (closer to zero) over the full ϕ range. It becomes increasingly difficult to judge an overall alignment quality at this level. For example, the reduced cosmic sample leads to a smaller amplitude of the systematic effects, most pronounced in the outer layers. However, the average precision for most wires gets slightly worse. In general, the RMS of the determined residual misalignment parameters of the wires is 9 to 12 μm. We will return to a more detailed look into the alignment result later. For now, we should only remind that no ideal optimization target variable[2] exists, and some optimization steps might be based on subjective reasoning.

Cosmic ray tracks are getting a large weight in the alignment solution thanks to their high momenta. Thus not many of them are needed to constrain some global distortions. On the other hand, cosmic tracks suffer from azimuthal and polar asymmetry. This was already studied in the baseline alignment, where this asymmetry is responsible for the induced deformation of the CDC layers in the y direction. A significant amount of tracks from the IP is needed to suppress this residual effect, especially if those are mostly low momentum hadrons. Also, in this case, we seem to observe that a large number of tracks from the IP will be needed to reduce these unwanted features of the cosmic sample. Let us vary the other samples to understand which features they influence and how the scale of

---

[2]For example, it is not clear how much a priori weight we assign to the quality of the wire alignment compared to the VXD alignment. Also, inner wires are likely more important for physics than outer wires. Some types of systematic biases might be more relevant than others, despite having the same scale, etc.
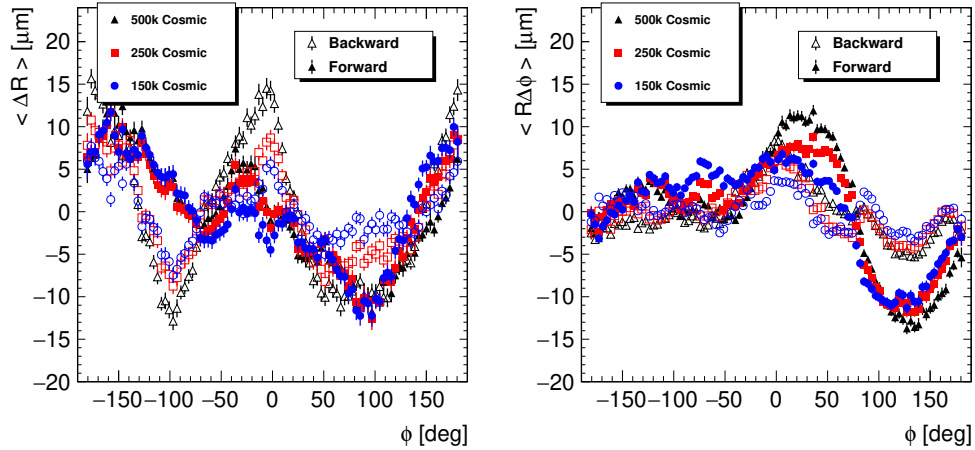
Figure 16.5: Impact of the size of the cosmic sample on the average residual wire misalignment in the radial (left) and $R - \phi$ direction (right) as a function of wire azimuthal angle. Open (filled) markers show backward (forward) wires.

the effect depends on the sample size.

The result of variation in the size of the hadron sample is demonstrated in Fig. 16.6. Its impact is primarily relevant for the radial systematics, and only a minor reduction of the amplitude (about 5 μm) is observed when two million instead of one million tracks are used. The result from variation (reduction) of the cosmic sample suggests that 4×10^6 tracks from the hadron sample could be optimal, but the improvement from the case of 2×10^6 tracks will be only minor.
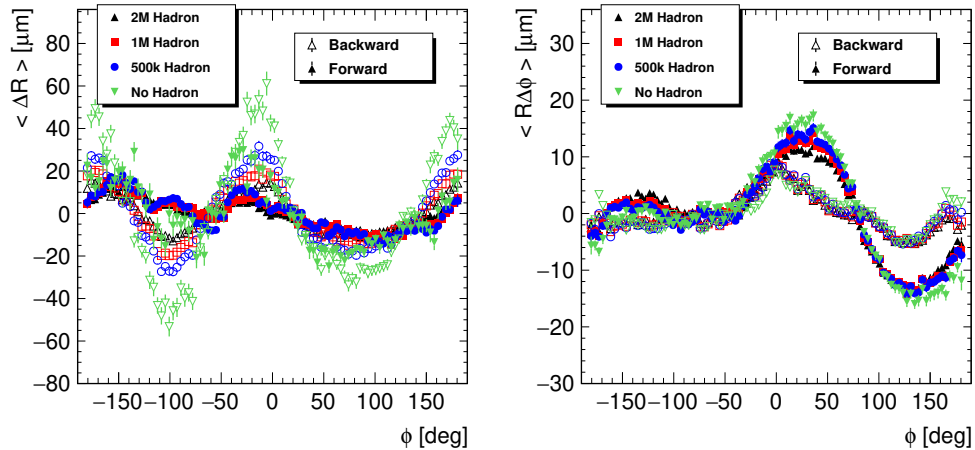
Figure 16.6: Impact of the size of the hadron sample on the average residual wire misalignment in the radial (left) and $R - \phi$ direction (right) as a function of wire azimuthal angle. Open (filled) markers show backward (forward) wires.

Finally, the size of the di-muon sample is varied. In Fig. 16.7, the default case with 5×10^5 di-muon events (black) is compared to twice (red) and half the amount (blue). With 1×10^6 events, the $R - \phi$ systematics (right) is further reduced to ~ 10 μm, while radial systematics is almost unchanged. We will return to the evaluation of the impact of such residual misalignment later, but already

at this point, any of the options should result in an acceptable alignment.
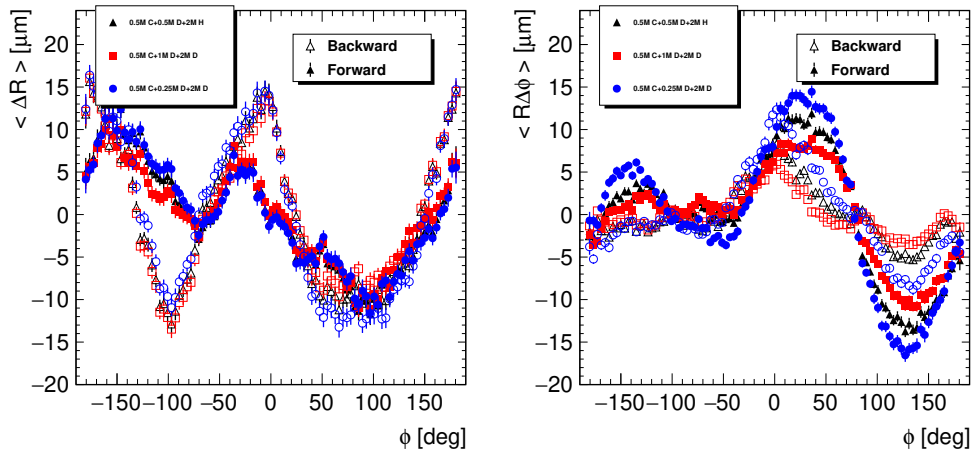
Figure 16.7: Impact of the size of the di-muon sample on the average residual wire misalignment in the radial (left) and $R - \phi$ direction (right) as a function of wire azimuthal angle. Open (filled) markers show backward (forward) wires.

One should not over-optimize at this stage. The MC data samples are not the exact equivalents of their data counterparts. Also, the exact number of the input tracks is not known beforehand, as different amounts can pass selection criteria. Especially the p-value and CDC hit quality requirements could reject a different amount of tracks depending on the CDC calibration. Therefore the sample size can vary by $\sim 30\%$ in practice.

This MC study suggests the ideal combination of the data samples is 5×10^5 cosmic ray tracks, 1×10^6 di-muon events, and 2×10^6 to 4×10^6 hadron tracks. This chapter will be continued using the default 2×10^6 hadron tracks, but for data, we use 4×10^6 hadron tracks. Such an amount of data can be reasonably obtained in less than two days at typical 2021 luminosities (1 fb^{-1}/day), taking a safety factor of two for the di-muon sample into account[3].

16.2.3 Interplay of Wire and Baseline Alignment Parameters

So far, we have only concentrated on the residual misalignment of the CDC wires. However, in each case, all the parameters of the baseline alignment were floating, too. This must be the case to resolve all correlations. On the other hand, the VXD and CDC layer alignment can be degraded with respect to the baseline case. The most apparent changes are more significant systematic deviations in the layer alignment, which is also one of the most problematic sets of degrees of freedom in the baseline alignment. The global residual misalignment of the CDC layers and VXD half-shells in x and y coordinate is demonstrated in Fig.16.8, where the default sample composition is compared to options with excluded samples, as we did for Fig. 16.4.

---

[3]The di-muon sample is driving the integrated luminosity requirement (currently). Tracks from hadronic decays are very abundant, and (usable) cosmic ray tracks are recorded at a frequency higher than 20 Hz, such that the required amount of data is acquired sooner than
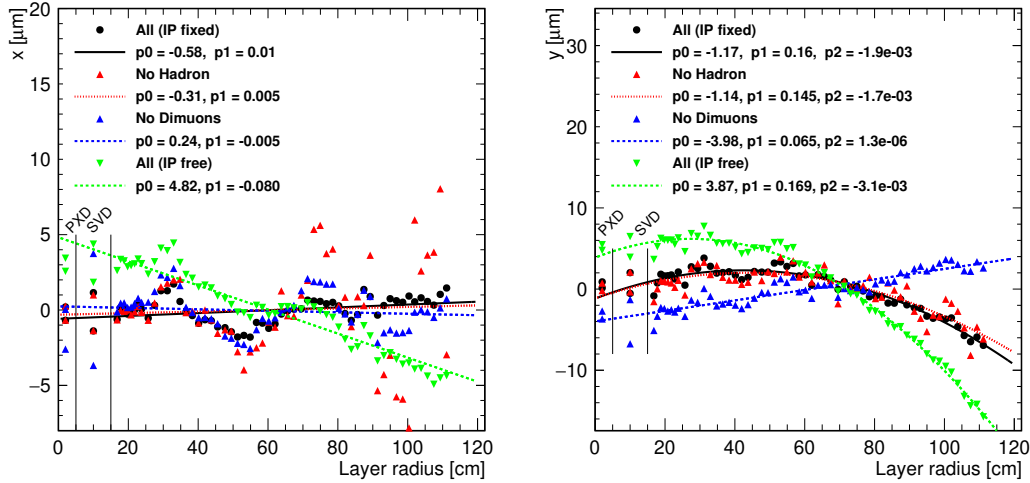
Figure 16.8: Residual misalignment of the VXD half-shells and CDC layers in the vertical (left) and horizontal direction (right). The VXD half-shells are put at an arbitrary radius for visualization. The default data sample composition (black) is compared to cases when the di-muon sample (blue) or hadron sample (red) are not used. The option with the default sample composition and floating IP position is shown in green. Linear or quadratic fits to the CDC layer points are shown (the VXD points are not included in the fit).

The large impact of the hadron sample for systematics in the global x direction is likely caused by the different topology (large curvature) of low momentum tracks compared to high energy muons from the di-muon and cosmic samples. We have also tested the possibility of keeping the IP position floating in the alignment. This obviously leads to larger but still acceptable systematics. In any case, we will keep the IP position fixed to a previously well-determined (and time-dependent) position and profit from this additional strong constraint. The possible unresolved correlations are expected to be negligible if the wire alignment is started from a fully converged baseline alignment. This way, the systematic global offsets from the baseline alignment will be imprinted in the full alignment. In principle, further iterations with time-dependent IP determination can remove any remaining effects.

Regarding the sensor level alignment, it was confirmed that all reasonable configurations with all data samples lead to local sensor alignment almost identical to the case where all wires are fixed. It can be shown that the additional degrees of freedom can truly degrade the internal VXD alignment. In Fig. 16.9, the residual misalignment for the VXD sensor rigid body parameters is shown for the default sample composition. This can be hardly distinguished from the case with fixed wires (the maximum deviation is only 0.2 μm larger, and the color pattern nicely matches visually). On the other hand, when the di-muon sample is excluded, not only the wire alignment gets significantly worse in Fig. 16.4, but also the internal VXD alignment is degraded significantly. This can be seen in Fig. 16.10, which has a twice wider range of the scales, and one can see the relative color pattern among the parameters changes. This effect must be caused

---

for the di-muon sample at the current SuperKEKB luminosity.
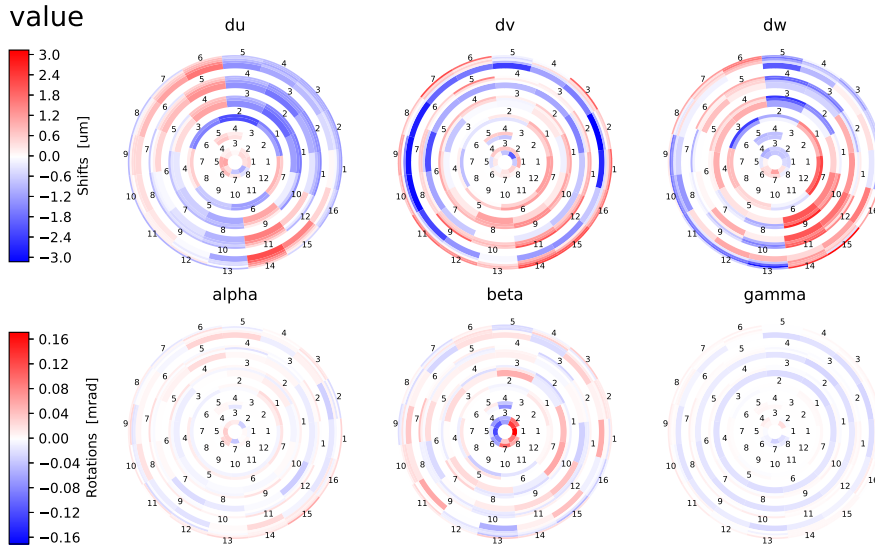
Figure 16.9: Residual misalignment for the VXD sensor rigid body parameters for the default data sample composition.

by the wire alignment as the cosmic and hadron samples together are sufficient to reach maximum rigid body parameter deviations of ≤ 4 μm with fixed wires. The alignment parameters of the VXD sensor surface deformations also experience some degradation in such cases. A more detailed overview of the residual misalignment of all the parameters will be given at the end of this chapter once the final alignment configuration is found[4].

16.3 Studies with Wire Misalignment

At this point, we are ready to test the procedure with a misaligned detector. A single realistic wire misalignment will be used for these tests. This misalignment was derived during preliminary studies of wire alignment on data taken before Summer 2020. The RMS of the wire misalignment parameters ranges from 39 to 48 μm (larger at the forward end-plate). It contains both statistical and systematic misalignment. However, the radius of all layers was fixed by constraints, and thus no radial misalignment is present on average in each layer. In addition, because this misalignment is derived from data, the wires in layer 54 are not misaligned, as this layer was deactivated.

In contrast to CDC wires, the CDC layers and the whole VXD are not misaligned, as one of the necessary checks is to confirm that the baseline alignment does not get significantly degraded by the inclusion of the additional thousands of degrees of freedom. This was checked in the previous section, and the optimal

---

[4]But the VXD sensor parameters change only negligibly, so these are skipped. Also, thanks to the di-muon sample, the precision is slightly better than for the baseline alignment, and thus, following our conservative philosophy, we keep the baseline values as our precision estimates for non-wire parameters.
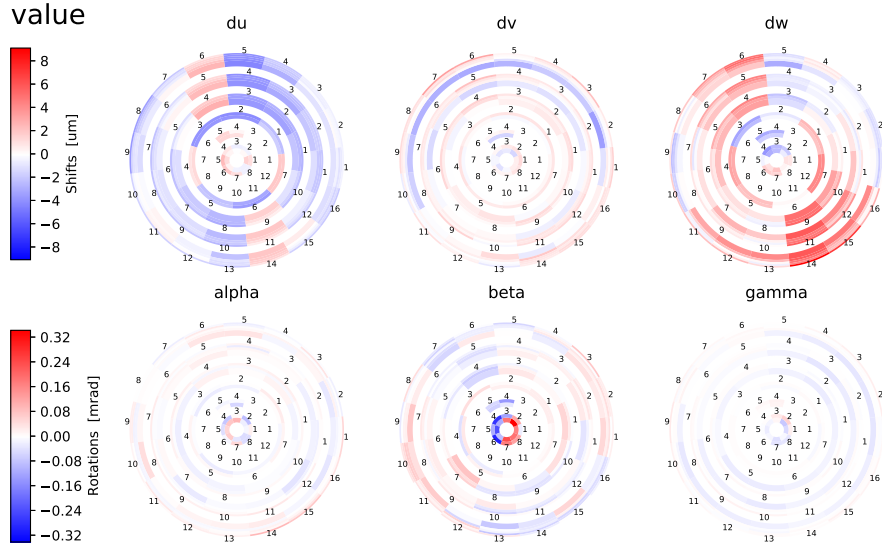
Figure 16.10: Residual misalignment for the VXD sensor rigid body parameters for the default data sample composition without di-muons.

configuration truly achieves almost identical alignment quality for the baseline parameters as in the case when the wires are fixed.
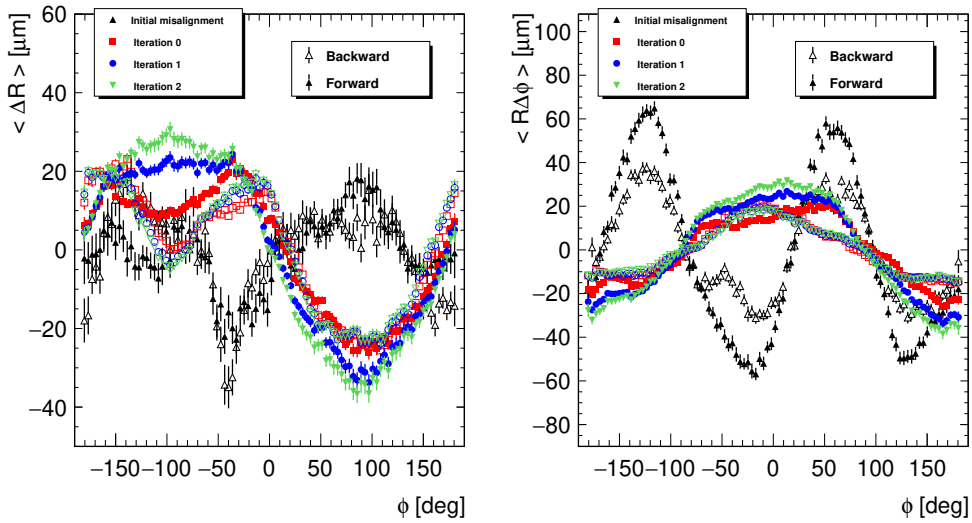
Figure 16.11: Average initial (black) and residual misalignment from the first (red), second (blue), and third (green) iteration of the alignment with the optimal MC sample composition and radial constraints for all layers in the radial (left) and $R - \phi$ direction as a function of wire azimuthal angle. Open (filled) markers show backward (forward) wires.

The alignment with the optimal sample composition from previous sections and radial constraints for all layers is iterated twice with the full reconstruction, starting with the initial realistic misalignment. The additional iterations were motivated by the observation of a weak mode appearance. While we cannot run

diagonalization for such a large problem to extract the eigenvectors corresponding to the lowest eigenvalues directly, we can iterate the alignment such that the weak mode gets more pronounced and could be extracted by looking at the result. The initial and residual misalignment after iterations 0, 1, and 2 are shown in Fig. 16.11. The initial systematic distortions are not present anymore, but a weak-mode deformation develops. The residual misalignment after the final iteration is also visualized in Fig. 16.12.
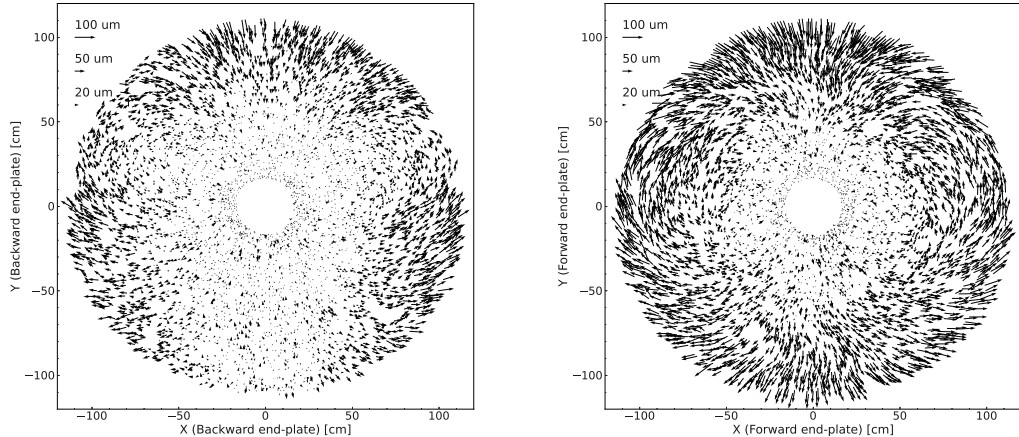
Figure 16.12: Residual wire misalignment for the third iteration of the alignment with the optimal MC sample composition and radial constraints for all layers on backward (left) and forward end-plate (right). Only 20% of wires are randomly shown by an arrow denoting the direction and size of the misalignment.

One can identify a clear systematic pattern in these figures, slightly different for the backward and forward end-plate wire ends. At the first order in Fourier decomposition, the weak mode behaves like an even (cosine-like) function of the azimuthal angle for the radial distortions and an odd (sine-like) function for the $R - \phi$ distortions. The average amplitude of this effect is about 30 μm over all wires and quickly grows, starting from the middle of the CDC, with some outer wires being misaligned as much as $\sim$ 100 μm. Note that the average of the radial wire movements in each layer is zero in these plots.

Two possible ways to mitigate this weak mode were considered and tested. The first idea was to fix all wires in some outer layer. If this layer was not initially misaligned (this is the case for layer 54 in our realistic testing misalignment), this results naturally in almost weak-mode-free alignment. The amplitude of systematic effects in our typical projection plots is suppressed to almost 5 μm, and the RMS of the residual wire misalignment is 8–9 μm. However, for this result, several hundreds of degrees of freedom are removed. Furthermore, in reality, this layer will be misaligned. To check a more realistic case, we fixed the outer layer 55 with misalignment derived from data. The result was not acceptable: the outer layer misalignment gets imprinted in the overall CDC deformation, resulting in the amplitude of average systematic effects in the $R - \phi$ direction close to 70 μm. Thus this idea was abandoned.

The second idea was to break the symmetry of the induced weak mode de-
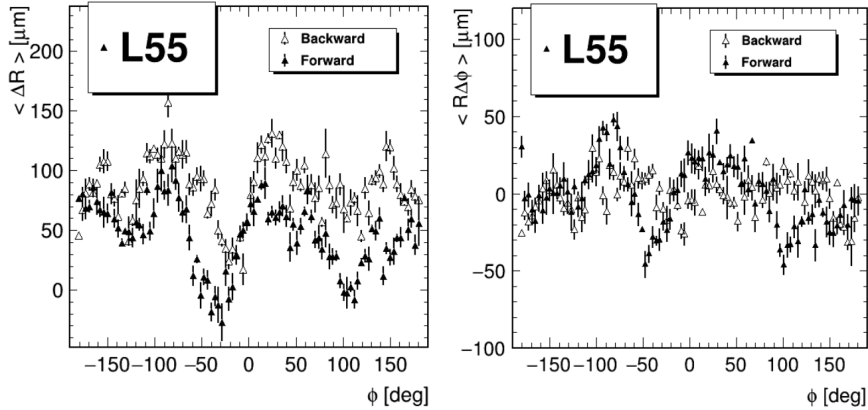
Figure 16.13: Average wire alignment constants projected into radial (left) and $R-\phi$ direction (right) as function of wire azimuthal angle determined for wires in the outer CDC layer in real data (`bucket 15`). All remaining wires' alignment is reset to survey positions. Open (filled) markers show backward (forward) wires.

formations by special constraints. The basic idea behind this approach can be understood from Fig. 16.11. In the left subfigure, one can require that the average change of the radius is not zero over all wires in the layer but separately for wires in the top and bottom hemispheres. This will automatically cancel the sine-like behavior, at least in this outer layer. Similarly, we can split the default constraint, which fixes the average rotation of wires in the layer, and include two constraints per end-plate instead, separately for wires in the left and right hemisphere. This will cancel the cosine-like behavior in the right subfigure. These hemisphere constraints are defined in Sec. 12.6.3. Because four default constraints are removed and replaced by eight new ones, there are four fewer degrees of freedom.

Of course, any actual misalignment that does not respect the hemisphere constraints will be again imprinted in the CDC deformation. Thus a quick preliminary check with real data was also performed. The wire alignment was reset to zero (surveys [21]), and starting from the baseline alignment, only the wire positions in the last layer were aligned. Assuming the baseline misalignment can be, in a reasonable approximation, decoupled from the wire alignment, no systematic effects stemming from the alignment method itself are present at this stage. The resulting determined misalignment in our traditional projections is shown in Fig. 16.13. We do not observe any indications that this misalignment does not respect the hemisphere constraints significantly, and thus the imprinted effects should be small. On the other hand, note that the estimated layer radius is different from zero on average. It is not possible to tell if this is a local effect (and layer 55 really has its radius larger by 50 μm) or a global effect, and the whole CDC radius is different. It, however, suggests the level of effects that we are going to neglect by imposing a layer radius constraint on just a single layer. This is entirely acceptable, as the corresponding relative change of the CDC radius is less than 1×10^{-4}. Based on physics performance studies, a fine correction to the measured momenta can be achieved by scaling the momenta of the final state particles at the analysis level.

From now on, the hemisphere constraints will be used. The MC misalignment study is repeated while imposing the hemisphere constraints on the (misaligned)
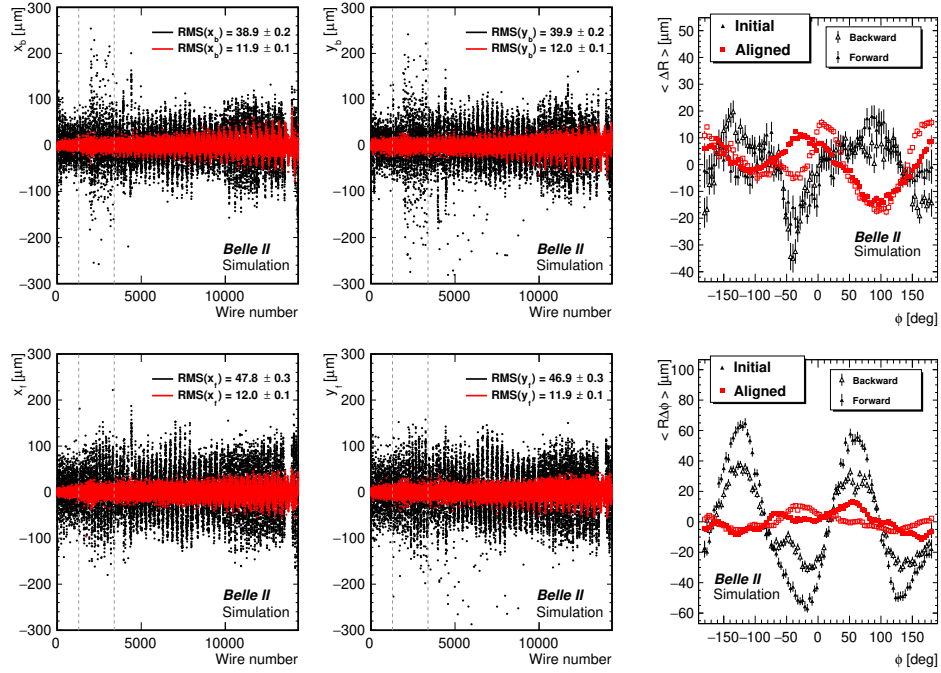
Figure 16.14: Initial (black) and residual (red) misalignment for the optimal MC configuration and hemisphere constraints. Parameters are shown wire-by-wire (right four plots for each wire parameter) and as projections of the residual wire misalignment into the radial (top right) and $R - \phi$ direction (bottom right) as a function of azimuthal angle.

outer layer 55. A test using the non-misaligned layer 54 showed only a small improvement (by 1 μm for RMS of wire parameters), again indicating it is unlikely that the imprinted effects will be severe. To take this into account anyway, we continue using hemisphere constraints for layer 55. In Fig. 16.15 the initial realistic misalignment is compared to the first iteration of our new alignment configuration. Next to the usual projection plots, the parameters are also compared wire-by-wire. The RMS of the distribution of the individual wire alignment parameters is shown too. The average misalignment is reduced significantly, down to ~ 12 μm. One can still see the remaining systematic effects in the radial coordinate—but with a maximal average amplitude of 20 μm, these are expected to be entirely negligible for physics.

The misalignment is almost negligible in the inner part but can grow up to 50 μm for outer wires. This can be clearly seen by splitting the residual misalignment projections by layers, for example, as in Fig. 16.16. The effect of removal of just four (two at each end-plate) degrees of freedom can also be shown by visualization of the residual wire misalignment in Fig. 16.15. This figure uses the same scale as Fig. 16.11, and the difference is evident. Note that some asymmetries might be induced by the misalignment of the realistically misaligned layer 55, which does not respect the hemisphere constraints exactly.

At this point, we shall briefly return to the main problem that motivated all this effort: the limitations of the approximate MINRES-QLP solution method. Parameter variations, like numerical tolerance, never led to a satisfying solution when starting with a realistic misalignment. While this method is successfully
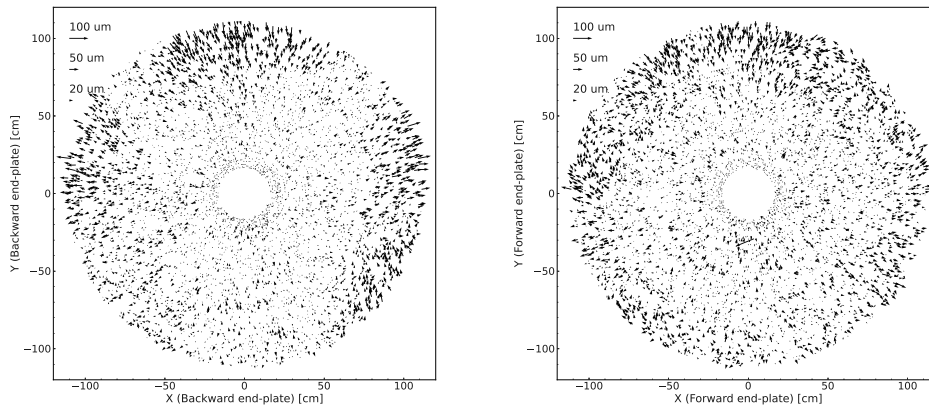
Figure 16.15: Residual wire misalignment for the first iteration of alignment for the optimal MC configuration and hemisphere constraints on backward (left) and forward end-plate (right). Only 20% of wires are randomly shown by an arrow denoting the direction and size of the misalignment.
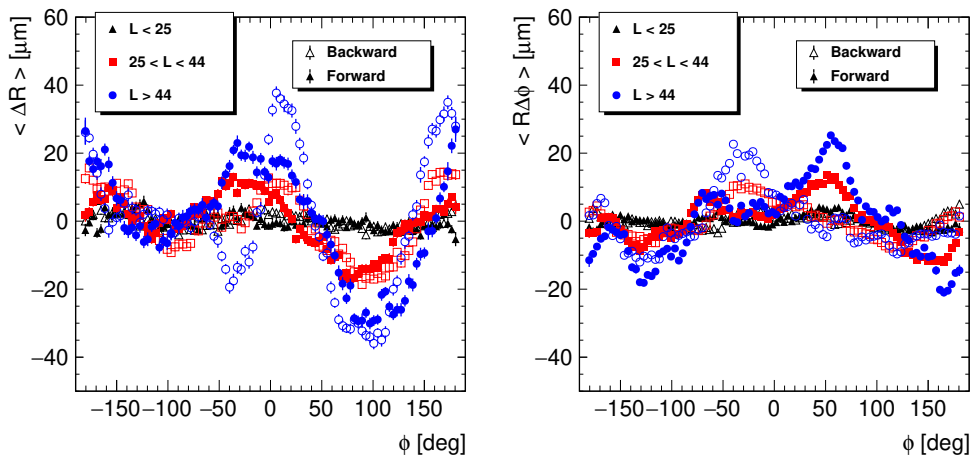
Figure 16.16: Residual wire misalignment split by layer in the radial (left) and $R-\phi$ direction (right) as a function of wire azimuthal angle for the first iteration of alignment for optimal MC configuration and hemisphere constraints. Open (filled) markers show backward (forward) wires.

used for the CMS tracker alignment [74], their global matrix is sparse. This is not the case for our alignment problem, with more than 90 % non-zero elements. As a demonstration, this approximate method is used in the exact same alignment configuration as we have just derived above. The result is shown in Fig. 16.17: the $R-\phi$ systematic pattern is not resolved, and the overall result for individual wires shows a final RMS from 28 to 41 μm.

The improvement with an exact solution method is striking and indicates possible limitations of the MINRES method in alignment problems without a sparse global matrix. The possibility to profit from the inclusion of all correlations among the measurements induced by tracks crossing different detector elements is a crucial advantage of the global Millepede II approach. However, it can be

284
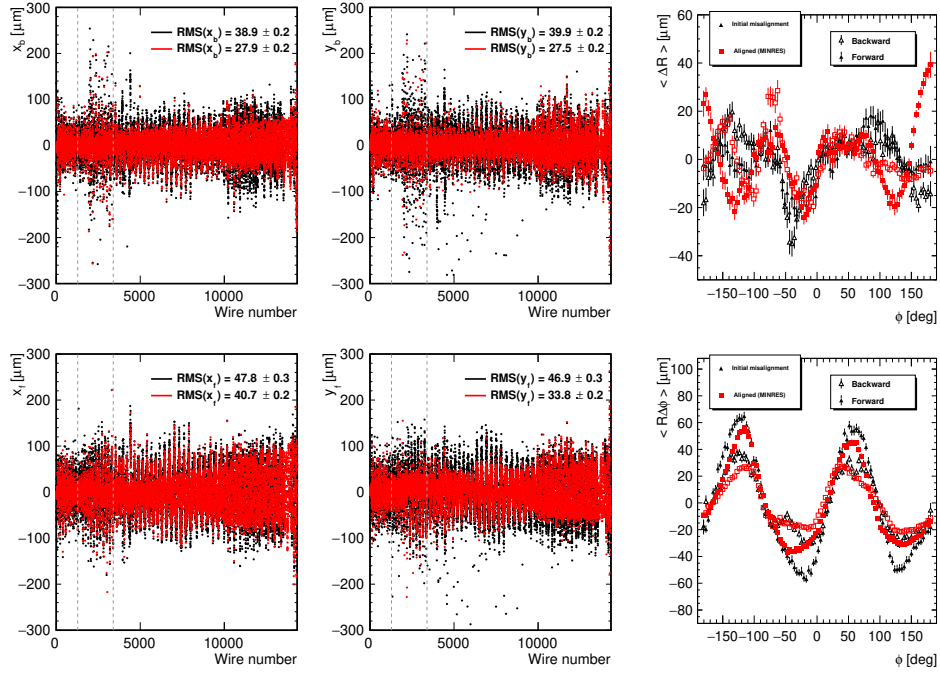
Figure 16.17: Results for MINRES-QLP solution method: Initial (black) and residual (red) misalignment for the optimal MC configuration and hemisphere constraints. Parameters are shown wire-by-wire (right four plots for each wire parameter) and as projections of the residual wire misalignment in the radial (top right) and $R - \phi$ direction (bottom right) as a function of azimuthal angle.

considered a happy coincidence that the size of our alignment problem gives a reasonable solution time on the currently available (high performance) hardware. Even with the most modern methods, the solution time still scales as the cube of the number of the alignment parameters. Thus, for example, for CMS, exact solution methods could be are again too slow practically, with more than several hundreds of thousands of parameters [85].

One last untested hypothesis should be checked before moving to studies with real data. The residual misalignment should not affect physics. This was also confirmed using a simple MC study. As a generator, a particle gun with pions in the full angular acceptance is used. The momentum is set to exactly $2\,\mathrm{GeV}/c$, and the tracks are generated exactly at the center of the global coordinate system. In Fig. 16.18, distributions of the reconstructed vertex parameters and momentum from 5×10^4 generated events are shown. The ideal case is compared to a reconstruction using the residual misalignment discussed above. The IP position is fixed to the true value, so no significant shift of the vertex parameter distributions is expected. The observed biases are caused by small shifts of the whole VXD detector together with CDC layers, up to 1 μm, and are only observable in such simplified MC studies, where the true (generated) vertex is known precisely. In reality, especially the shift in the z-coordinate has no measurable reconstruction effect for data unless really large. A more interesting distribution is the reconstructed momenta. No change in the bias is observed, and the width does not increase significantly. The degradation of the momentum resolution is entirely negligible. We have also checked, for example, the azimuthal dependence of the
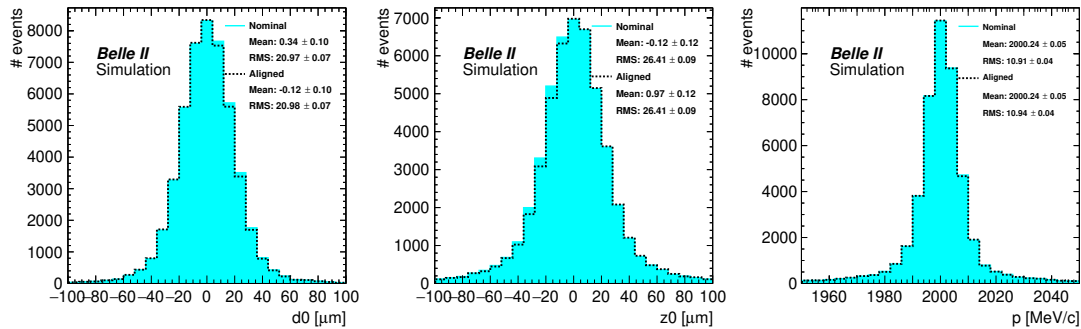
Figure 16.18: Results of simple MC validation: reconstructed particle transverse impact parameter d_0 (left), longitudinal impact parameter z_0 (middle) and momentum (left) for ideal geometry (cyan) and geometry with residual misalignment (black) from the optimal MC configuration.

reconstructed momenta. No statistically significant differences to the ideal case have been observed.

16.4 Studies with Real Data

For the studies in this section, data from the last bucket (`bucket 15`) before the 2020 Summer shutdown are used. The data sample composition for alignment follows the results of the MC study:

- 5×10^5 events from the cosmic skim,

- 10^6 events from the tight di-muon skim, and

- 4×10^6 events from the hadron skim.

The prompt alignment and calibration are used as a linearization point, but all wire alignment parameters are reset to zeros, removing any previously determined values. Only the survey measurements are implemented as wire displacements, and tensions [21]. The previously determined time-dependent IP position from the prompt alignment is used as a fixed constraint.

The cosmic validation was used to compare the performance of the various alignment options. The comparison of the prompt and 2020 reprocessing alignment, which is based on the output of the studies in this chapter, is discussed in detail in Chapter 17. Here, we concentrate on the difference between the default alignment configuration derived from the MC studies and the final configuration with the best performance on data. Note that there is some freedom in our choice of constraints. For example, fixing the radius for some layer might result in a less biased alignment if that particular layer happened not to be significantly misaligned in its radius. Effectively, such a layer then becomes the reference for the radial scale. This can only be optimized by comparing the performance of different alignments of real data. Such a process can be quite subjective and time-consuming. In general, we require that no available validation plot deteriorates significantly with the new alignment.
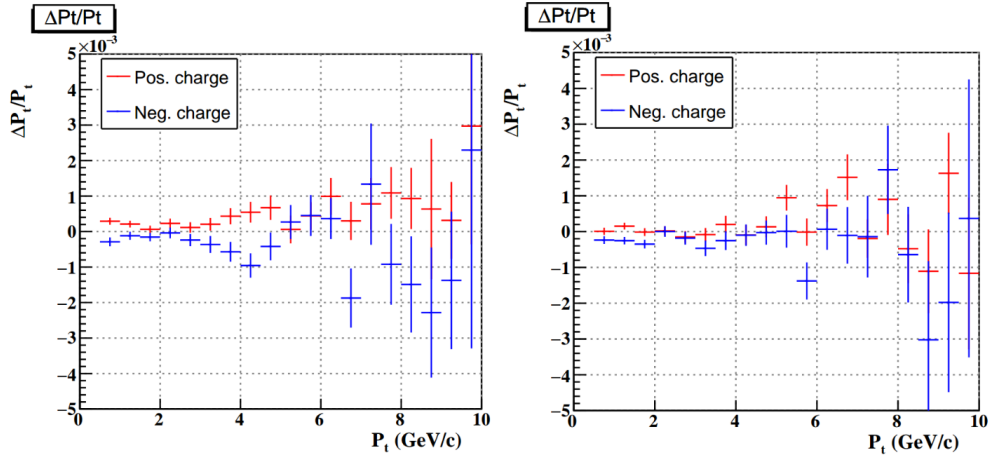
Figure 16.19: Standard CDC cosmic validation plots for tracks with at least four VXD hits for the optimal MC configuration (left) and configuration with the best data performance (right), showing relative transverse momentum bias as a function of the transverse momentum for positively and negatively charged tracks.

In the previous section, we derived the optimal alignment configuration based on MC studies. It was assumed that radial degrees of freedom have a small physics impact, and thus all layer radii could be kept fixed by constraints. Whereas some of the observed problems in the validation of the prompt alignment results were reduced with the new, full realignment, several observables showed worsened performance. The original wire alignment used for the prompt calibration was derived without the special constraints and thus had larger freedom to reduce possible true misalignment of the chamber.
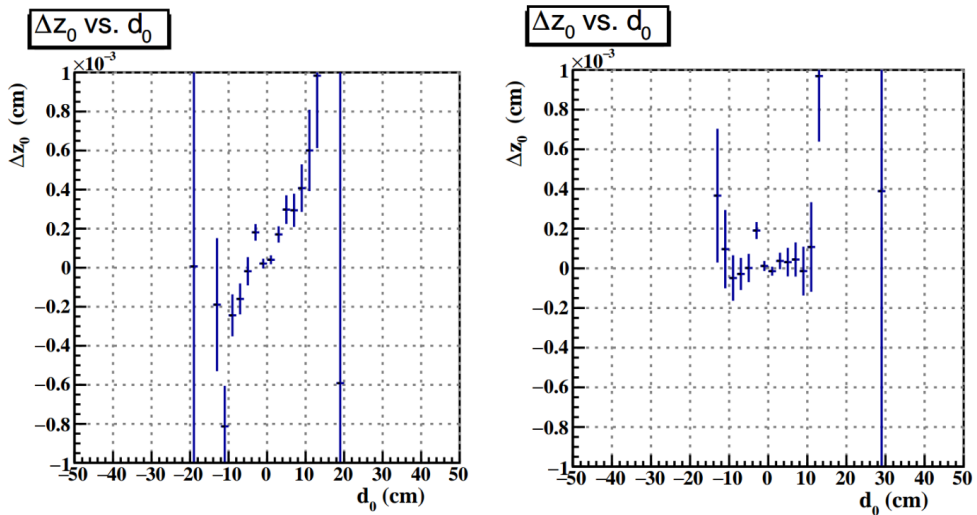
Figure 16.20: Standard CDC cosmic validation plots for tracks with at least four VXD hits for the optimal MC configuration (left) and configuration with the best data performance (right), showing z_0 bias as a function of the transverse impact parameter d_0 of the cosmic tracks.

We tried all the initially discussed radial constraint options, as well as remov-

287

ing some constraints for CDC layers. Only one configuration led to no significantly worse performance than the existing prompt alignment while solving some of the previously observed problems in the alignment validation. This best configuration required removing the CDC twist constraint and z-offset constraint. Layer radius constraint is required only for layer 53. The radial constraints for the whole CDC are also used. The hemisphere constraints are applied to the outer layer 55. The most significant differences from the optimal MC configuration can be seen in Figs. 16.19 and 16.20.
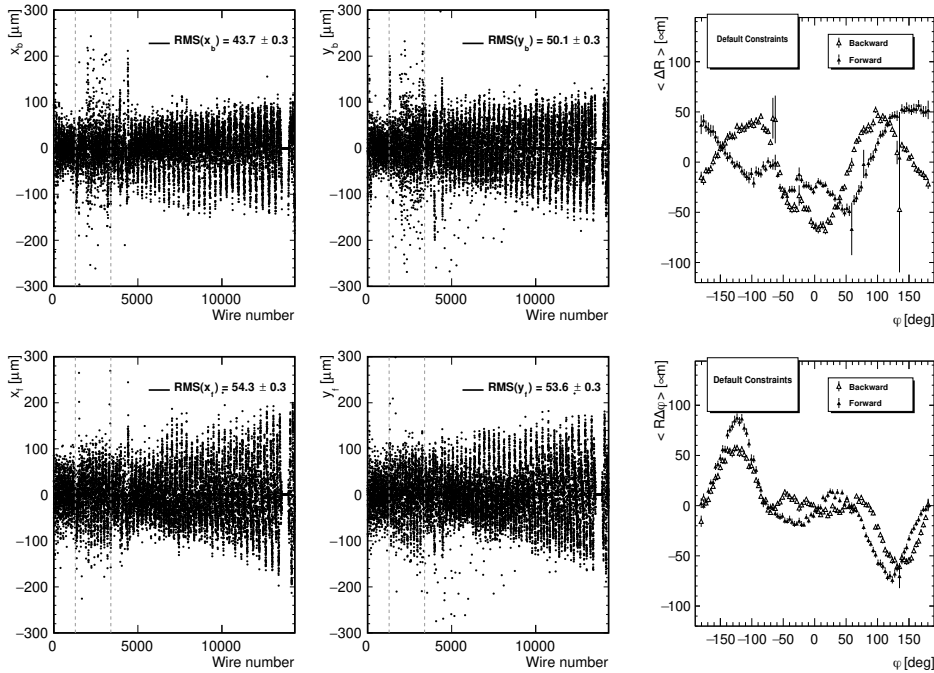
Figure 16.21: Determined wire alignment parameter for real data in `bucket 15` and the optimal MC configuration. Parameters are shown wire-by-wire (left four plots for each wire parameter) and as projections of the residual wire misalignment in the radial (top right) and $R - \phi$ direction (bottom right) as a function of azimuthal angle.

The determined alignment constants for the optimal constraints from the MC study are shown in Fig. 16.21. The alternative configuration with the best data performance is shown in Fig. 16.22. The large differences are most noticeable in the wire-by-wire figures, while the tendency in the projections is at least partially similar. The difference between these two alignments is essentially contained in the radial degrees of freedom. The alignment estimates a significant radial deformation in the inner and conical parts of the CDC. This can be better seen in the corresponding projection in Fig. 16.23 (left), where the average radial misalignment of wires reaches almost 200 μm in the conical section.

These radial deformations are probably connected with additional distortions of the chamber, namely a tiny twist between the backward and forward end-plate. When the twist constraint is removed, due to the interplay of the z-position measurement and stereo layer rotation, the z-offset constraint should be removed, too, while the IP position is fixed as a reference. The original alignment configuration from the MC studies did not lead to a significant change in the layer
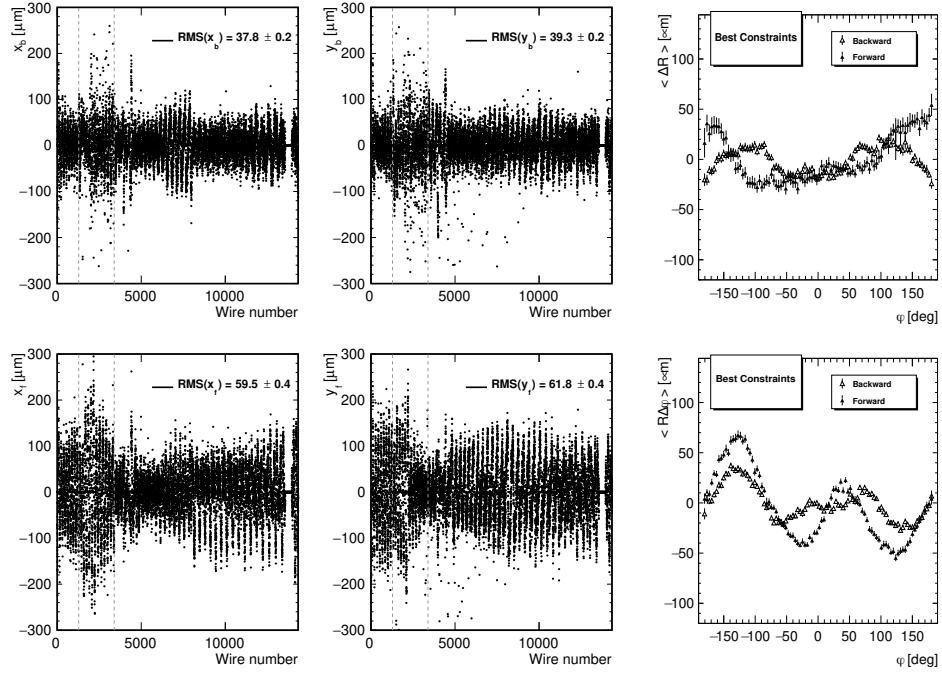
Figure 16.22: Determined wire alignment parameters for real data in `bucket 15` and the configuration with the best data performance. Parameters are shown wire-by-wire (left four plots for each wire parameter) and as projections of the residual wire misalignment in the radial (top right) and $R - \phi$ direction (bottom right) as a function of azimuthal angle.

alignment parameters determined by the baseline prompt alignment. However, significant differences appear after the removal of the constraints, as can be seen in Fig. 16.24.
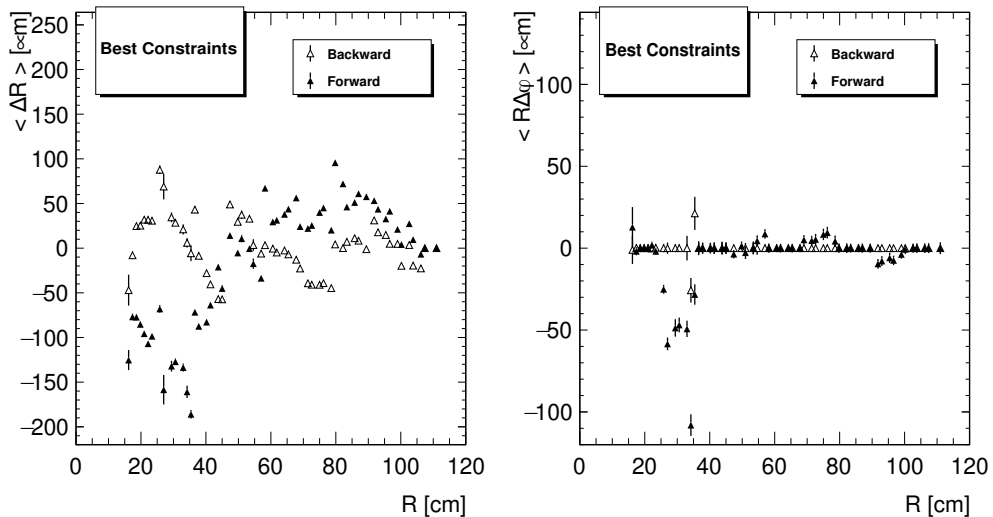
Figure 16.23: Determined wire alignment parameters for real data in `bucket 15` and the configuration with the best data performance. Shown are projections of the residual wire misalignment in the radial (left) and $R - \phi$ direction (right) as a function of wire radius.

289

These differences are primarily localized in the stereo layers. We believe the determined alignment parameters correspond (at least partially[5]) to reality as the validation results confirm a better performance. In addition, an independent study of radial deformations in the inner part of the chamber confirmed the alignment results. A more detailed look revealed a five-fold symmetry in the radial deformation of the inner layers. This symmetric pattern exactly corresponds to the deformation of the CDC inner plate that occurred during CDC construction. This structure was also confirmed by a study of material distribution using hadronic interaction vertices and in a variation of wire gains. We believe the deformation of the inner plate leads to modified electrostatic forces on the inner wires, projecting the five-fold deformation into the wire misalignment.
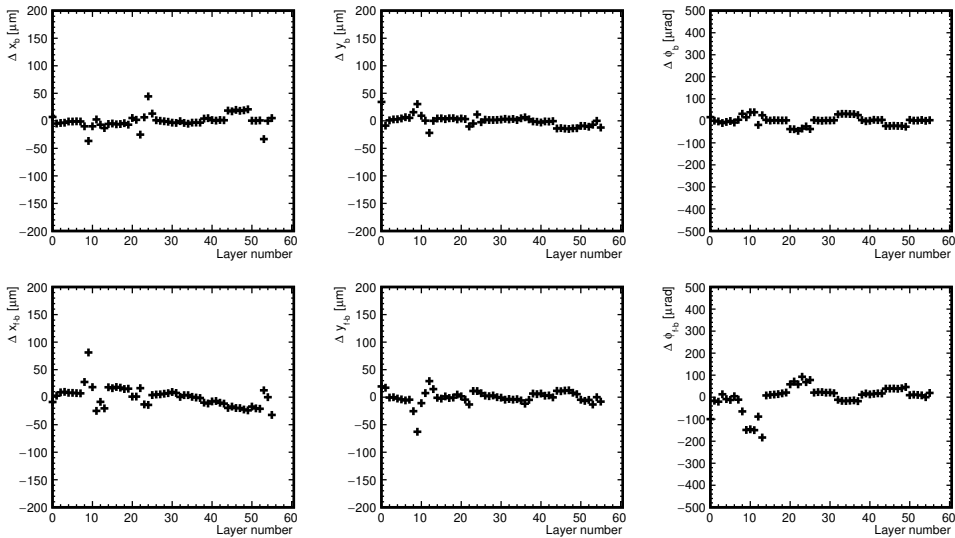
Figure 16.24: Difference between determined CDC layer alignment parameters for real data of `bucket 15` with baseline alignment and the configuration with the best data performance using released z-offset and twist constraints.

The most significant improvement with the best alignment configuration is observed in the dependence of the Δz_0 bias as a function of $\tan \lambda$ of cosmic ray tracks. We suspected such a bias is potentially dangerous for further high precision time-dependent CP-violation measurements and thus investigated ways to reduce it. While the exact mechanism of how this inconsistency arises is not understood, it is almost removed by the realignment, see Fig. 16.25 (right). In addition, we can reproduce its effects to a large degree on MC, see Chapter 17 or Fig. 13.5. As a brief demonstration that wire alignment can have an impact on precision vertex measurements, we have also realigned the baseline parameters with CDC wires fixed to surveys [21]. In Fig. 16.25 (left), these can be seen as blue points with a systematic variation having about 5 μm amplitude. No better performance can be achieved without CDC alignment at the wire level. While such effects would be acceptable in previous experiments, future precision measurements could deteriorate visibly if the full VXD+CDC alignment is not regularly performed. This becomes especially important when the CDC defor-

---

[5]We suspect inconsistencies caused by the additional CDC end-plate deformation due to weight of the VXD, not implemented in the wire displacements, namely the wire z-positions.
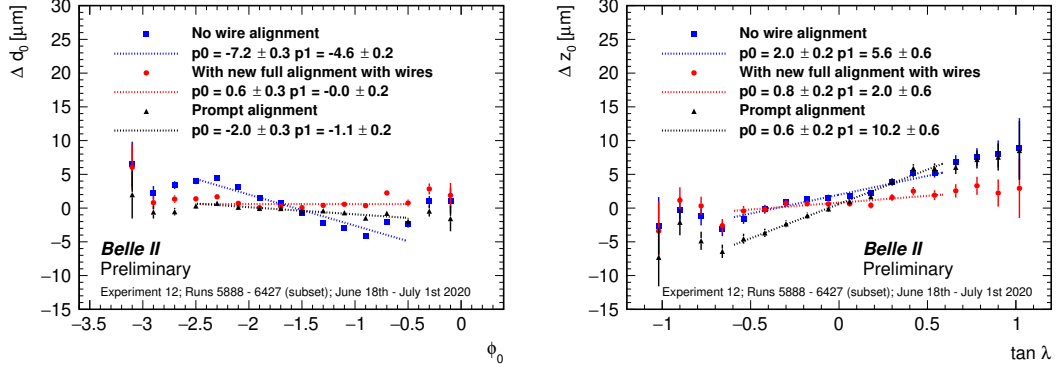
Figure 16.25: Cosmic validation of real data from `bucket 15` for tracks with at least four VXD hits showing biases (differences of upper and lower arm) of the transverse impact parameter d_0 as a function of ϕ_0 (left) and longitudinal impact parameter z_0 as a function of $\tan \lambda$. The case with wire alignment reset to surveys (blue) is compared to the preliminary prompt alignment (black) and the full alignment with the best data performance (red). Linear fits to the slopes of the biases are superimposed.

mation changes significantly with time, which is, unfortunately, the case for the Belle II CDC.

16.5 Final Alignment Overview

In the previous section, several initially assumed constraints were removed to reach optimal data performance. It is desirable to repeat the MC study for this final alignment configuration to verify that no additional significant systematic effects appear. Thus the final alignment configuration was executed, starting with the realistic wire misalignment. The wire-by-wire results are shown in Fig. 16.26. The RMS of the residual misalignment for wire positions is naturally slightly larger due to the additional radial systematics but still fully acceptable—the wires are aligned with an average precision of 13 to 14 μm, slightly better on the backward end-plate.

The projections of the initial and residual wire misalignment into the natural coordinate system allow decoupling of the individual systematics. This time, we show all four projections in Fig. 16.27. In addition to the previous optimal result on MC with radial constraints for all layers, the corresponding systematics reappear in ΔR vs. R, while the rest is almost unaffected. The variations of the wire misalignment in the radial direction can also be noticed in Fig. 16.28 in the conical regions, in addition to the already discussed systematic pattern visible mainly in the outer layers.

Regarding the VXD alignment parameters, no significant degradation is observed with the final alignment configuration when compared to the previous results. This is also the case for CDC layer alignment, which suggests that the twist constraint is, in fact, redundant and that the twist can be resolved by the cosmic sample. Nevertheless, the observed residual systematics for layer parameters shown in Fig. 16.29 are slightly larger and different than in our previous
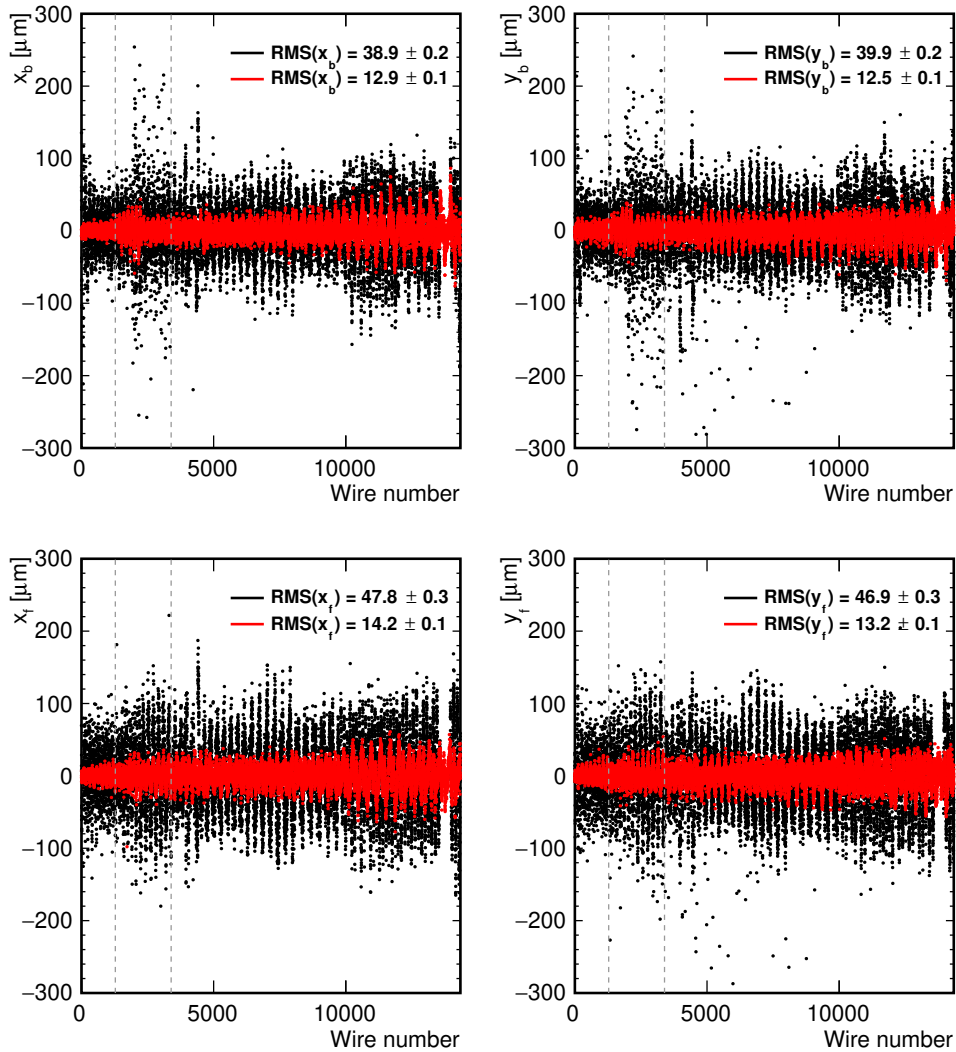
Figure 16.26: Initial (black) and residual misalignment (red) for the configuration with the best performance on data. Wire shifts at backward (top) and forward (bottom) end-plates are shown for each wire, together with the RMS of the distribution.

discussion. Notable is the slight average twist (right-most bottom plot) and the well-known deformation in the vertical direction (middle, top plot). The vertical systematics has a different sign with respect to Fig. 16.8, but a similar scale. Such effects are fully acceptable and practically unobservable as the CDC spatial resolution is about 100 μm. It is important to note that the starting misalignment does not respect the hemisphere constraints, and thus some imprinted systematics could make some of the results slightly worse. Nevertheless, such a realistic scenario is more suitable for estimating the alignment precision.

16.6 Summary and Prospects

The results of this chapter suggest that the derived alignment configuration can easily recover from any statistical misalignment as well as local CDC deforma-
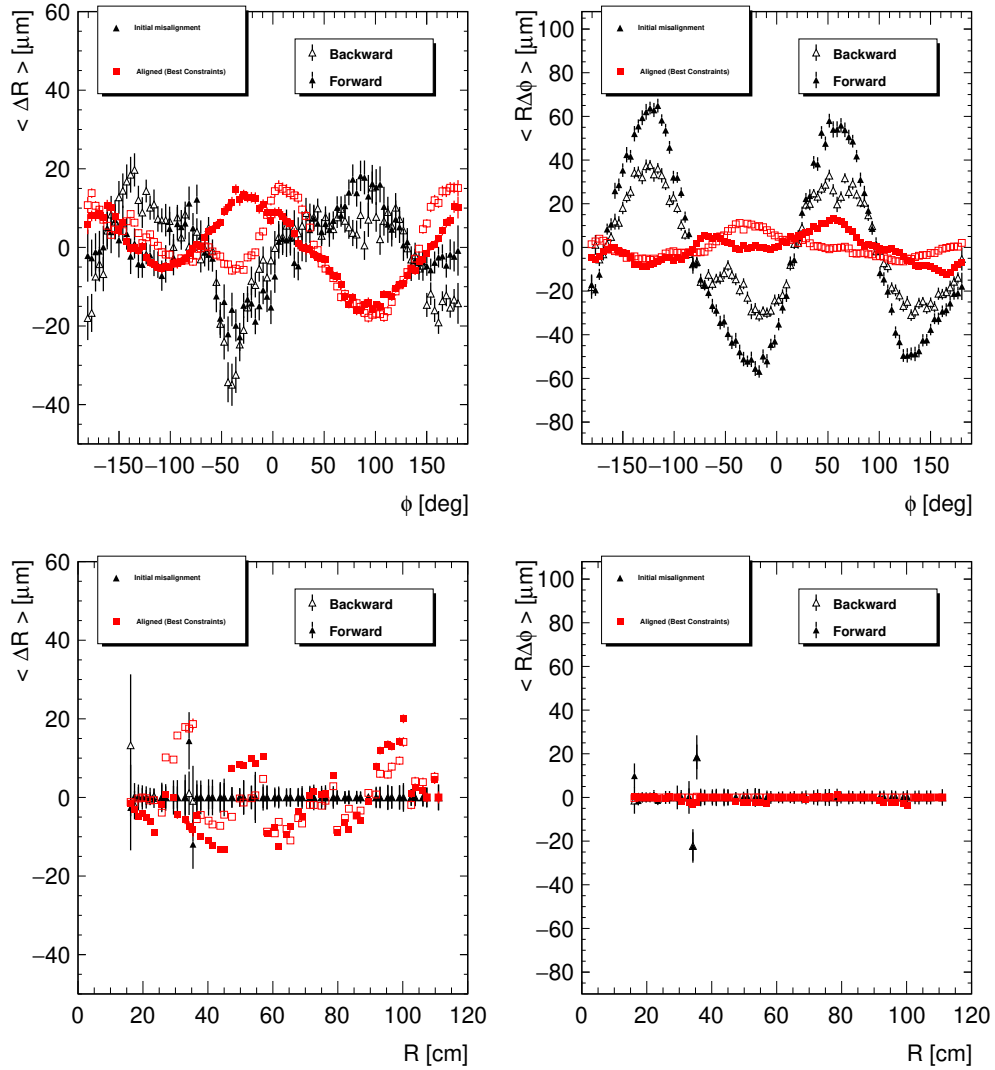
Figure 16.27: Initial (black) and residual misalignment (red) for the configuration with the best performance on data. Projections of average misalignment in the radial (left) and $R - \phi$ direction (right) are shown as a function of wire azimuthal angle (top) and radius (bottom). Open (filled) markers show backward (forward) wires.

tions. Some global distortions are expected to be resolvable, too. For example, an elliptical deformation with its long semi-axis along the horizontal direction would respect the hemisphere constraints and should be fully resolvable. On the other hand, by shifting the deformation phase by 90°, such a distortion would not respect the constraints *maximally* and could never be resolved with our method. Similarly, if the radius of the CDC is different, the method cannot correct for such an effect. On the other hand, especially in the two examples just given, a simple correction at the physics analysis level can be applied to account for such effects when observed, e.g., as an azimuthal dependence of the reconstructed invariant masses of unstable particles.

The exact same alignment configuration is used for the full realignment in the 2021 data reprocessing. On top of an average alignment per bucket, run-dependent alignments need to be derived. For this purpose, the number of floating
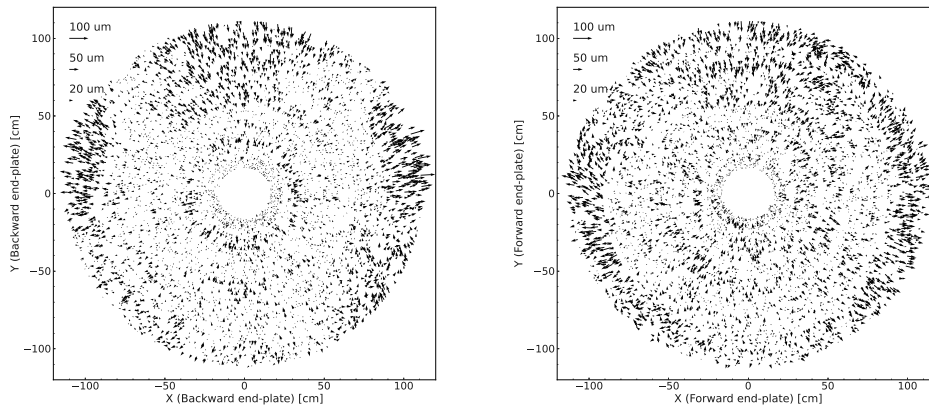
Figure 16.28: Residual wire misalignment for the configuration with the best data performance (starting from realistic wire misalignment). Only 20% of wires are randomly shown by an arrow denoting the direction and size of the misalignment.
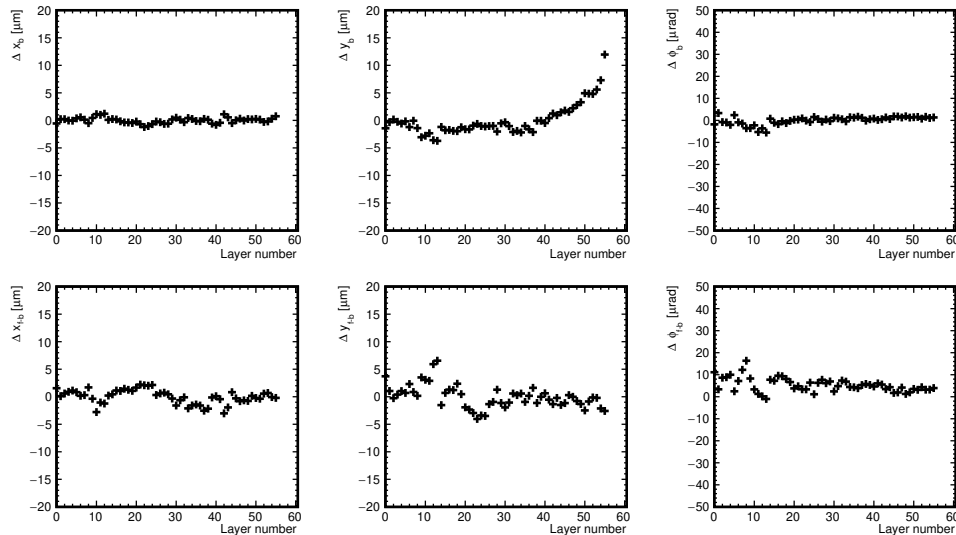
Figure 16.29: Residual misalignment for CDC layer parameters for the configuration with the best data performance. Note that while starting from realistic wire misalignment, layers are not misaligned initially.

parameters must be in a feasible range. Based on the previous investigations by the author and other collaboration members, it is necessary to perform a more frequent alignment of the VXD half-shells and PXD sensors. The time dependence of the internal SVD alignment, on the other hand, seemed relatively stable. As we already observed that most global movements of the VXD are correlated to CDC deformations, the CDC layer alignment also needs to be time-dependent. With such a limited number of degrees of freedom (and so many fixed parameters serving as a reference), the solution is expected not to suffer from significant weak modes. For the run-dependent alignment, only the di-muon data sample combined with cosmic-ray tracks is used. The default mix of the samples includes all available di-muon events mixed with about 1/5 of cosmic

294

events. Based on a simple MC study, the required statistics to reach statistical precision at the level of $\leq 1\mu$m for VXD sensor parameters and $\leq 3\mu$m for VXD half-shell z-positions (which have the lowest statistical precision) is about 8×10^4 events.

CHAPTER **17**

Prompt and Reprocessing Alignment Performance

In this chapter, the performance of the official alignments used in physics analyses will be reported and compared to Monte Carlo simulations with the help of the validation methods introduced in Sec. 13.5. We will attempt to explain the origins of some discrepancies and possible ways to mitigate them. The reported results were produced during the 2021 data reprocessing (known as `proc12`) for one continuous block of data recorded in Spring 2020 (experiment 12). The main goal was to understand alignment performance over several months of data taking and evaluate the improvement from prompt to reprocessing alignment. The gained experience should lead to further improvements for future operations at higher luminosities.

The results in this chapter do not aspire to be a definitive evaluation of the detector performance. One reason is a tight schedule and data availability in the official 2021 reprocessing. Most of the data for this chapter were processed just after the alignment finished and before the rest of the calibration chain was completed. The actual difference from the official physics performance is, however, negligible.

In several figures, we will also compare the results to an MC simulation with the prompt misalignment model extracted during the reprocessing; see also Sec. 13.6.4.

17.1   Validation with Cosmic Rays

Muons from cosmic rays penetrating the Belle II detector constitute a unique validation data sample. The momentum spectrum goes far beyond the maximal momenta of particles produced in collisions. Particles with high momentum are less affected by multiple scattering effects, and in the infinite momentum limit, the resolution of the track parameters is entirely dominated by the resolution of the detector itself. On the other hand, the angular coverage of cosmic tracks is limited as the detector is installed in an underground hall, and horizontal cosmic rays are thus scarce.

The idea of cosmic validation is to reconstruct the cosmic particle trajectory as two independent tracks, split at the POCA to the origin. The reconstructed track parameters are then compared, and their difference is a measure of the track parameter resolution. We will concentrate on full cosmic tracks, which cross all three tracking sub-detectors: PXD, SVD, and CDC. This way, we can scan the phase space of track parameters covering most of the tracks relevant for physics

analyses. For particles with longer lifetimes like K_S^0 meson or Λ baryon, one may enlarge the space by relaxing the requirement on hits in the PXD and thus allowing for a larger d_0 and z_0 of the cosmic ray tracks. The data sample is based on the cosmic skim recorded during collisions. Due to the strict requirement of at least one PXD hit and at least four VXD hits on each cosmic arm, the total collected data sample is limited. About 130×10^3 such full tracks are collected in the data for experiment 12. Note that this does not correspond to the total amount of initially recorded cosmic ray tracks. Pre-scales are applied to limit the total amount of data at the calibration center, which results in approximately 10×10^6 triggered cosmic events available per bucket. The MC sample of a similar size as the real data is generated using the Belle-based cosmic generator (see Sec. 11.7. The data-driven background for this specific experiment is simulated too.

To select a clean data sample, we require exactly two reconstructed (un-merged) tracks in an event, and for each track $|z_0| > 0.3$ cm to suppress contamination from collision events. This requirement can be clearly seen in Fig. 17.1 (top right). In this figure, the distributions of the five helix parameters (p_t is shown instead of ω) and the number of degrees of freedom of the tracks are compared in data and MC. The histograms are shown separately for the upper and lower track arm. One can generally see a nice correspondence, but the data momentum spectrum is softer than the MC one. This would then result in the natural degradation of the data/MC agreement in most validations where tracks with all momenta are combined. In principle, we could reweight the MC to match the data p_t distribution. However, even in such cases, an additional degradation is expected due to the lower average number of degrees of freedom (and thus fewer CDC hits on tracks). Thus one should expect slightly worse performance on data, even for a (purely theoretical) perfect alignment and calibration.

17.1.1   Helix Parameter Resolutions

The standard Belle II cosmic reconstruction attempts to derive the correct orientation of the two track halves before the track fitting. This is in contrast to the initial standard reconstruction, which incorrectly assumes both arms originate from the IP. Thanks to the proper orientation of the track halves, the material extrapolation and energy loss are calculated in the correct direction of propagation[1]. This also means that all five helix parameters are expected to be the same on average for both arms at any point along the particle trajectory, and no sign flip occurs for d_0.

To avoid issues with different momentum spectrums in data and MC, we use a transverse reference momentum of 5 GeV/c by selecting only tracks with 4.5 GeV/$c < p_t < 5.5$ GeV/c. At such momenta, the spread for most helix parameters reaches almost the plateau of the minimum influence of multiple scattering and thus can be used to estimate the detector resolution. The difference of upper and lower arms' helix parameters $h \in \{d_0, z_0, \phi_0, \tan\lambda, \omega\}$, defined as

$$\Delta h = h^{\mathrm{up}} - h^{\mathrm{down}}, \tag{17.1}$$

---

[1]In fact, for some cosmic tracks, the distinction between the upper and lower arm becomes ambiguous for the tracking software.
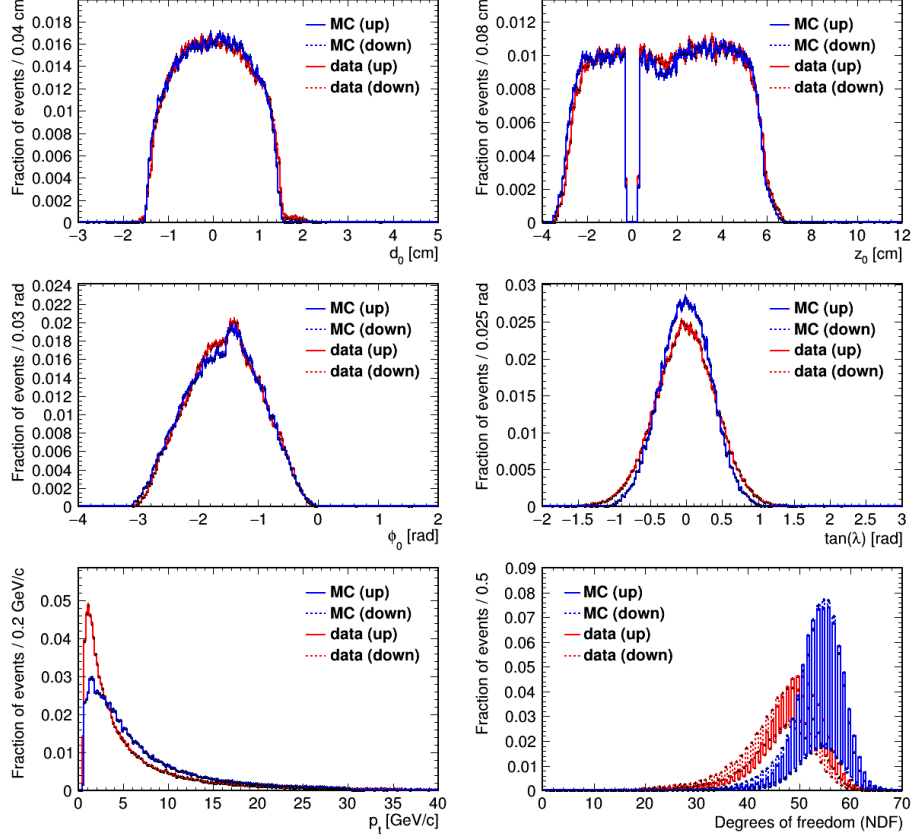
Figure 17.1: Distributions of the five helix parameters (transverse momentum p_t is shown instead of ω) and the number of degrees of freedom (number of 1D measurements minus five) for cosmic validation data sample for data (red) and MC simulation (blue). The parameters of the lower and upper arms of the split cosmic tracks are shown separately.

can be used to estimate the associated detector resolution and bias as

$$
\begin{aligned}
\text{resolution}^2 &= \text{Var}[\Delta h/\sqrt{2}] \\
\text{bias} &= \text{mean}[\Delta h/\sqrt{2}].
\end{aligned}
\tag{17.2}
$$

We do not explicitly show the bias in the following sections, but Δh, to make the discrepancies better visible with respect to the resolution. In contrast to the resolution, here, the factor of $1/\sqrt{2}$ is used by convention. In this section, we estimate the mean and variance from a Gaussian fit to the distribution of $\Delta h/\sqrt{2}$ within a range containing 90% of events, symmetric around its mean.

The resolution for the five helix parameters for simulated tracks is shown in Fig. 17.2. Notable are the small biases in vertex parameters, namely for d_0, where the bias of -0.7 μm is significant and corresponds to about 8% of the 9 μm resolution. We did not study the origin of such a bias but can observe a similar effect in $e^+e^- \rightarrow \mu^+\mu^-$ validation, and it will be visible also in the following section. This illustrates that even the MC reconstruction is not perfect and effects below the 1 μm level are irrelevant.

The helix parameter differences for the prompt alignment are shown in Fig. 17.3. The resolutions are already close to the MC expectations. Only a small 1.2 μm
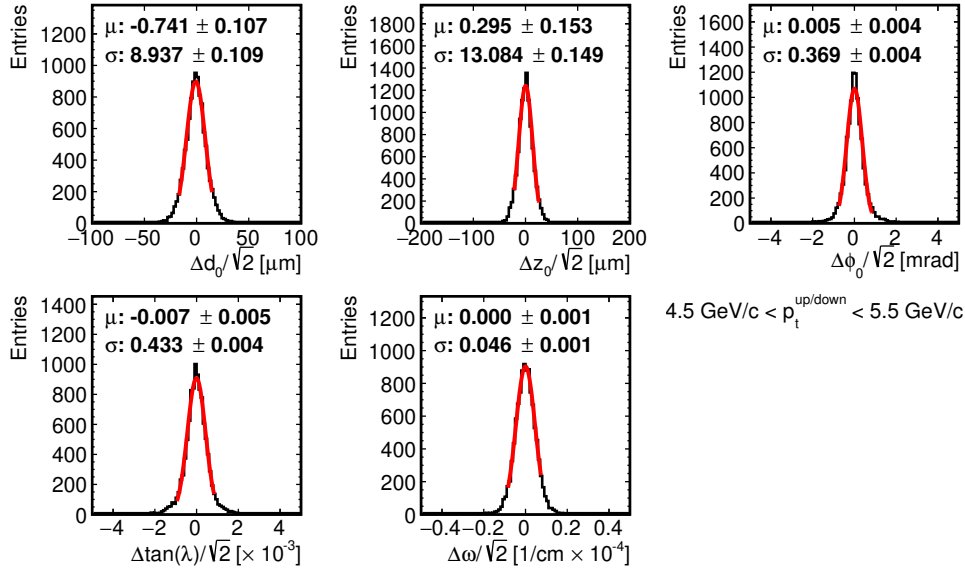
Figure 17.2: Resolution in the five helix parameters for the MC cosmic data sample without any misalignment. Only tracks with $4.5\ \text{GeV}/c < p_t < 5.5\ \text{GeV}/c$ are used. The red lines are results of a Gaussian fit to a range containing 90% of the events. The denoted parameters μ and σ correspond to the mean and the width of the fit function, with their estimated statistical uncertainties.
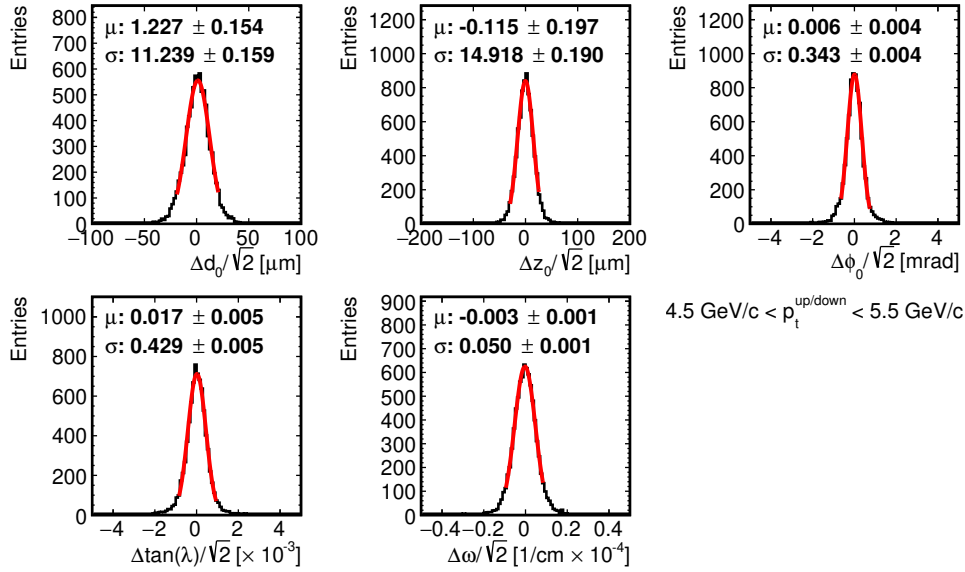
Figure 17.3: Resolution in the five helix parameters for the data of experiment 12 with prompt alignment. Only tracks with $4.5\ \text{GeV}/c < p_t < 5.5\ \text{GeV}/c$ are used. The red lines are results of a Gaussian fit to a range containing 90% of the events. The denoted parameters μ and σ correspond to the mean and the width of the fit function, with their estimated statistical uncertainties.

bias is visible in d_0, whose resolution is about 2 μm worse than for MC, similarly as for z_0. This corresponds to about 26% and 14% worse data resolution for d_0 and z_0, respectively. This would already not meet the original precision requirements—that data is not more than 20% worse than MC. For the helix

angles ϕ_0 and $\tan \lambda$, the data resolutions seem actually even slightly better than in simulation at our reference momentum of 5 GeV/c, which is only a result of an imperfect fit of the MC distributions, while the resolution of the track curvature is about 8% worse. CDC calibrations and the gas composition drive the worse momentum resolution. Further improvements in the momentum resolution do not seem to be achievable through a better alignment.
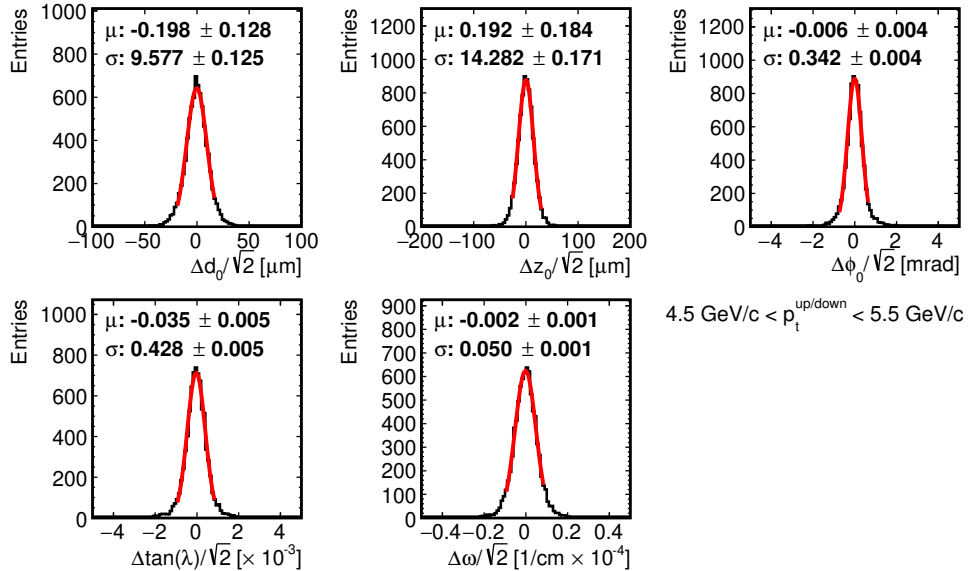
Figure 17.4: Resolution in the five helix parameters for the data of experiment 12 with reprocessing (`proc12`) alignment. Only tracks with 4.5 GeV/c < p_t < 5.5 GeV/c are used. The red lines are results of a Gaussian fit to a range containing 90% of the events. The denoted parameters μ and σ correspond to the mean and the width of the fit function, with their estimated statistical uncertainties.

In Fig. 17.4, the resolutions after reprocessing are shown. The d_0 bias is removed, and the d_0 and z_0 resolutions for data are now only by 7% and 9% worse than MC, significantly improved by the reprocessing. A more detailed study over a wider momentum range to extract vertex resolutions will be discussed in Sec. 17.1.3. In general, the reprocessing performance is better in all validations. An exception is a bias in $\tan \lambda$, which develops in the reprocessing and is not present in the prompt alignment. This bias is statistically significant but constitutes only about 8% of the actual $\tan \lambda$ resolution at 5 GeV/c, and its potential physics impact for mostly lower momentum tracks (with corresponding worse $\tan \lambda$ resolution) is likely negligible.

The appearance of $\tan \lambda$ bias is believed to be connected to deeper issues and inconsistencies beyond the alignment degrees of freedom. We will return to it in the next section, where the biases will be analyzed as a function of helix parameters, allowing us to observe systematic effects that might be averaged out and thus not visible in the resolution plots.

17.1.2 Helix Parameter Correlations

As any difference between the reconstructed helix parameters should, in principle, only be caused by the detector resolution, it should follow a normal distribution

with a mean consistent with zero. This was more or less confirmed in the previous section. This may hold even despite the presence of a significant misalignment. Any functional dependence of the biases on helix parameters indicates a systematic misalignment or other problem influencing tracks systematically. If uncorrected and significant, an associated systematic error for physics analyses may need to be derived.

The correlations are produced as simple profiles. Each bin shows an average over of the parameter differences. Two additional selections remove far outliers: $|\Delta z_0| < 200$ μm and $|\Delta d_0| < 100$ μm. In the previous section, one cannot see such events by eye, justifying the cut. We also require the transverse momentum of both arms to be higher than 0.6 GeV/c to remove the low momentum tail of the momentum spectrum—present in data but not in MC.
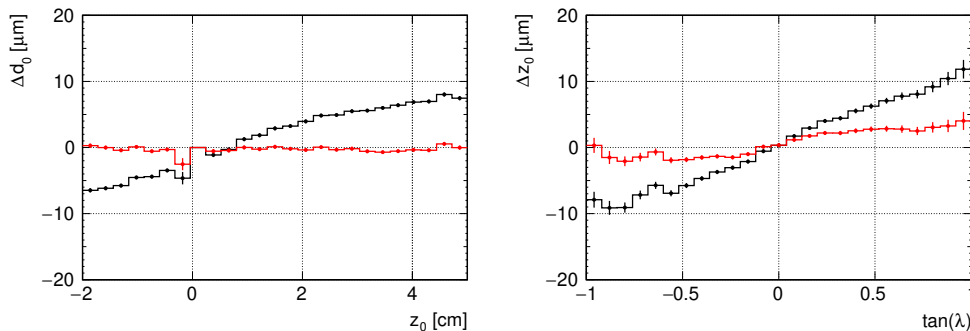
Figure 17.5: The major systematic effects observed in the prompt alignment (black) and in the reprocessing (`proc12`, red) for experiment 12 data using the cosmic validation. The average difference in the estimated d_0 (left) and z_0 (right) of the two arms is averaged in bins of z_0 (left) and $\tan \lambda$ (right). The error bars show the statistical error of the average in each bin. Bins around $z_0 = 0$ are affected by a cut removing tracks from collisions (background for our validation).

In the ideal case, Δh should be consistent with a constant zero function. Fig. 17.5 shows the most striking correlations in the prompt alignment and the improvement with the reprocessing. The dependence of the d_0 bias on track z_0 is completely removed and consistent with a perfect alignment. The correlation of the z_0 bias is significantly reduced, but some remaining effects can be seen, slightly larger for the forward tracks. The local biases at the level of $2 - 3$ μm correspond to at most 20% of z_0 resolution (which gets worse for the forward tracks, and thus the bias gets even less significant) and are expected to have a negligible impact on physics.

Five helix parameters imply 25 corresponding different validations. These plots are compared for prompt alignment and MC in Fig. 17.6. The results after reprocessing are compared with the ideal MC in Fig. 17.7.

Firstly, some imperfections are again observed on MC, indicating what level of biases is likely negligible. The previously reported d_0 bias is with some effort visible in the first row—basically constant as a function of all helix parameters, except ω. Surprisingly, this bias is almost zero for negative and about 2 μm for positive tracks, nearly independent of the actual curvature.

After reprocessing, only small remaining correlations are present in some validations. In general, we observe improvements in all validations, where some issues

are observed in the prompt alignment, which concerns namely the first two rows with vertexing systematics. The small sine-like dependence remaining in the Δz_0 vs. ϕ_0 correlation is reduced, but not entirely. We will return to biases of this type later. This particular dependency is suspected to be a remnant of a typical time-dependent deformation at the SVD sensor level. These improvements are of high importance to reach the ultimate performance mainly for future high precision time-dependent *CP*-violation measurements.

On the other hand, we have found at least two slight degradations. The $\tan\lambda$ bias was already discussed and can be with some difficulty observed also by comparing the third row in Fig. 17.6 and Fig. 17.7. In addition, we observe a slight increase in $\Delta\omega$ vs. ϕ_0 dependence. This could potentially be caused by the wire alignment and global CDC deformation, which cannot be entirely corrected by the alignment procedure. This effect should have only a small physics impact, which can be corrected ad hoc at the analysis level, as was already discussed in Sec. 16.3

17.1.3 Impact Parameter Resolution for Physics-like Tracks

This section will discuss a more elaborate analysis of the resolution in the two most crucial helix parameters, d_0 and z_0. From the 25 subfigures of the previous section, we conclude that the reprocessing alignment should provide unbiased measurements, free of any significant systematic errors. These validations use only limited selections to identify any problems. To evaluate detector performance for physics, one must be stricter and make the cosmic sample as similar as possible to the actual collision tracks, for which the performance should be estimated. We also intend to compare the results to the previously derived performance figures from simulations (see Fig. 3.7), which used a simple particle gun with muons originating from the IP rather than cosmic ray simulation.

Most tracks originate from the inside of the beam pipe. Its material represents the limiting factor for maximum achievable resolution due to multiple scattering and a generally soft spectrum of tracks interesting for physics. Some cosmic tracks hitting the PXD do not cross the beam pipe. For such tracks, the resolution seems significantly better. Thus we require $|d_0| < 1$ cm for both arms, which is the beam pipe radius. Another contamination seems to come from cosmic tracks with a large z_0, which cross the Titanium coating of the beam-pipe and thus are also not representative of a typical physics track. These tracks are removed by imposing -2 cm $< z_0 < 4$ cm for both arms. The reasoning behind these cuts can be best understood when the estimated helix errors from the tracking are investigated. As an example, the reprocessing alignment is used in Fig. 17.8 to compare the estimated d_0 tracking errors as a function of pseudo-momentum. We do not use the tracking errors in our evaluations as there are still some issues, and fine-tuning is needed to reach unit pulls. However, these tracking errors encode information about the material crossed during track extrapolation and thus can be used to identify tracks passing more or less material than expected. These selections leave the distribution without indications of significant contamination by non-standard tracks present initially.

We report the resolution as a function of pseudo-momentum, defined sepa-
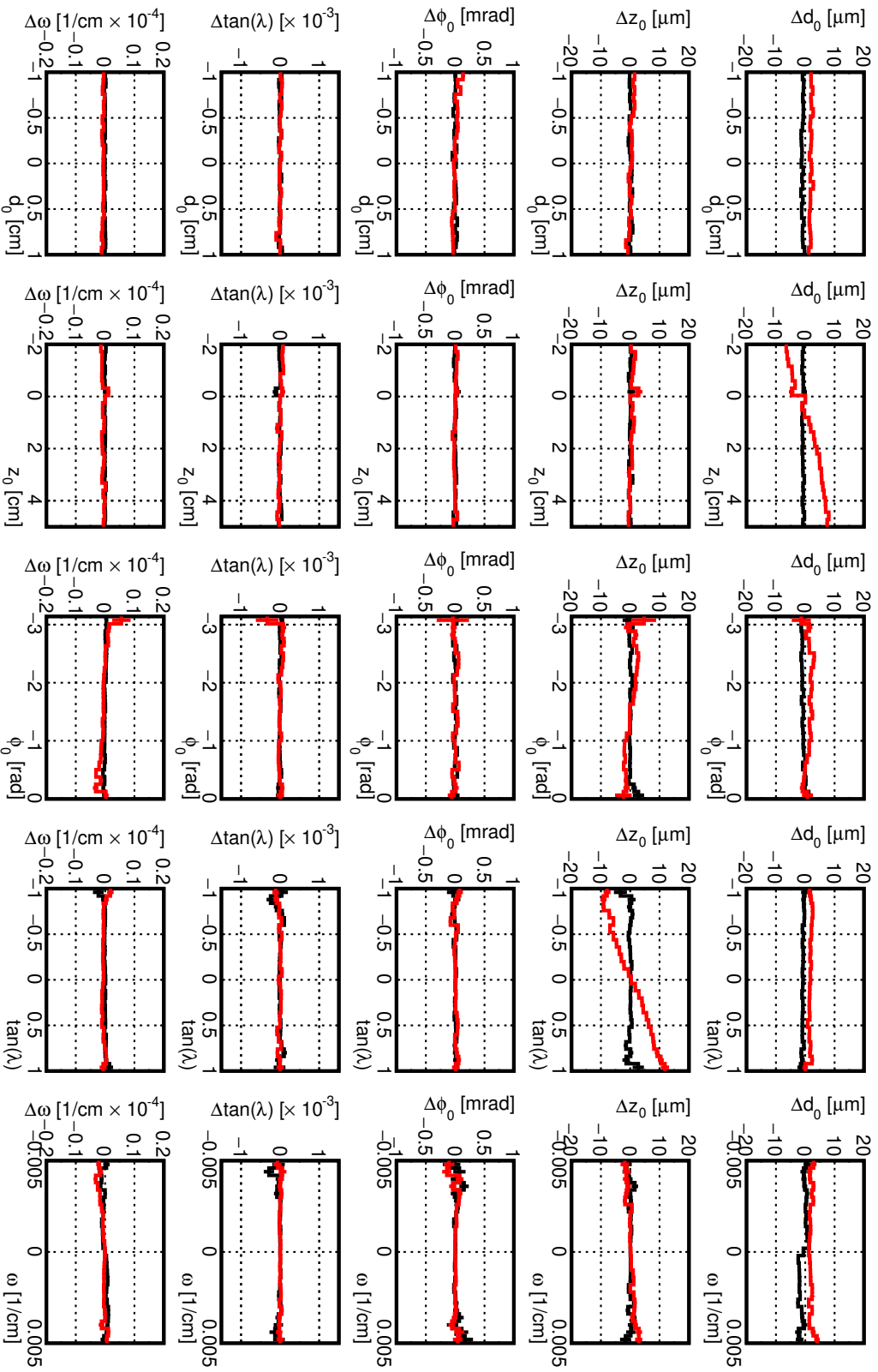
Figure 17.6: Correlations of biases of the helix parameters as functions of the helix parameters compared for MC without misalignment (black) and data of experiment 12 reconstructed with prompt alignment.

304
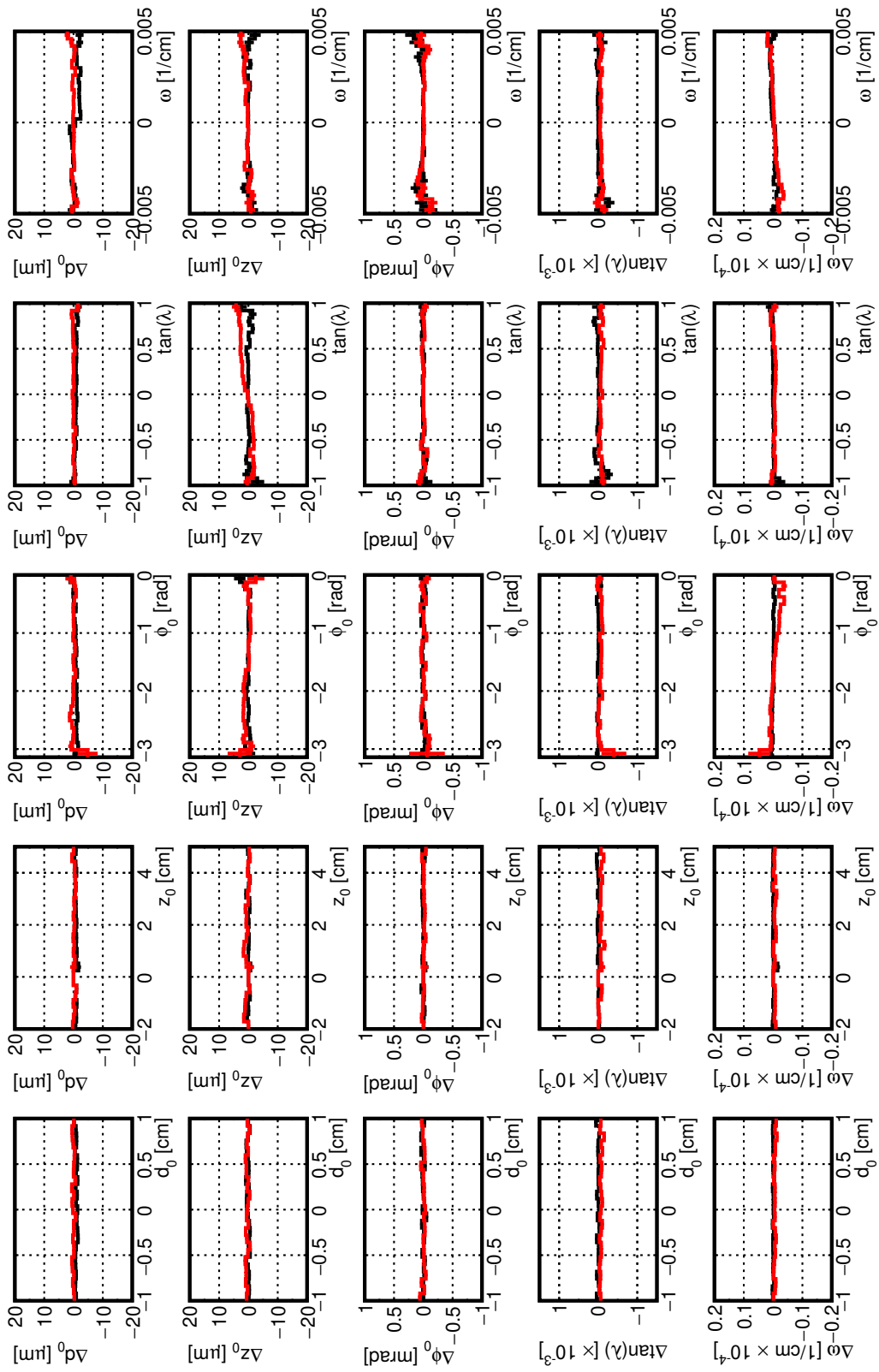
Figure 17.7: Correlations of biases of the helix parameters as functions of the helix parameters compared for MC without misalignment (black) and data of experiment 12 reconstructed with reprocessing (proc12) alignment.
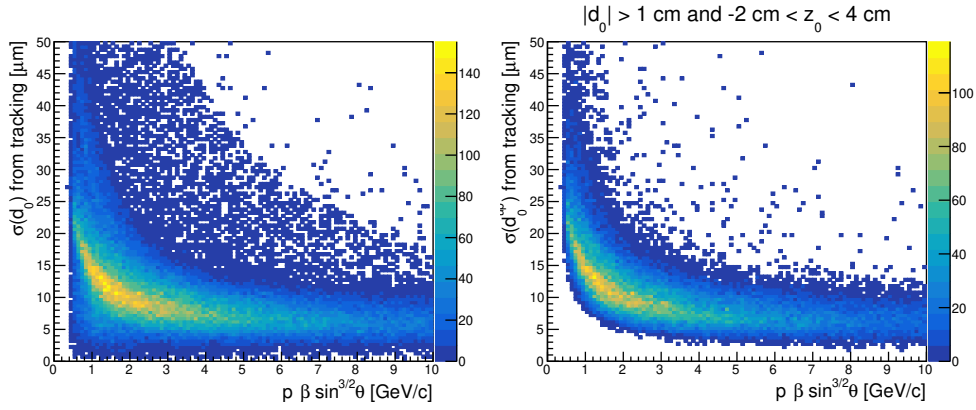
Figure 17.8: Estimated errors from tracking for d_0 for data of experiment 12 with reprocessing alignment (difference to MC or prompt alignment is small) as a function of pseudo-momentum. The distribution before (left) and after cuts (right) to remove tracks not hitting the beam-pipe or hitting its Titanium coating are shown.

rately for the transversal and longitudinal projection as

$$
\begin{aligned}
\tilde{p}_{d_0} &= p\beta \sin(\theta)^{3/2} \\
\tilde{p}_{z_0} &= p\beta \sin(\theta)^{5/2},
\end{aligned}
\tag{17.3}
$$

which reflect the geometrical dependence of the amount of crossed material (and level arm) as a function of track polar angle[2] and can be used to decouple this effect for a more elaborate resolution estimate. By using the pseudo-momenta, the resolutions can be parametrized as

$$
\sigma^2(\tilde{p}) = a^2 + \frac{b^2}{\tilde{p}^2},
\tag{17.4}
$$

where a stands for the detector resolution and b parametrizes the material budget affecting smearing by multiple scattering, which is proportional to the inverse of the particle momenta. Thus in the high momentum limit, the multiple scattering term gets negligible, and the resolution reaches a plateau, corresponding to the detector resolution.

The width of the $\Delta d_0/\sqrt{2}$ and $\Delta z_0/\sqrt{2}$ distributions for tracks passing all selections is evaluated with the σ_{68} method per each bin of the pseudo-momentum: a symmetric range around the mean of the distribution is built such that 68% of events lie within it. Half of this symmetric range length is then reported as a resolution estimate. This method is more robust against outliers and has become a standard in the Belle II Collaboration. The error for the estimate in each bin is estimated as $\sigma_{68}(\Delta h/\sqrt{2})/\sqrt{N}$, where N is the number of events in the corresponding momentum bin.

In Fig. 17.9, Eq. 17.4 is used to extract the resolution from a fit to points in separate pseudo-momentum bins. The ideal MC case is compared to the prompt and reprocessing alignment. The multiple scattering factors b agree in data and

---

[2]This parametrization assumes that the material can be described as a concentric tube, which is not exactly the case, especially with slanted SVD sensors.
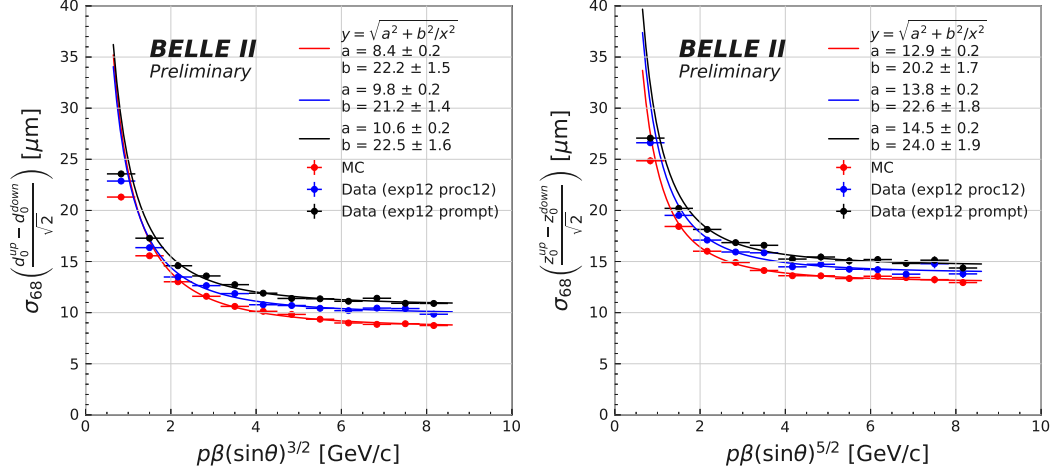
Figure 17.9: Estimated resolution in the cosmic validation for d_0 (left) and z_0 (right) in bins of pseudo-momentum for MC without misalignment (red) and experiment 12 data with prompt (black) and reprocessing (blue) alignment. Vertical error bars are plotted but not visible. The lines are the results of χ^2 fits—the estimated parameters of the fits and their statistical errors are shown.

MC. This means that the material distribution is well parametrized in the simulation.

Regarding the estimated resolutions, a significant improvement is achieved by the reprocessing. In general, the observed reprocessed data resolutions are approximately one micrometer larger than for the ideal simulation. The situation is slightly worse for d_0, where data are about 17% worse than expected. This was already observed in previous analyses in the Collaboration and is at least partially explained by a worse intrinsic resolution of the SVD sensors. A coordinated effort of the SVD group to make simulation closer to reality is still in progress. Thus it is likely that alignment is not responsible for the entire discrepancy. Such a discrepancy is still acceptable, especially at this early stage of the experiment. We arrived at a different relative degradation estimate when the resolution was evaluated at 5 GeV/c reference momentum. All these methods have some systematic errors (correlated across different data samples) that result in wide variations when differences in parameters are compared and expressed relative to the MC estimate.

The estimated reprocessing z_0 resolution is already very close to the ideal MC and improves significantly from the prompt alignment. Here we suspect the alignment to further reduce time-dependent variations in the future, but any such improvement will have only a minimal physics impact. This is also true for d_0, as most physics analyses use tracks with momentum smaller than 2 GeV/c.

Statistical and Systematic Errors

To estimate the statistical error of the determined impact parameter resolutions, we generate 100 bootstrap replicas (with replacement) of the data set and retrieve uncertainties from the standard deviation of the distributions of the parameters a and b.

For systematic uncertainties, we consider the following sources:

- The number of pseudo-momentum bins in the resolution plots, which is varied from 12 to 10, 20, and 30.

- The maximum range of pseudo-momentum, which is varied from 8.5 GeV/c to 10, 15, and 20 GeV/c.

- The number of bins for resolution per each pseudo-momentum bin, varied from 24×10^3 to 10×10^3, 12×10^3, and 48×10^3.

- The range used to evaluate the resolution in each pseudo-momentum bin, varied from 600 μm to 200 μm, 400 μm, and 1000 μm.

The last two items estimate systematics stemming from the σ_{68} method, which is evaluated on histograms instead of raw arrays of data. However, the largest systematic uncertainty for the a term stems from the pseudo-momentum range (wider ranges give slightly better resolutions). The b term is sensitive to the binning in pseudo-momentum (finer binning gives larger estimates of b). We use the maximum observed deviation from the nominal fit value in each category and sum the contributions in quadrature.

In summary, we obtain as the detector impact parameter resolution the following values for experiment 12 data after the (`proc12`) reprocessing

$$\sigma_{d_0} \text{ (proc12)} = (9.8 \pm 0.1 \pm 0.7) \ \mu\text{m} \oplus (21.2 \pm 0.6 \pm 2.3) \ \mu\text{m} \cdot \text{GeV}/c/\tilde{p}_{d_0}$$
$$\sigma_{z_0} \text{ (proc12)} = (13.8 \pm 0.2 \pm 0.4) \ \mu\text{m} \oplus (22.6 \pm 1.1 \pm 7.1) \ \mu\text{m} \cdot \text{GeV}/c/\tilde{p}_{z_0},$$
$$(17.5)$$

where the first uncertainty is statistical (from the bootstrap method) and the second systematic. For prompt calibration of the same data, we get

$$\sigma_{d_0} \text{ (prompt)} = (10.6 \pm 0.1 \pm 0.6) \ \mu\text{m} \oplus (22.5 \pm 0.8 \pm 3.4) \ \mu\text{m} \cdot \text{GeV}/c/\tilde{p}_{d_0}$$
$$\sigma_{z_0} \text{ (prompt)} = (14.5 \pm 0.2 \pm 0.4) \ \mu\text{m} \oplus (24.0 \pm 1.2 \pm 6.0) \ \mu\text{m} \cdot \text{GeV}/c/\tilde{p}_{z_0}$$
$$(17.6)$$

and finally, for MC simulation, the estimates yield

$$\sigma_{d_0} \text{ (MC)} = (8.4 \pm 0.1 \pm 0.5) \ \mu\text{m} \oplus (22.2 \pm 0.6 \pm 1.8) \ \mu\text{m} \cdot \text{GeV}/c/\tilde{p}_{d_0}$$
$$\sigma_{z_0} \text{ (MC)} = (12.9 \pm 0.1 \pm 0.3) \ \mu\text{m} \oplus (20.2 \pm 0.8 \pm 3.2) \ \mu\text{m} \cdot \text{GeV}/c/\tilde{p}_{z_0}.$$
$$(17.7)$$

This means the data/MC disagreement improves from $(26 \pm 10)\%$ to $(17 \pm 11)\%$ and from $(13 \pm 4)\%$ to $(7 \pm 4)\%$ for d_0 and z_0, respectively, due to the reprocessing alignment.

The method is not suitable for estimation of the b parameter, which depends on a steep slope at low pseudo-momenta and would deserve, among other things, a finer binning in this region. However, as the alignment practically contributes only to the a term, this method is sufficient for our purposes.

17.2 Validation with $e^+e^- \rightarrow \mu^+\mu^-$ Events

Di-muon events have two important advantages over the cosmic rays. They cover the full angular acceptance and are produced at a much higher rate, which will

further increase by more than an order of magnitude when the full luminosity is reached. We have already utilized this sample for the final run-dependent stage of the reprocessing alignment. With higher luminosity, the time resolution could be further increased for the price of higher computing resources dedicated to the alignment, possibly executed online in the future.

As the dedicated tight di-muon skim already represents a clean sample, we do not impose any further selections besides the following:

1. Exactly two charged tracks reconstructed in an event.

2. For both tracks $p > 1.0$ GeV$/c$, $|dz| < 0.2$ cm, and $dr < 0.5$ cm, where dz and dr are longitudinal and transverse impact parameters relative to the IP position.

3. For the reconstructed invariant mass of the pair to fall in $[9.5, 11]$ GeV$/c^2$.

In contrast to the previous chapter, no background is simulated for MC, which we use for comparison. We expect its effect to be small, if not negligible, for the figures in this section. The MC was, however, simulated realistically with most missing and dead channels. We simulate the MC only in a particular run range for the comparison.

Because this validation was produced right after the alignment finished, the IP calibration was not yet recalculated after the new alignment was derived. While run-dependent constant offsets are determined in the final reprocessing alignment stage, these are discarded for the beam spot payloads. The reason is a missing implementation to sum the corrections with intra-run-dependent IP calibration. The IP calibration procedure should be repeated with the final alignment in any case. This, however, means that we must determine the beam spot position during the validation procedure to perform a pivot transformation of the helix parameter of the two muons before their comparison. This is necessary because the default pivot for the helix in the Belle II tracking is the origin of the coordinates, while the actual IP position in our global reference frame is displaced from the origin by several hundreds of micrometers. It then further oscillates around this position and can also experience larger shifts when adjusted by the accelerator operators, for example, to find a *sweet spot* for the highest luminosity.

We do not attempt to determine a fine intra-run dependent beam spot position but evaluate its mean in the x, y, and z coordinates as a median of the corresponding distributions after a vertex fit. We do so for all events in each data file[3]. With the average IP coordinates, the pivot transformation is performed for the helix parameters of both muons. Only the transformed helix parameters are used in all shown plots.

In contrast to correctly oriented cosmic ray tracks, the two muons have, by definition, an opposite sign of d_0, and thus, we are interested in the distribution of the quantities

$$\Sigma d_0 = d_0^+ + d_0^-$$
$$\Delta z_0 = z_0^+ - z_0^-,$$

(17.8)

where the sum and difference are taken between the helix parameters of the positive and negative muon.

---

[3]On average, each data file contains 2000 to 6000 events, but much smaller files are possible.

17.2.1 Impact Parameter Resolutions

The resolutions in d_0 and z_0 are evaluated using the σ_{68} method in bins of the positive track's azimuthal angle ϕ_0 and polar angle, expressed as $\tan\lambda$. The distributions for prompt and reprocessing (denoted as `proc12`) alignment are compared to the ideal MC and MC with the prompt misalignment model (see Sec. 13.6.3). A range of muon momenta is summed together in these validations (where the high momentum muons dominate), so the estimated resolutions are slightly worse numerically than, for example, for the reference momentum used in the cosmic validation.
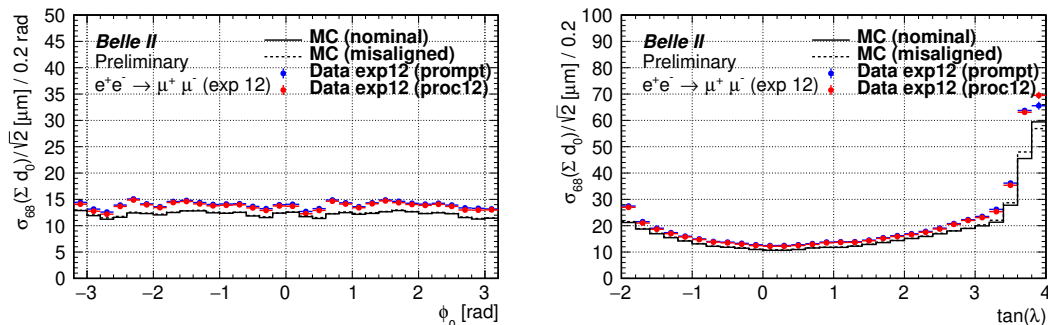
Figure 17.10: Estimated resolution in d_0 for the di-muon sample as a function of positive track ϕ_0 (left) and $\tan\lambda$ (right). The resolution estimated with σ_{68} method is compared for MC without misalignment (solid black line), MC with the model of prompt misalignment (dashed black line), and data of experiment 12 with prompt (blue) and reprocessing (red) alignment. Error bars correspond to the errors of the standard deviation for values in each bin. For clarity, the error bars are only shown for data; they are similar for MC.

In Fig. 17.10, the d_0 resolution is shown. Only a minimal improvement is achieved by the reprocessing. We see a slight general offset from the expected MC resolution, larger at large polar angles, where the resolution gets quickly worse (as expected). This might be caused by residual sensor deformations beyond the current parametrization or a larger influence of the SVD, which has probably significantly worse intrinsic resolution than expected on MC in this direction (discussed later). In the ϕ_0 dependence, one can also see that some subtle patterns from MC are reproduced in data. Similar conclusions can be drawn for the z_0 resolution, shown in Fig. 17.11. Here, the improvement from prompt alignment can be clearly seen.

The very small remaining discrepancies can be caused by an average uncorrected random time-dependent misalignment or some (possibly) constant systematic effect. For example in the ϕ_0-dependent resolution (see left in Fig. 17.10 for d_0 and Fig. 17.11 for z_0), the best data/MC correspondence seems to occur near $\phi_0 = 0 \pm \pi$. The larger difference in other regions, especially for z_0 resolution, seems to be caused by specific deformation effects evolving in time. This will get more evident in the next sections.
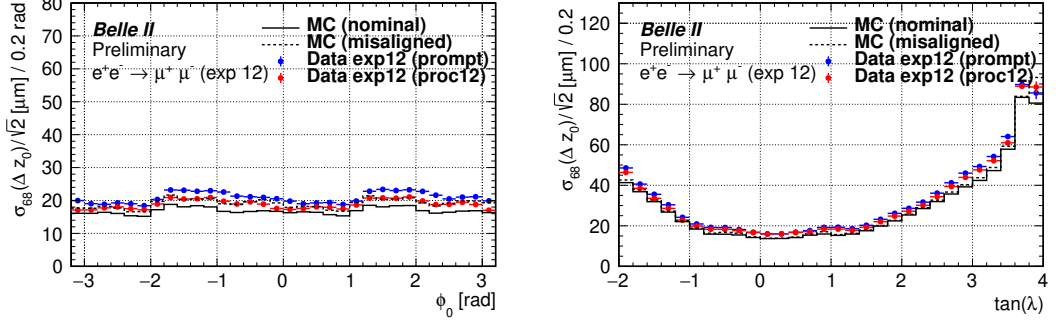
Figure 17.11: Estimated resolution in z_0 for the di-muon sample as a function of positive track ϕ_0 (left) and $\tan\lambda$. The resolution estimated with σ_{68} method is compared for MC without misalignment (solid black line), MC with the model of prompt alignment (dashed black line), and data of experiment 12 with prompt (blue) and reprocessing (red) alignment. Error bars are errors of the standard deviation for values in each bin. For clarity, the error bars are only shown for data; they are similar for MC.

17.2.2 Impact Parameter Biases

The vertex parameter biases provide a more robust look into the possible alignment problems, which easily get averaged out over tracks going in various directions. In Fig. 17.12, the d_0 bias is shown as a function of the positive muon ϕ_0 and $\tan\lambda$. A slight positive bias is again present in the MC, similarly to the cosmic ray validation. Whether the origin of this bias is the same is not currently known. We use it as an argument to neglect biases of a similar scale observed on data.
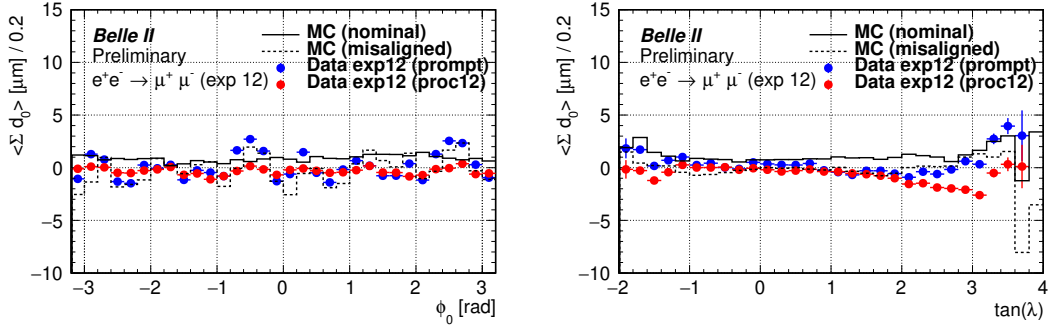
Figure 17.12: Estimated biases in d_0 for the di-muon sample as a function of positive track ϕ_0 (left) and $\tan\lambda$. The biases are compared for MC without misalignment (solid black line), MC with the model of prompt alignment (dashed black line), and data of experiment 12 with prompt (blue) and reprocessing (red) alignment. Error bars are errors of the means in each bin. For clarity, the error bars are only shown for data; they are similar for MC.

Regarding the Σd_0 vs. ϕ_0 dependence, some residual misalignment with a structure resembling the PXD ladders can be observed in the prompt alignment. After reprocessing, this effect is suppressed well under one micrometer on average. Some slight dependence can be observed for $\tan\lambda$ in both MC and the reprocessing, albeit with an opposite sign. The prompt alignment seems surpris-
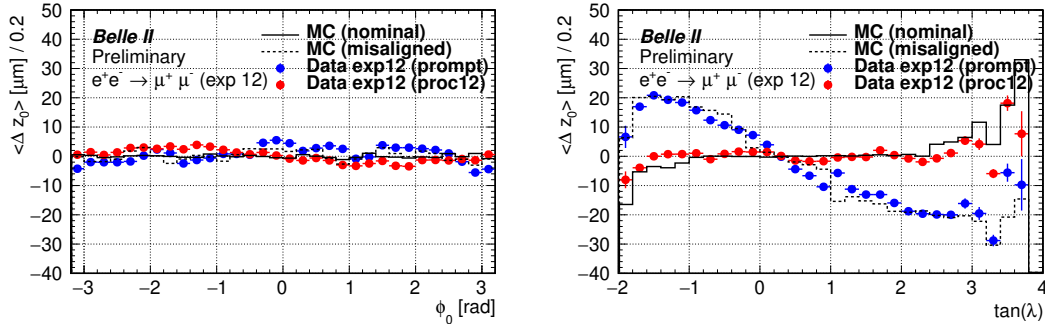
Figure 17.13: Estimated biases in z_0 for the di-muon sample as a function of positive track ϕ_0 (left) and $\tan\lambda$. The biases are compared for MC without misalignment (solid black line), MC with the model of prompt alignment (dashed black line), and data of experiment 12 with prompt (blue) and reprocessing (red) alignment. Error bars are errors of the means in each bin. For clarity, the error bars are only shown for data; they are similar for MC.

ingly better than the reprocessing, but only in the regions where the resolution is already worse, see Fig. 17.10 (right).

The z_0 biases, shown in Fig. 17.13 are more interesting. The large slope in the right figure is entirely removed by the reprocessing. The local biases have a similar scale to the ideal MC. In addition, the prompt result can be accurately reproduced by the prompt misalignment model. This slope is likely related to the corresponding dependence in the cosmic ray validation. In the ϕ_0 dependence (Fig. 17.13 left), a small oscillation is visible with $\sim 3~\mu$m amplitude and nodes of zero bias at $\phi_0 = 0 \pm \pi$, corresponding to the slight variations of z_0 resolutions. This suggests that some specific deformation remains on average, which could not be corrected by the time-dependent reprocessing alignment.

In conclusion, the vertex parameter bias study shows that the average final performance is good, sometimes appearing even less biased than the nominal MC. No serious issues have been identified, even at this more detailed level. To observe the remaining inconsistencies in detail, time-dependent validation must be utilized.

17.2.3 Run Dependence of Average Biases and Resolutions

The time dependence of the average d_0 and z_0 biases is shown in Fig. 17.14. We evaluate the bias per each data file and plot them binned every ten runs. Instead of the time axis, run numbers are used. One can still see gaps usually corresponding to maintenance days, where non-physics runs are taken. Two significant shifts up to 4 μm occur for the d_0 bias, which is recovered in the reprocessing, whose points are generally closer to and consistent with zero within 1–2 μm. For z_0, the improvement is much less visible. The luminosity was lower for runs < 2000, which results in larger fluctuations and statistical errors (taken as errors of the mean). The MC data are generated only for a subset of runs for a reference, shown as black points in the plots.

In Fig. 17.15, the run-dependence of the resolutions is shown. A small im-
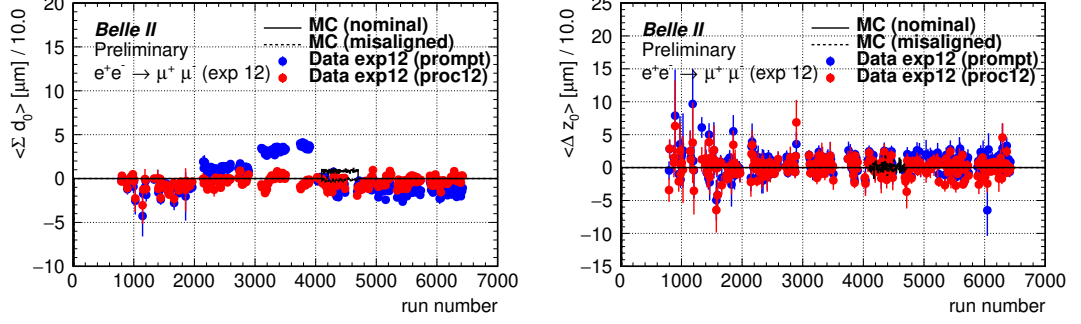
Figure 17.14: Estimated track parameter biases in d_0 (left) and z_0 (right) for the di-muon sample as a function of the run number of experiment 12. Estimated bias values are averaged in bins of ten runs—error bars show errors of the averages. Biases are compared for MC without misalignment (solid black line), MC with the model of prompt alignment (dashed black line), and the data of experiment 12 with prompt (blue) and reprocessing (red) alignment.

provement by about 0.5 μm, consistent over the entire experiment period, can be observed for the d_0 resolution. Much larger effects can be seen in the z_0 resolution. The resolution in the prompt alignment gets significantly degraded after run 3000. The fact that the repeated bucket-by-bucket baseline alignment does not remove this effect suggests that it could be related to an evolving CDC deformation at the wire level, which is only corrected in the reprocessing.
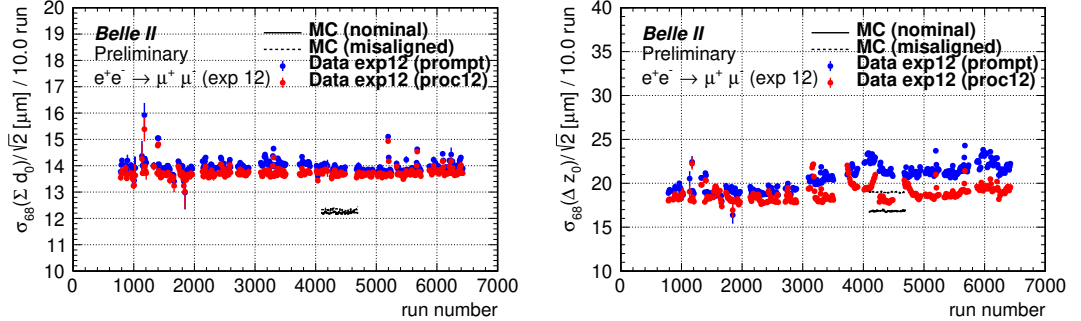
Figure 17.15: Estimated track parameter resolutions in d_0 (left) and z_0 (right) for the di-muon sample as a function of the run number of experiment 12. Estimated resolution values are averaged in bins of ten runs—error bars show errors of the averages. Resolutions are compared for MC without misalignment (solid black line), MC with the model of prompt alignment (dashed black line), and the data of experiment 12 with prompt (blue) and reprocessing (red) alignment.

A significant improvement is achieved for the z_0 resolution in the reprocessing. But for some periods, we see a gradual increase or decrease of the resolution by up to about 5 μm. Such periods usually directly follow after a gap in the runs, meaning a maintenance day or some major event, like QCS magnet quench, followed by a longer break in the physics data-taking, happened in the gap. In addition, one can observe a slight continual worsening of the best achieved reprocessing z_0 resolution towards the end of the data-taking period. This degradation may be related to increasing backgrounds at higher luminosity.

Note that the run-dependent alignment is performed for pixel sensors, VXD half-shells, and CDC layers, but not for the SVD sensors, which are still aligned only once per bucket, even in the reprocessing. It seems that these uncorrected evolving degradations are caused by systematic biases stemming from a continuous deformation of the SVD. Whether this is directly connected to the CDC deformation is not yet clear. So far, we have had difficulties decoupling the different deformations and correlating them to machine and detector events, except for a couple of major ones. From the evolution of the z_0 resolution in the prompt and reprocessing alignment, it is easy to see why it is so. Instead of the continuous increase seen in the reprocessing in the problematic periods, the prompt alignment shows a step-like behavior. The problem is that more severe misalignments occur, which make more subtle effects challenging to resolve. Only after the reprocessing of the complete experiment 12 and this validation have we started to suspect time-dependent systematic deformation in the SVD inside the buckets to be non-negligible for the overall vertex performance. This was not obvious until the majority of other misalignment effects were removed by the reprocessing.

17.2.4 Investigation of the Remaining Time Dependence

Let us briefly look at an example of movements determined by the run-dependent alignment. We do not have a convenient set of degrees of freedom to describe the whole VXD displacement as a rigid body, but we construct such parameters by averaging alignment corrections to all four half-shells. To a first approximation, these move together, and in addition, this movement is correlated to deformations of the CDC. Such a situation was completely unexpected before the first data and when the alignment parametrization was defined.
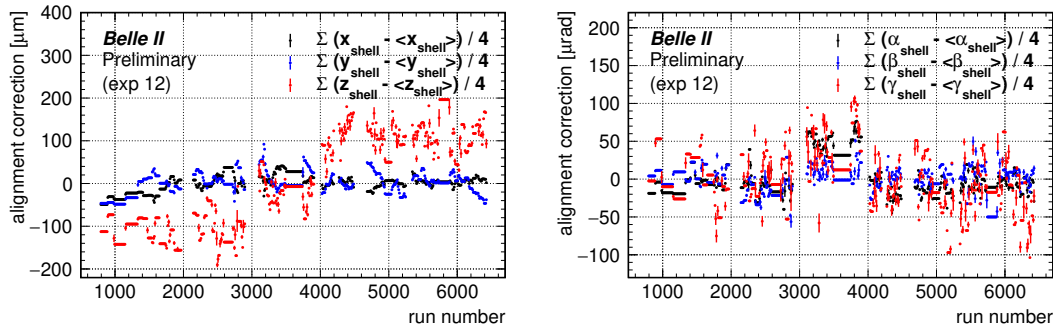
Figure 17.16: Averaged reprocessing alignment corrections in experiment 12 as a function of run number for shifts (left) and rotations (right), showing estimated coherent movements of all VXD half-shells with respect to the run-independent prompt alignment. Corrections are averaged in bins of ten runs. The error of the average is shown as an error bar.

For each half-shell parameter, the mean over all corrections determined in experiment 12 is first subtracted. An average over the four half-shells is then taken. Its mean is shown in Fig. 17.16, binned every ten runs. The short horizontal lines of connected points mean that the payload IOVs were extended, either (1) due to a break in the data taking (without bucket boundary) or (2) because multiple runs were merged to reach sufficient statistics. Case (2) happens mostly for the

initial runs < 2000, and case (1) is typical for the higher run numbers. A gap without points indicates a bucket boundary. One can identify seven original buckets in these figures. In practice, the gap is filled with the last available alignment corrections.

The angle corrections show no apparent trends, except for the step for runs 3000–4000. This will be our candidate serving as an example of a problematic run range. Naturally, the precision for angle γ is lowest due to the shortest corresponding level arm (VXD radius vs. length).
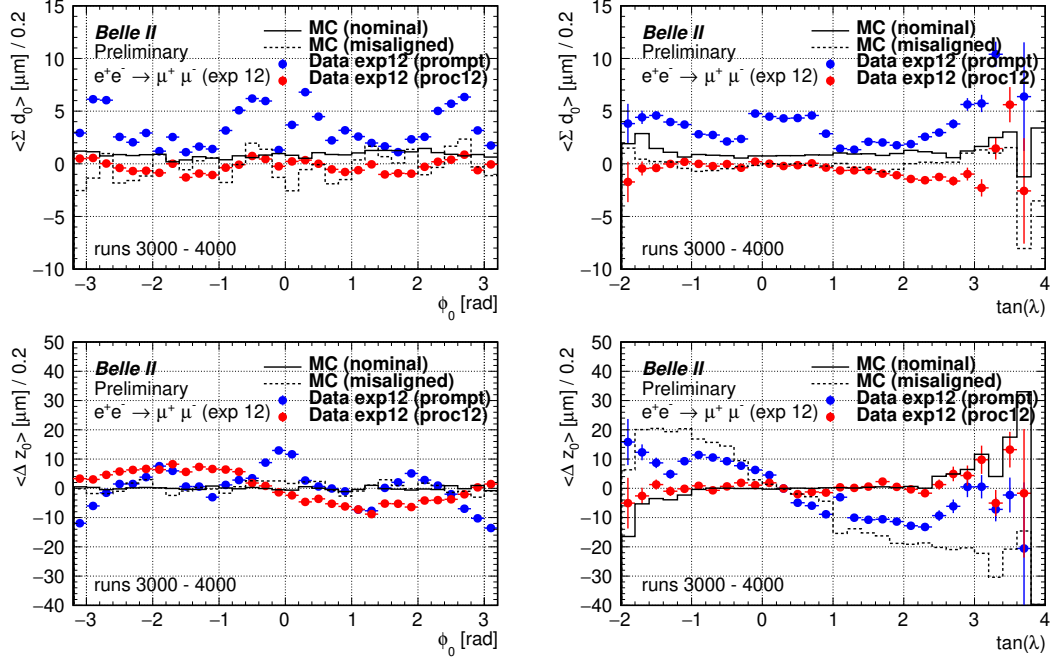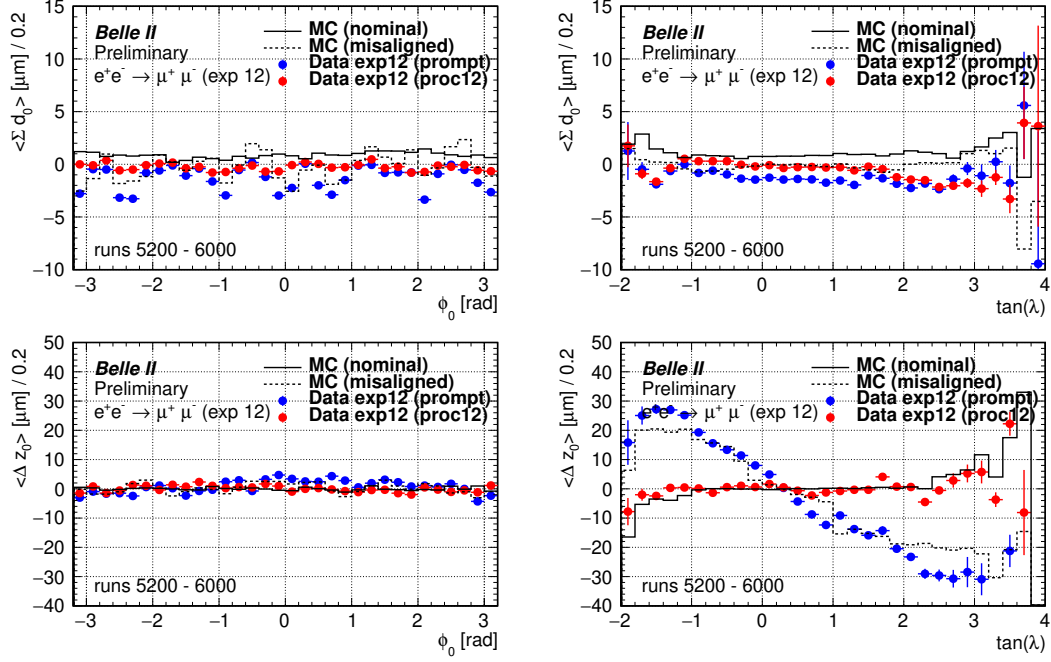

Figure 17.17: Estimated biases in d_0 (top) and z_0 (bottom) for the di-muon sample in run range 3000–4000 of experiment 12 as a function of positive track ϕ_0 (left) and $\tan\lambda$ (right). The biases are compared for MC without misalignment (solid black line), MC with the model of prompt alignment (dashed black line), and the data of experiment 12 with prompt (blue) and reprocessing (red) alignment. Error bars are errors of the means in each bin. For clarity, the error bars are only shown for data; they are similar for MC.

Regarding the shifts of the half-shells, the most significant deviations are observed for the z-coordinate. To be able to absorb mainly thermal deformations of the accelerator and beam-line structures, to which the vertex detector is attached, flexible connections were designed, providing some freedom in the z-direction. We can observe that during experiment 12, two major steps in the z-position of the VXD occur, from -100 μm to zero and $+100$ μm, where again, the run range 3000–4000 is specific. Otherwise, the VXD z-position is stable within about 50 μm, which could be explained by expansions/contractions in the flexible mounts. In the future, possible correlations to various temperature measurements should be investigated. However, it is notable how large corrections need to be derived by the alignment procedure for this coordinate and how small the effects are on the overall performance. For the prompt alignment, there would be only seven constant lines for each bucket in this figure, completely ignoring the additional variations. However, as shown in the previous sections, the

average performance is already good. This demonstrates how unimportant the correlated global movements of the VXD and CDC layers (almost identical to our typical weak modes) are in practice. From MC alignment studies, we estimate that the sensitivity to (systematic residual misalignment in the direction of) such deformations is about an order of magnitude better (smaller) than the observed corrections in reality.

The CDC deformation seems to happen mainly in the vertical direction, and the VXD moves mostly coherently with the first CDC layer. This can be seen in several slopes of the blue points in Fig. 17.16 (left), usually after a gap (maintenance day). It is now almost certain that this deformation mainly happens when the superconducting accelerator optics (QCS magnets) is switched off. The absence of forces from the QCS leads to a quick relaxation. At that point, no physics data are taken. While cosmic data might be recorded after the QCS is de-excited; usually, local runs for internal sub-detector calibrations are taken. The physics data taking is then typically restarted near the evening of each maintenance day. Before this, the QCS is turned on, which restores the magnetic forces. However, it seems that the deformation evolves continuously over several days, which can be seen as the slopes in variations of the y-position of the VXD within about 80 μm range. For the x-direction, the range of movements is about twice smaller in some problematic periods and seems to be correlated to the y-movement.

Most interesting is the apparent correlation of the large continuous y-movements and the change in the z_0 vertex resolution in Fig. 17.15 (right). The CDC layers are aligned at the same frequency as the VXD half-shells and PXD sensors. Thus the remaining uncorrected misalignment must occur in other degrees of freedom. The alignment determines an average over the full bucket for the SVD sensors and CDC wires. Large deviations from the average in the unstable periods then cause systematic effects, which contribute to a worse resolution. This especially happens when multiple periods, including a maintenance day, are merged into one bucket. This was sometimes done to process the data more quickly, but as can be seen, it can lead to significant degradation. By splitting the problematic buckets, one could further suppress this effect.

The validations of vertex biases are shown separately for the most problematic run range (3000–4000) in Fig. 17.17. Large systematic effects can be seen in the prompt alignment, significantly suppressed in the reprocessing. However, the bottom left figure showing z_0 bias as a function of positive track azimuthal angle ϕ_0 exhibits the already well-known (minus) sine-like dependence with about 10 μm amplitude. This represents a typical deformation effect left uncorrected by the reprocessing and is probably the worst scenario as it is observed mainly with one sign and does not entirely average to zero over long periods. Fortunately, these periods of continuous deformations represent only a limited fraction of all data, and, thus, the average effect observed for experiment 12 is well within acceptable limits. It is desirable to limit the number of events leading to change in the deformation (change of the QCS magnetic field) and to split data into buckets using a higher granularity. In such a way, these effects can be suppressed automatically, without resorting to complex alignment strategies that may still not reach optimal splitting of the validity of the alignment constants.

As a demonstration of a period of equivalent length, but without any major issues, we also show the validation for run range 5200–6000 in Fig. 17.18. After

reprocessing, the resulting performance is excellent, without indications of any systematic effects, practically identical to the ideal MC simulation.

Figure 17.18: Estimated biases in d_0 (top) and z_0 (bottom) for the di-muon sample in run range 5200–6000 of experiment 12 as a function of positive track ϕ_0 (left) and $\tan \lambda$ (right). The biases are compared for MC without misalignment (solid black line), MC with the model of prompt alignment (dashed black line), and the data of experiment 12 with prompt (blue) and reprocessing (red) alignment. Error bars are errors of the means in each bin. For clarity, the error bars are only shown for data; they are similar for MC.

The remaining average ϕ_0-dependent z_0 bias in the problematic periods seems to be caused by deformations in the SVD leading to movements of the sensors in the z-direction, but further investigations are needed to understand what happens. If such deformation can be described by a single or just a couple of parameters (which seems, however, currently unlikely), we could add them to the set of alignment parameters and determine them with much lower statistics (and thus higher frequency) than required for a full realignment of the SVD sensor-by-sensor. The PXD is going to be replaced in 2022 or 2023, and hopefully, with redesigned cable layout, the deformations of the CDC should not occur anymore. So it seems premature to re-parametrize the alignment at this point, even when a model of the deformation is derived.

17.3   Validation with Track-To-Hit Residuals

In this section, we return to the very basics of alignment: the residuals. All (new) issues in this chapter are caused by the complex time dependence of the vertex detector misalignment. We performed a number of preliminary studies on a limited dataset (one bucket) before the reprocessing. As will be shown, many effects could be significantly suppressed by the reprocessing, and this is

observed consistently over the whole experiment 12. To limit the number of figures, we often use two representative periods: a short `bucket 11` and a long `bucket 15` (last bucket of experiment 12). The long bucket represents a more typical situation in the current data taking. The short bucket could, on the other hand, represent an ideal situation where the bucket size is limited to only several days.
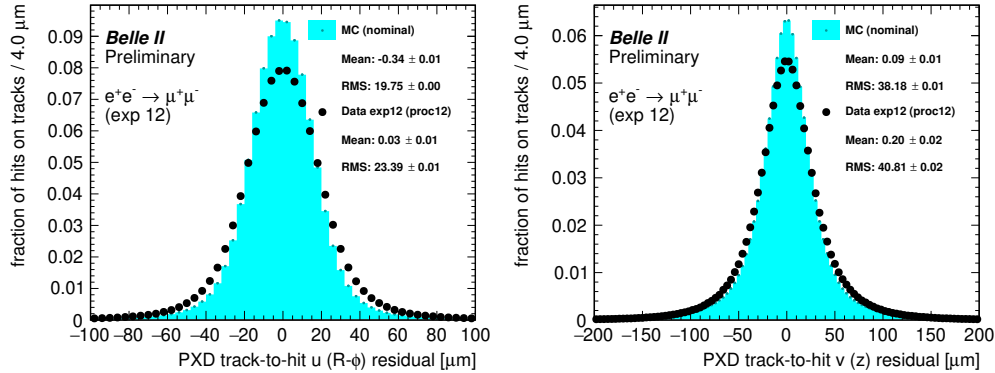
Figure 17.19: Unbiased track-to-hit residuals in u (left) and v (right) from standard tracking (with DAF) integrated over all PXD sensors and all di-muon validation files in experiment 12, using the reprocessing alignment (black points—error bars shown, but not visible). The data distribution is compared to MC simulation without misalignment (cyan).

Before moving to individual buckets, let us briefly review the residual distributions integrated over the entire experiment 12 and all sensors. For pixel sensors, the integrated residuals are shown in Fig. 17.19. For these figures, only the final reprocessing alignment is used. The improvement over the prompt alignment is marginal (usually sub-micron for widths) in such integrated distributions. No particular requirements are applied on the track or hit quality beyond the standard selections of the tight di-muon sample. For MC, no background is simulated, and thus it is possibly slightly optimistic. Nevertheless, very good correspondence is achieved for the v (z) PXD residuals, which are only approximately 5% worse than MC. For u residuals, we again observe worse data/MC correspondence. But we suspect this effect, in fact, to come from a significantly worse intrinsic SVD resolution in the u-coordinate. The residual distributions integrated over all SVD sensors are shown in Fig. 17.20. Whereas the v coordinate is about 20% worse than MC, the u residuals are wider by 50%.

This large discrepancy has two possible reasons: significant misalignment or overestimated MC resolution of the SVD sensors (or issues only on data reconstruction—for example, due to background, which are unlikely to be responsible for the full difference). As all the distributions are nicely centered, and this is also the case for any individual SVD (or PXD) sensor integrated over all data or over any bucket, this could only be caused by time-dependent instabilities (inside buckets). These must be well averaged but large enough to effectively smear the sensor resolutions. As we will see in the next section, we can really observe such instabilities. However, their scale can be quantified and does not explain the discrepancy of the SVD u ($R - \phi$) resolution between data and MC.

318
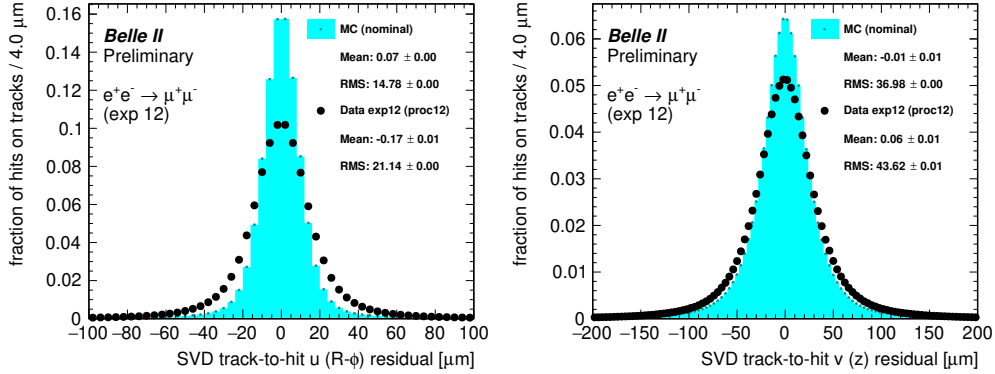
Figure 17.20: Unbiased track-to-hit residuals in u (left) and v (right) from standard tracking (with DAF) integrated over all SVD sensors and all di-muon validation files in experiment 12, using the reprocessing alignment (black points—error bars shown, but not visible). The data distribution is compared to MC simulation without misalignment (cyan).

17.3.1 Time Dependence of PXD Sensor Residuals

As we already suspected that PXD sensors move significantly, every sensor was aligned with the VXD half-shells in each block during reprocessing. With $\sim 50 \times 10^3$ di-muon events, the alignment precision for sensor parameters is already at $1\,\mu$m level or better. Thus we can afford to align PXD sensors too.

Using the two buckets as examples, we have collected the medians of histograms of the track-to-hit residuals per every tight di-muon skim file with about $2000 - 6000$ events. For the u-residuals, the prompt and reprocessing alignment is compared for the two buckets in Fig. 17.21. The blue (prompt) points seem to follow almost linear paths along the three days of the small bucket. The amplitude of this movement can be even larger than $5\,\mu$m, and the forward sensors (number 1) are slightly more affected. Most of these movements are significantly suppressed by the reprocessing, but short periods can be observed for some sensors, where probably due to suboptimal merging, the alignment fails to correct for the movement entirely (for example, the brief peak for sensor 1.7.2). For the long bucket, over two weeks, even larger excursions (over $15\,\mu$m) of the sensors can be observed and are mostly removed, but again with some problematic periods (e.g., for sensor 1.5.1).

The results for the v-coordinate, shown in Fig. 17.22, indicate much more significant instabilities of the PXD sensors. The largest movements can be observed for sensor 2.4.1 in the second (incomplete) PXD layer, where the medians of the residual distribution can be more than $30\,\mu$m away from zero in some periods, and one can definitely observe continuous movements in the prompt alignment. While such movements are largely suppressed by the reprocessing, the alignment is still not fast enough to follow them. Some isolated uncorrected points are likely a result of the sub-optimal merging of smaller runs with a possible time gap, where the large movements happen, but data are not taken. We will attempt to quantify the initial (prompt) and remaining (after reprocessing) residual instabilities in Sec. 17.3.2, but one can draw qualitative conclusions already from these time-dependent validations. Somehow unexpected is the very different behavior
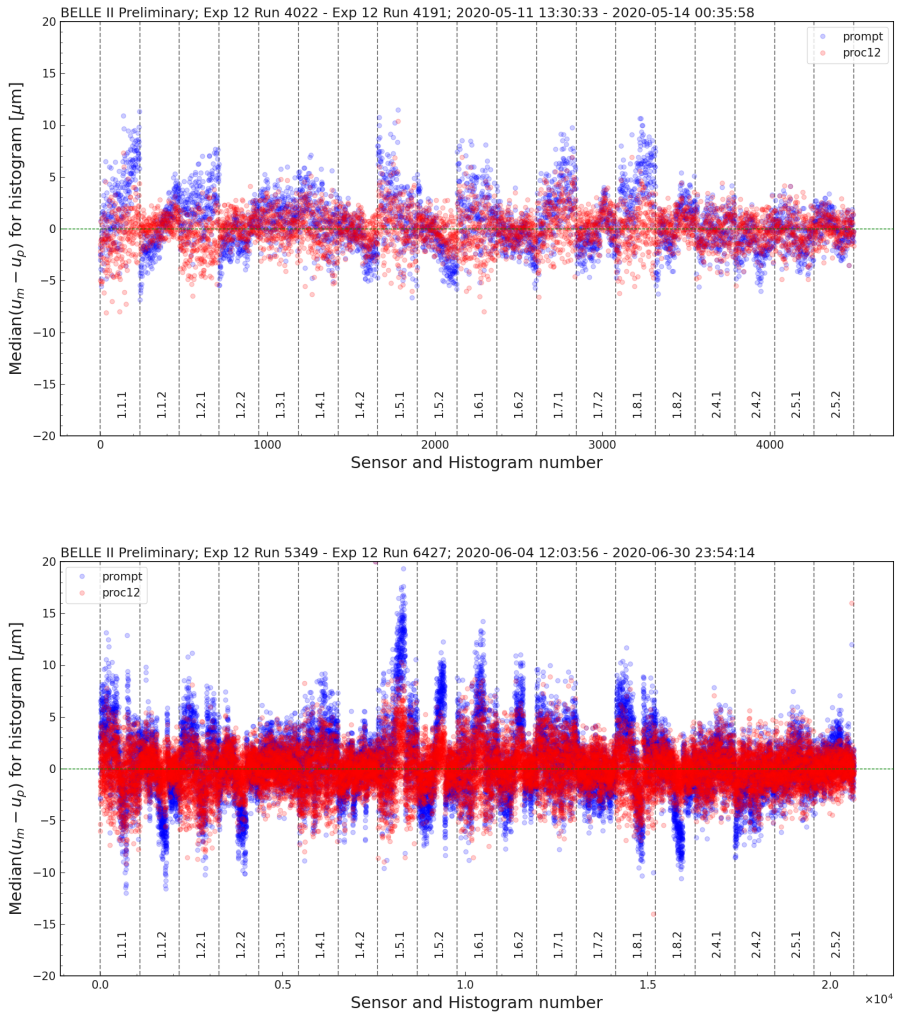
Figure 17.21: Time-dependence of medians of track-to-hit u-residuals in the short `bucket 11` (top) and long `bucket 15` (bottom) of experiment 12 compared for prompt (blue) and reprocessing (red) alignment. Each point corresponds to the median of the histogram in a file with > 2000 di-muon events, error bars not shown. The points are ordered in run number (and file number) for each PXD sensor. The sensor identifiers are shown at the bottom.

of the forward (number 1) and backward (number 2) PXD sensors. Due to lower statistics recorded in the forward sensors (about six times fewer hits than number 2 sensors for the di-muon sample), the spread of the points is naturally larger. But we can only identify continuous movements at our resolution in the forward sensors.

An initial attempt to derive a quantitative estimate of the residual instabilities and their possible physics impact was to just integrate the distributions in the time-dependent plots over all periods (in a bucket) and all PXD sensors. The histograms of all the collected medians for the two example buckets are shown in Fig. 17.23. A significant reduction of the standard deviation of the medians is observed after reprocessing, but as can be seen for the v-coordinate, even after the reprocessing, the RMS is significantly larger for the long bucket (4.74 μm) than
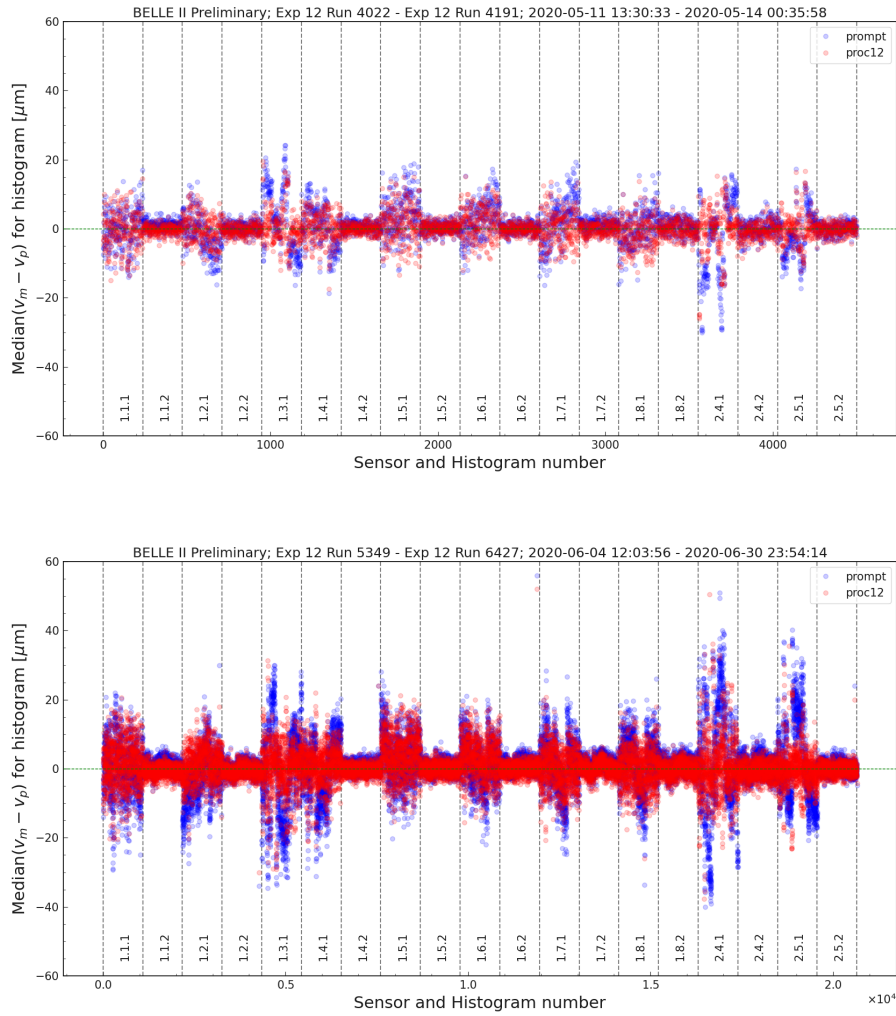
Figure 17.22: Time-dependence of medians of track-to-hit v-residuals in the short `bucket 11` (top) and long `bucket 15` (bottom) of experiment 12 compared for prompt (blue) and reprocessing (red) alignment. Each point corresponds to the median of the histogram in a file with > 2000 di-muon events, error bars not shown. The points are ordered in run number (and file number) for each PXD sensor. The sensor identifiers are shown at the bottom.

for the shorter bucket (3.85 μm). For the u-coordinate, the residual instabilities after reprocessing seem much closer for the two buckets. This again indicates that the time resolution of the alignment is still not sufficient to remove the time dependence. But as we have already learned, this seems to be the case only for some sensors. Moreover, the values after the reprocessing already suggest that the remaining effect on physics should be negligible according to our previous studies [71, 72], with RMS of the medians being less than 3 μm and 5 μm in u and in v, respectively.
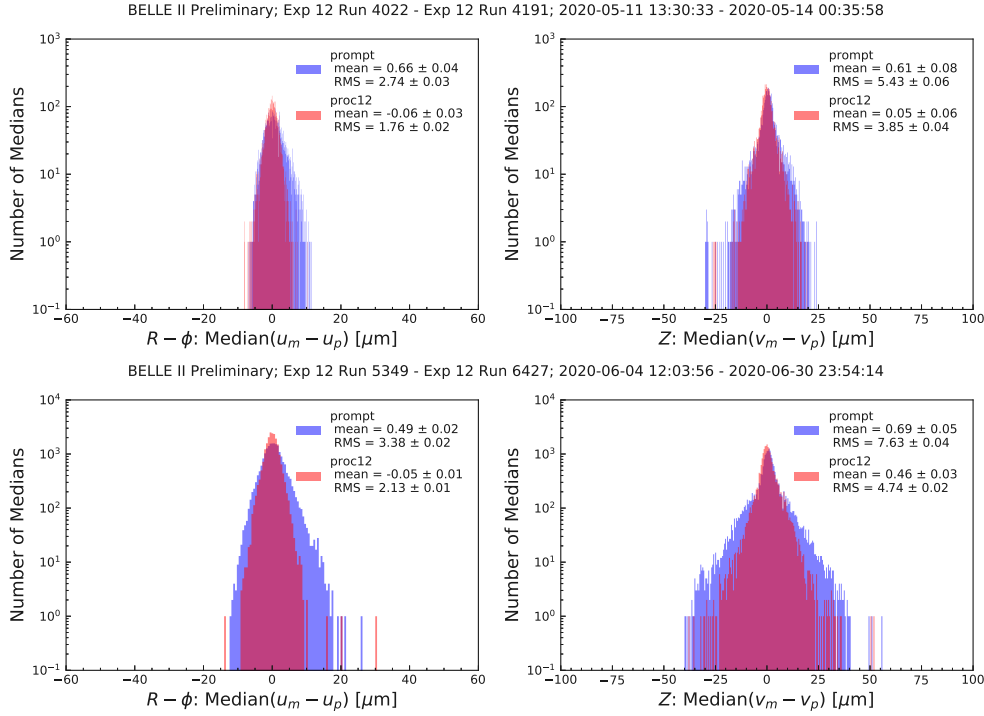
Figure 17.23: Distribution of medians of track-to-hit u (left) and v (right) residuals integrated over all PXD sensors in the short `bucket 11` (top) and long `bucket 15` (bottom) of experiment 12 compared for prompt (blue) and reprocessing (red) alignment. The mean and RMS of the distributions are shown together with their statistical errors. Only files with > 2000 di-muon events are used.

Correlations of PXD Sensor Residual

The time-dependent plots also suggest that movements among some sensors are highly correlated, at least before the reprocessing. We have thus plotted the medians of one sensor versus the other for all sensor combinations. For entirely random movements, the points should not exhibit any significant correlation. We extract the correlation factors from all these plots and all buckets. One can generally observe many interesting patterns in such correlation plots. As an example, for correlations in the u-coordinate, the extracted correlation factors for the long bucket are compared in Fig. 17.24 for the prompt and the reprocessing alignment.

The sensors of the second layer are localized in the lower right 4×4 block of the correlation matrix. The large red and blue blocks, visible nicely in the prompt alignment, correspond to divisions into half-shells. One cannot decouple half-shell movements from individual sensor movements easily in such plots, especially before the time-dependent (relative) half-shell movements are suppressed in the reprocessing. This can be clearly observed in the bottom figure, where the estimated correlations are reduced for all sensors, from more than 70% to less than about 30%. Very similar patterns can be observed in the correlations for long buckets with higher luminosities. For the initial three buckets, it seems the half-shell movements are still present, probably because two long periods are needed to accumulate the required statistics for one alignment block. Towards the end of the data-taking period, the removal of correlations seems to be more

efficient. Thus we show the last bucket as it should be more representative of future data taken at higher luminosities. For two short (3 and 5 days long) buckets, the prompt correlations are almost identical to those after the reprocessing, showing similar patterns. After the reprocessing, the remaining correlations in the u-coordinate do not indicate serious issues.

For the v-coordinate, very different correlation patterns, shown in Fig. 17.25, are observed. The sub-division into half-shells is not really apparent. This could indicate that the instabilities observed for sensors in the u-coordinate are mostly caused by small relative coherent movements of all sensors in a given half-shell or all sensors in the PXD with respect to the SVD. But in the v-coordinate, much more significant correlations are observed for very different and quite specific combinations of the sensors. Namely, the sensors 1.3.1 and 1.4.1 in the first layer are correlated to the sensors in the second layer and especially to the forward sensors 2.4.1 and 2.5.1. Sensor 1.3.2 is missing because it is not working. This makes a gap (with sensor material present), and thus the residuals could be correlated (at least partially) as an effect of the tracking, possibly projecting instabilities from the second layer to the first layer (or vice versa).

Whereas this analysis cannot de-correlate such effects, the alignment method could do so in principle, with a sufficient time resolution. Despite for the v-coordinate, a general decrease in the correlations coefficients for all sensors is observed, this is less significant than for the case of the u-coordinate, especially for the problematic correlations. For example, the correlation coefficient between sensors 1.4.1 and 2.4.1 is reduced only slightly—from -95% to -91%.

The correlation coefficients themselves do not reflect the size of the movements but only the degree to which they are related. As we have seen, the scale of the median variations was significantly reduced on average by the reprocessing. Thus the remaining correlations themselves are likely not dangerous but indicate that those sensors could move together. If it only remains as a result of only slightly too long alignment blocks, we should be able to see in the determined reprocessing alignment constants what degrees of freedom are genuinely responsible, see the next Sec. 17.4. Here we are more interested in the final effect we can observe with our collected validation data sample.

17.3.2   Time Stability of the VXD Sensor Alignment

Variations in the means or more robust medians of residual distribution for the VXD sensors can be attributed to the combination of random fluctuations due to limited statistics $\sigma^{\texttt{stat}}$ and true variations of the sensor positions in time combined with stochastic contribution, e.g., due to high-frequency[4] vibrations, corresponding to instabilities $\sigma^{\texttt{time}}$ in the two sensitive coordinates. The time instabilities $\sigma_u^{\texttt{time}}$ and $\sigma_v^{\texttt{time}}$ are of interest in this section. As the alignment parameters are constant within a range of one or more runs (or even a whole bucket for the SVD sensors and CDC wires), these variations correspond to instabilities in the sensor alignment and contribute to a worse per-sensor resolution, if these are well averaged out over time. If the movements result in a non-Gaussian distribution on average (over large data sets), this can also contribute to a worse statistical

---

[4]Much smaller than our alignment resolution, being several hours on average in the later buckets of experiment 12.
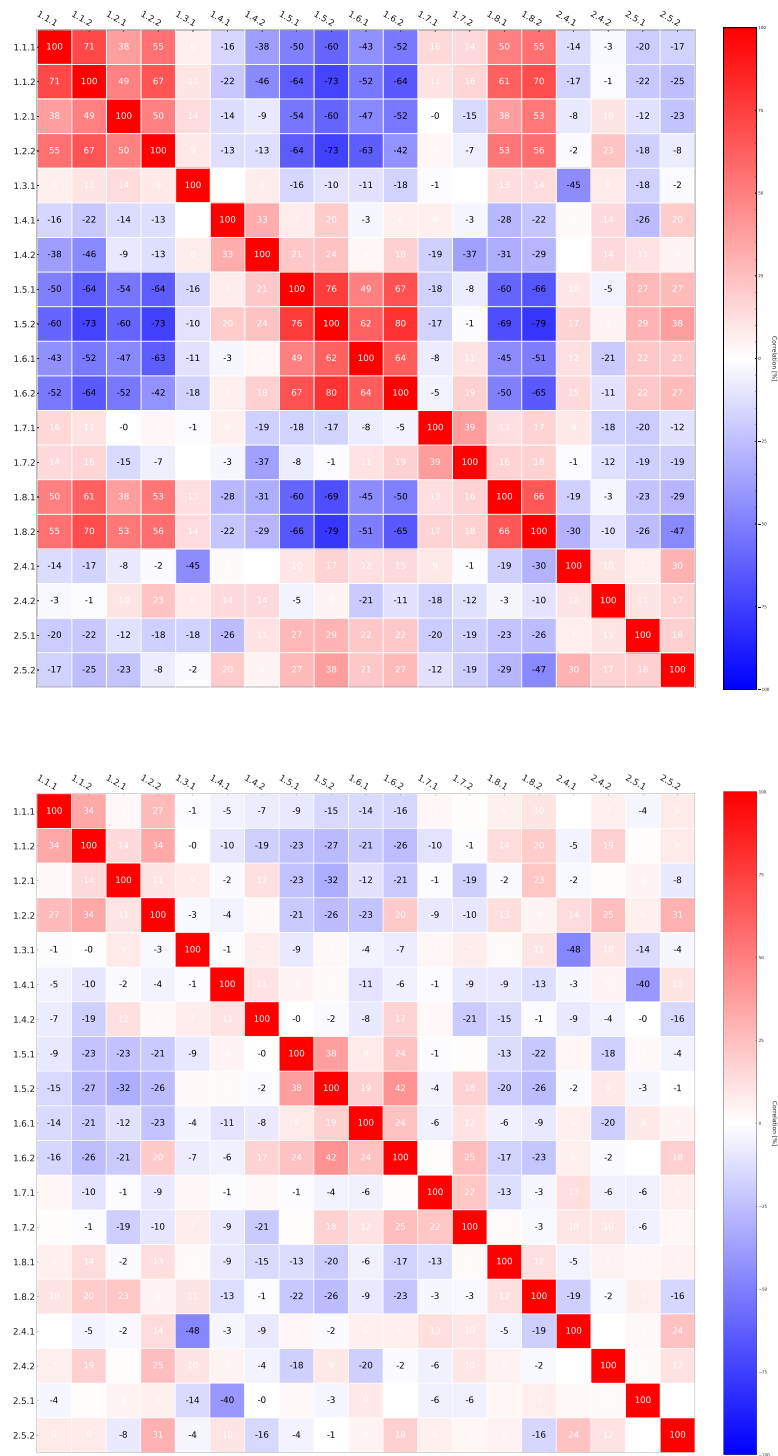
Figure 17.24: Correlations of medians of track-to-hit u residuals among all pairs of PXD sensors in long `bucket 15` of experiment 12 compared for prompt (top) and reprocessing (bottom) alignment.
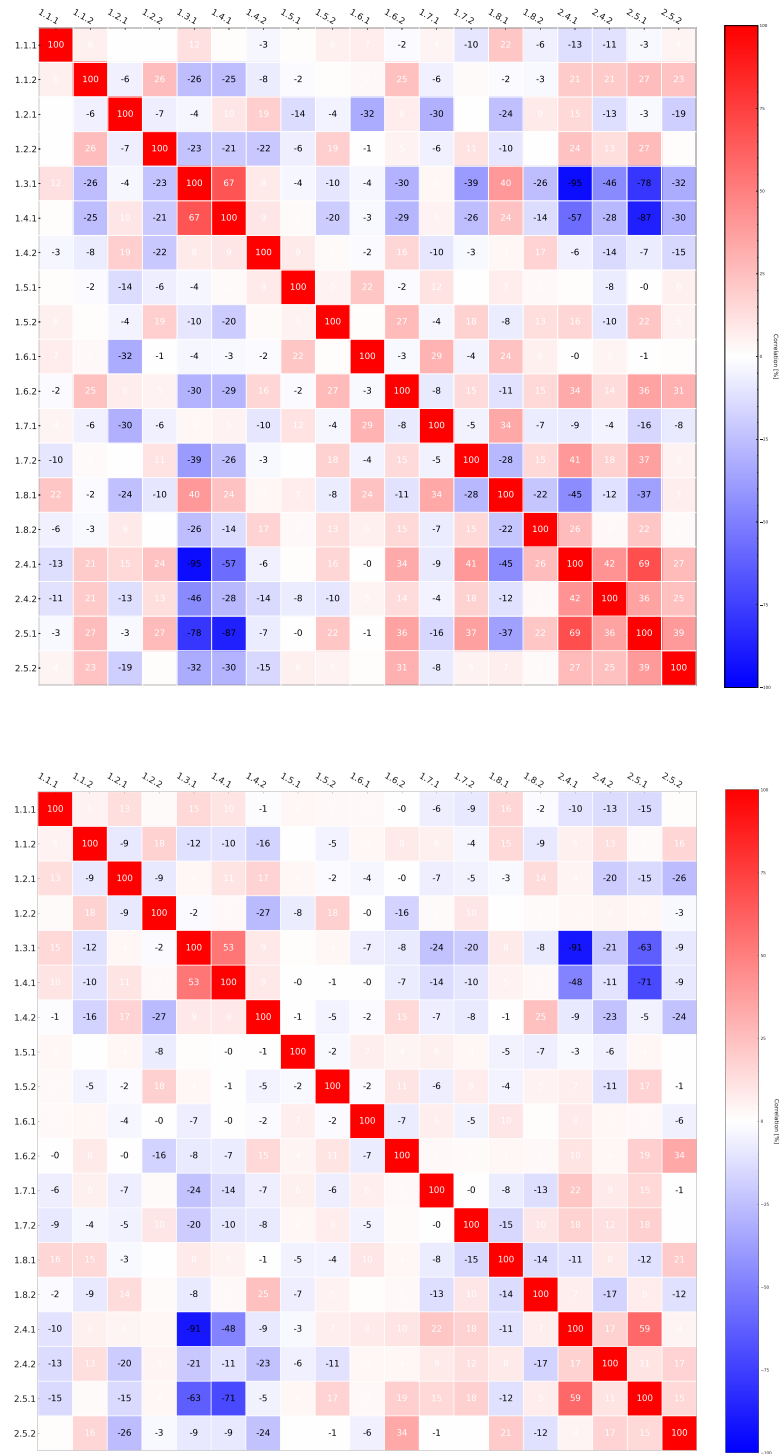
Figure 17.25: Correlations of medians of track-to-hit v residuals among all pairs of PXD sensors in long `bucket 15` of experiment 12 compared for prompt (top) and reprocessing (bottom) alignment.

alignment precision. For the PXD, we will thus only quantify the instabilities observed in the previous section but consider all buckets. In addition, instabilities will be evaluated for all VXD sensors to provide an estimate of the total effect on the observed sensor resolution.

Unfortunately, the analysis is not well suited for the outer SVD layers, where some sensors have only a limited number of hits. This will result in larger estimated uncertainties of the method. The data set is split into files with about $4000 - 6000$ di-muon events on average. All smaller files are discarded and for each residual histogram evaluated, at least 20 entries are required.

The analysis is first performed per each sensor individually. Given sensor and coordinate u or v, the median of the residuals (from a histogram collected by a DQM module) $\texttt{med}[r_{u,v}]$ is evaluated. The median is used instead of the mean at this point for better robustness against outliers.

The instability for each sensor is then evaluated by taking averages over all files in experiment 12 as

$$\sigma_{u,v}^{\texttt{time}} = \sqrt{\sigma_{68}^2[\texttt{med}[r_{u,v}]] - \texttt{med}^2\left[\frac{\sigma_{68}[r_{u,v}]}{\sqrt{N}}\right]}, \qquad (17.9)$$

where $\texttt{med}\left[\frac{\sigma_{68}[r_{u,v}]}{\sqrt{N}}\right]$ is our estimate of $\sigma^{\texttt{stat}}$ and N is the number of entries in each histogram where the median is evaluated.

Although we generally follow the formalism of Ref. [118], we have replaced most estimators with their more robust counterparts. The σ_{68} method is used instead of the standard deviation to estimate the error of the mean of the residual distribution for each file. As residuals have tails from multiple scattering and δ-electrons and uncorrected movements can be present, their distributions might be non-Gaussian[5]. The approach of taking median over all files to estimate $\sigma^{\texttt{stat}}$ is used to cope with variations in the sample (file) size, which is however small, and no small files are used. A dedicated data sample would be more suitable, with more even distribution of the number of events per file and ideally also pre-scaled in $\cos\theta$. However, in this section, we are limited to the validation data set collected during the reprocessing. The used file size is the maximum available at the calibration center for the tight di-muon skim. For other data samples, no such reliable simulation is available. In theory, the Bhabha events would be more suitable (due to the much higher cross-section) and can be simulated reliably, but these are significantly pre-scaled at the calibration center.

To estimate the systematic errors of this method, we have used the same process to derive time instabilities for sensors in MC simulation. As we do not simulate any time dependence, these results should be consistent with zero. This is, however, generally not the case. We have experimented with many combinations of estimators. It is notable that when the standard deviation is used instead of the σ_{68} method to estimate $\sigma^{\texttt{stat}}$, the MC underestimates the instabilities, giving negative results of generally corresponding size as with the σ_{68} method. Constructing an unbiased estimator for such a large range of observed statistics per sensor is difficult.

---

[5]This estimate assumes Gaussian-distributed residuals. The error of the mean is used instead of the error of the median, which is less trivial to calculate.

For highly populated sensors, the MC bias of our estimator is less than 1 μm. Once the average histogram statistics becomes lower, the bias grows and also gets generally less stable across the ladders. It grows up to about 4 μm for some sensors in the SVD layer six and the most backward sensors. It also oscillates by about 2 μm towards lower values. These sensors also receive the lowest amount of hits, and the bias could be caused by failing approximations in the estimations of the statistical error contribution. However, the MC is simulated without small files and with four times larger average statistics per file (about 15000 events), and thus such effects should be suppressed. This seems not to be the case, and much larger statistics would be needed. Even with large statistics, very subtle dependence on the sensor occupation and possibly the geometric position of the sensor in combination with the correlations in the track sample could remain.

Thus, instead of attempting to derive an unbiased estimator of the instability, we use the MC without misalignment to sample the expected distributions of the residuals observed in the data. For this reason, all MC files are merged to produce high statistics templates of the expected shape of the residual histogram for each sensor (without time instabilities). For each residual histogram observed in the data, a new histogram is sampled from the MC template with the same number of events as observed in the data. To allow for an unbiased estimator, at least in principle, we have to allow also negative values for $\sigma_{u,v}^{\texttt{time}}$. When a negative argument under the square root is obtained, $-\sqrt{-\dots}$ is returned instead.

The instability is then evaluated both for data histograms and for the sampled histograms from MC. We test this method by estimating the instabilities of the ideal MC. As already mentioned, this method estimates a non-zero positive instability for almost all sensors. As discussed below, we take a mean over the ladders to suppress the instabilities across the ladders. The standard deviation of the estimates over the ladder is taken as a statistical error. The sampled MC estimates averaged over ladders, and the MC instabilities match almost perfectly. The largest deviation of the MC estimate from the re-sampled MC estimate, about 0.3 μm, is taken to be the systematic error of the method. We add it in quadrature to the estimated error band for all re-sampled MC estimates averaged over ladders. In fact, we expected some finite resolution of this method as the input residual histograms are binned in bins of 1 μm width.

Finally, as already mentioned, once the instability, its error, and the corresponding re-sampled MC estimate for each sensor are evaluated, the results are averaged over ladders in each layer separated by sensors' z-position (the so-called sensor number, where sensors with sensor number 1 are the most forward). This separation is the most natural, as in general, the time-dependent effects seem to mostly depend on the z-coordinate of the sensors and layer number. As a function of ladder number, the only larger systematic deviations of the estimated instabilities occur for the PXD around the position of the disabled sensor in layer 1 and above in layer 2. A much less significant ladder-dependent effect is seen in SVD layer 3. This can also be seen in Fig. 17.26, where the mean instabilities are illustrated for each sensor in the first three layers individually. More outer layers do not show any significant systematic variation over ladders.

The final instabilities (and corresponding errors) are determined as averages
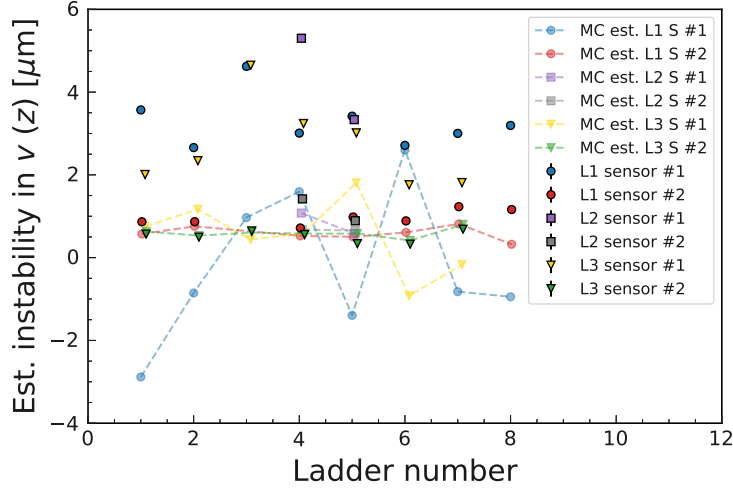
Figure 17.26: Estimators of instability $\sigma_v^{\mathtt{time}}$ for every sensor in the first three VXD layers for di-muon validation data (points with edge) in experiment 12 (after reprocessing) and MC estimates $\sigma_{\mathrm{MC},v}^{\mathtt{time}}$ (without any misalignment– points without edge, connected with dashed line for clarity) re-sampled to mimic statistics per every histogram observed on data. The points have no associated estimated errors.

over all ladders:

$$\langle \sigma_{u,v}^{\mathtt{time}} \rangle = \mathtt{mean}[\sigma_{u,v}^{\mathtt{time}}]$$
$$\sigma \langle \sigma_{u,v}^{\mathtt{time}} \rangle = \mathtt{std}[\sigma_{u,v}^{\mathtt{time}}]. \tag{17.10}$$

As we have just evaluated the bias of our MC estimator, we directly subtract it from the data estimates and do so for every average over ladders. The results are shown in Fig. 17.27 and also compared for the prompt and reprocessing alignment. While after reprocessing, most instabilities do not seem significant within the resolution of our method, this is not the case for several sensors in the prompt alignment.

For the u-coordinate, very good time stability can be generally observed, well under 3 μm for all sensors except the most backward sensors in SVD layer six. Such a level of random residual misalignment is known to have a negligible physics impact, and no corresponding degradation of resolution is expected. A small, general trend can be observed with growing instability towards the forward sensors (and backward for SVD layers > 4). But here, the estimated error of the method, shown by the gray bands (corresponding to one and two standard deviations of the averages over ladders for the re-sampled MC estimates), is large. The systematic error of 0.3 μm is added in quadrature to all error bands, primarily affecting the error bar size of the most precisely measured sensors for the u-coordinate.

In the longitudinal v (z) coordinate, the instabilities are significantly larger. However, practically all estimates are well within 5 μm needed for negligible resolution degradation in the z-coordinate after the reprocessing.

The instability growing towards a larger sensor radius (and size) and distance from the center of the structure seems natural. Generally, the larger effect at the forward structures is probably related to the way the vertex detector is attached
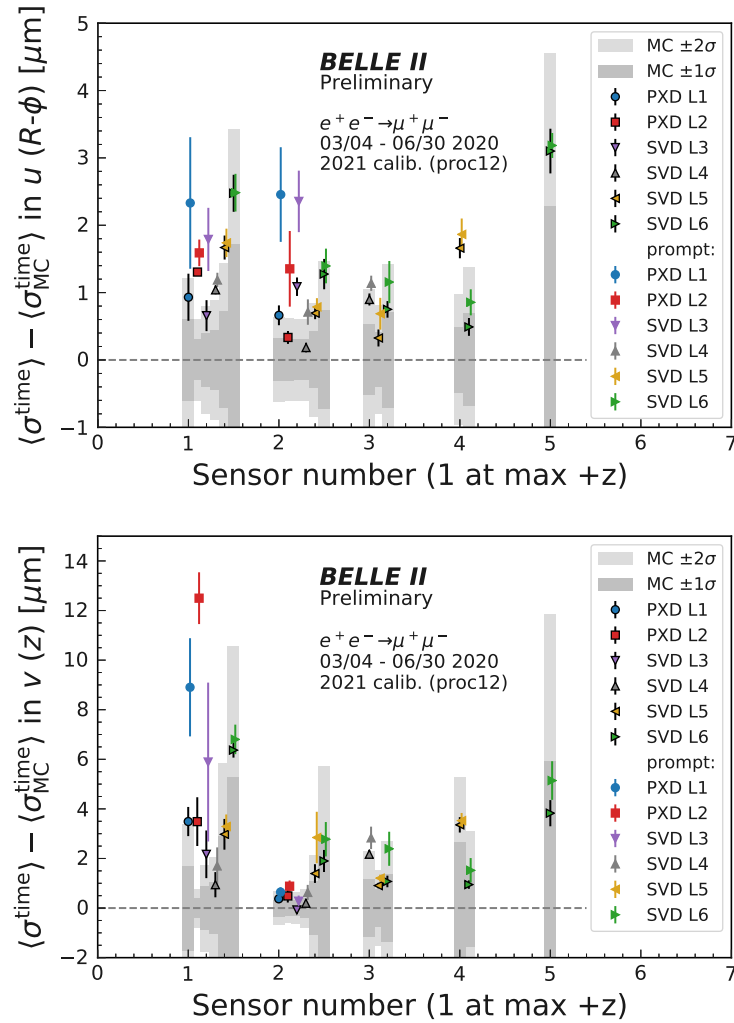
Figure 17.27: Estimated unbiased average instability for the VXD sensors sep-arated into layers as a function of sensor number in u (top) and v (bottom) coordinate. The difference of the data estimator averaged over ladders $\langle\sigma_{u,v}^{\mathtt{time}}\rangle$ and the (biased) MC estimator $\langle\sigma_{\mathrm{MC},u,v}^{\mathtt{time}}\rangle$ is plotted. Data are shown as points with error bars, separately for prompt (no edge) and reprocessing (points with edge and black error bar) alignment. The error bars of the data points show the standard deviation of estimated values across all sensors in ladders with the same sensor number (and layer). The gray bands show one and two standard deviations of the re-sampled MC estimates across ladders (like for data), which mimic statistics for every histogram observed on data. The error bands of the MC estimates (which are trivially exactly zero after bias subtraction) include a systematic error of 0.3 μm added in quadrature to the standard deviation of the MC estimates.

to the rest of the construction. However, one should note that the instabilities in the outer SVD layers are not significant within the estimated resolution of the method. The re-sampled MC without any misalignment would have the (bias-subtracted) mean precisely at zero. However, the values would oscillate around zero with the standard deviation shown by the gray bands. The measured estimates on data are, in fact, positive for every single sensor. Thus on average, non-zero positive instabilities are observed with a high significance. After the bias subtraction, only the most stable sensors with number 2 in SVD layer three have a slightly negative estimate, consistent with zero within the estimated errors. We have thus only indications that the instabilities actually grow for the outer sensors, given the large estimated uncertainties.

What is, however, slightly worrisome is the large instability of the forward PXD sensors, namely in the (incomplete) second layer. These are also significant within our resolution for the u-coordinate. The estimated remaining time instabilities in the range $3 - 5$ μm could lead to a very small degradation of the z_0 resolution in the forward direction. It is not clear if the absent sensors cause these effects, and the overall structure of the complete PXD with both layers fully equipped with sensors could be more stable. The larger instabilities observed in the forward sensors could be related to the elastic connection of the support structures to the silicon elements and some thermal tensions. While the backward sensors are connected to the forward ones, these seem to move somewhat independently, possibly being slightly rotated around the joints' axis and/or deformed. In the following section, we will try to shed some light on these movements.

Such instabilities could become a limiting factor for the complete PXD z_0 resolution or for any further upgrades, but fortunately seem to be reduced to an acceptable level by the time-dependent alignment. It should be investigated to what operational conditions these instabilities are correlated to mitigate them. One possible cause could be thermal expansions/contractions from beam heating or switching high voltage of the PXD and/or SVD on and off. This can happen regularly at the start and end of data-taking, often in the afternoon and morning JST, or after a lost accelerator beam or other issues requiring turning off PXD low voltage to prevent detector damage from accelerator background conditions.

Nevertheless, the expected physics impact of the estimated remaining instabilities should be only very small. These should only be observable as a very slight degradation of the resolution in the v (z) coordinate. This is also true for the whole SVD. Thus the larger width of the u-residuals for the SVD observed in the data cannot be caused by residual alignment instabilities. This is yet another confirmation that the MC simulation for the SVD sensors is too optimistic, and additional effort is needed to bring the simulation closer to reality. It is possible that this difference could be observed as degraded d_0 resolution for $V0$ particles (K_S^0 and Λ^0) without PXD hits on tracks. Otherwise, as was shown in the previous section, the overall effect on the resolution for tracks from IP is still acceptable, and data are less than 20% worse than MC simulation.

17.4 Time-dependent Alignment Constants

In this section, we will look at the extracted time-dependent constants themselves. There are about 300 run blocks in experiment 12 reprocessing, each with 786 determined parameters (VXD half-shells, CDC layers, and PXD sensors up to cubic deformations), while using 30 constraints (24 for VXD and 6 for CDC— twist constraint is not used). By presenting and evaluating the results in detail, we could easily make this thesis twice longer and still leave plenty of room for end-less hypothesizing about the explanations of our observations. Instead, we have demonstrated good physics performance using detailed validations in the previous sections. Here, we will look into some of the determined constants just briefly to better understand the time-dependence in the PXD at the sensor level. After relative half-shell positions, these are the most important degrees of freedom for the precision measurements and, unfortunately, also the most unstable.
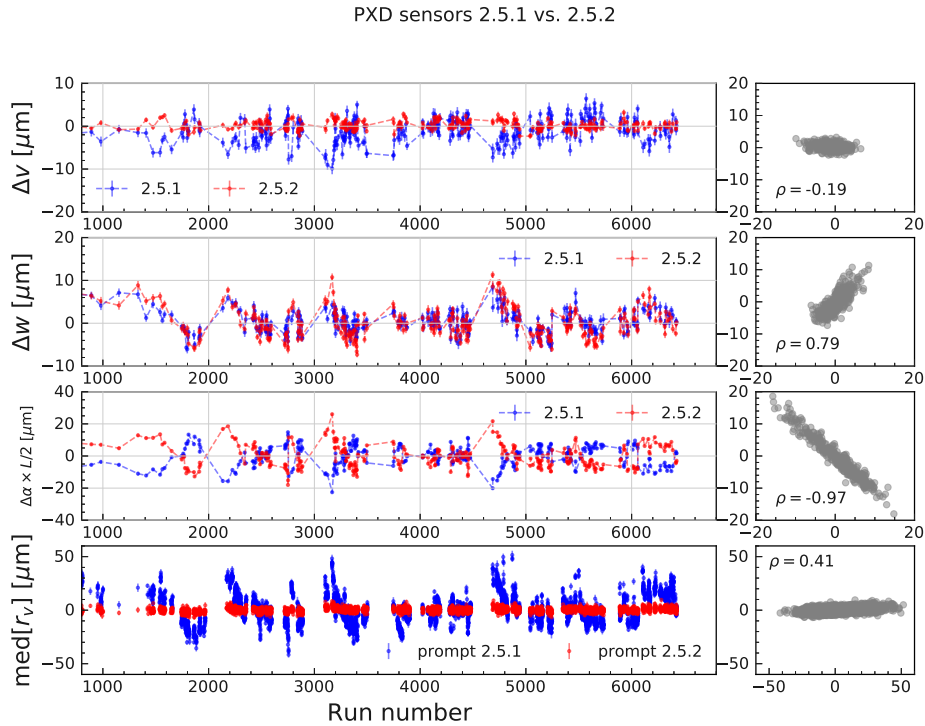
Figure 17.28: Determined run-dependent alignment corrections (first three rows) in v, w and α and medians of residuals with the prompt alignment (bottom row) in the v-coordinate for PXD sensors 2.5.1 and 2.5.2. The panels on the right show the correlations and correlation factors of the points on the left.

We have observed the most significant instabilities in the second PXD layer in the previous sections. We were wondering why only the forward sensors seem to be affected. This strange inconsistency is visible in the v residuals, residual correlations, and estimated instabilities. All these methods use track-to-hit residuals and cannot decouple the correlated changes in different degrees of freedom. However, alignment does precisely this. We can look at the most unstable parameters v, w, and α for the sensors 2.5.1 and 2.5.2 of the same ladder in layer two. Instead of α, we translate this parameter into the corresponding shift at the edge

of the sensor by multiplying α by the half-length of the sensor. The alignment corrections as a function of run number are shown in Fig. 17.28, where the last row shows the medians of the v-residuals for a reference.

We can see that the apparent calmness of the backward sensors was only an illusion. The reason is that di-muon tracks originating from the IP hit the backward sensors approximately in the middle of their length. As roughly the same amount of tracks points forward and backward, the residuals will average these two effects. But for the forward sensors, all tracks go forward from the IP, and thus the averaging is not in effect here, and we can observe the biases with average residuals over the sensor and all di-muon tracks. The parameters α are the most unstable—moving sensor edges in the range of ± 20 μm—and highly anti-correlated. Furthermore, these movements are obviously connected with the shifts in w, with roughly ± 10 μm range. Regarding the v parameter, the forward sensor seems to be more affected. The total combined movement results in the PXD ladder's bowing around the sensors' glue joint. The changes in v might be signs of the glue joint expanding or contracting. However, this could also be caused by external effects coming from the mechanical support. In any case, we can see that despite the residuals can have means even $+50$ μm for the forward sensor, this actually results from a combination of misalignment in several different degrees of freedom. These are themselves more stable, but the total effect is rather significant if ignored.
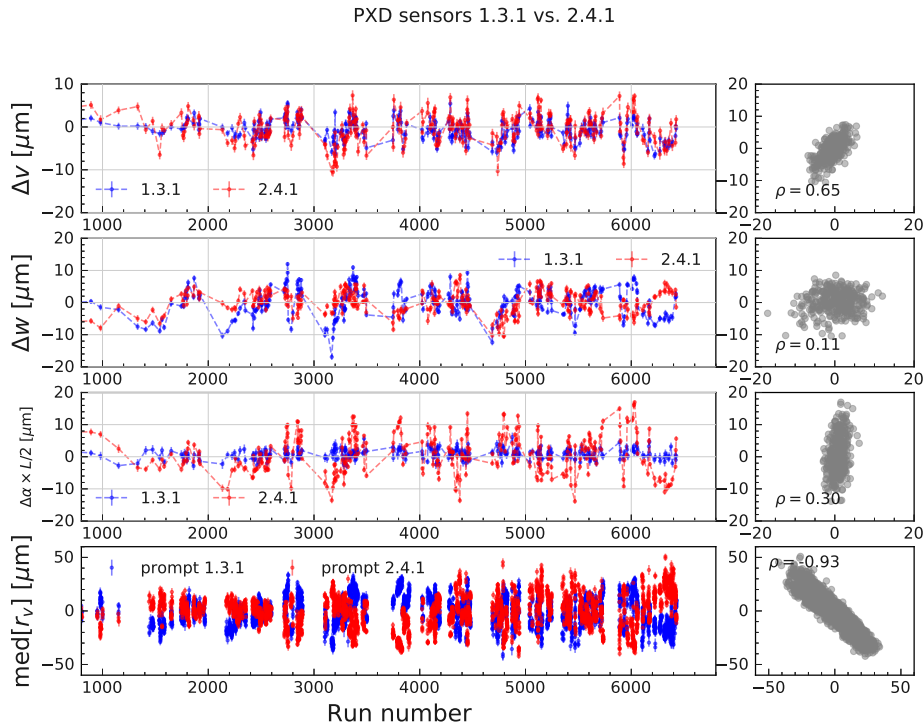
Figure 17.29: Determined run-dependent alignment corrections (first three rows) in v, w and α and medians of residuals with the prompt alignment (bottom row) in the v-coordinate for PXD sensors 1.3.1 and 2.4.1. The panels on the right show the correlations and correlation factors of the points on the left.

Let us look at one more example of time-dependent PXD sensor alignment

constants to explain one of the previously observed mysteries: the large correlation between v-residuals for sensors 2.4.1 and 1.3.1, which lie in different layers. We show comparison of their movements and median residuals in Fig. 17.29. The large anti-correlations of the residual medians are not clearly present among the alignment parameters. Thus the residuals are likely correlated due to the tracks, which mostly hit this combination of sensors. The instabilities in the first layer are much smaller for α, but we can observe a similar scale for the v and w movements as in the previous example.

These two examples demonstrate an additional benefit of the reprocessing alignment. It can help us to understand our detector from a mechanical point of view. For this, we will need to investigate the corrections as a function of time, rather than a run number, to be able to overlay them with, e.g., continuous temperature measurements. Moreover, the run-dependent alignment for the PXD sensors is certainly not an overkill. The combined instabilities can lead to residual biases several times larger than the sensor resolution, especially for tracks entering under a large incidence angle.

17.5   CDC Performance

The purpose of this brief section is to look at the performance of the CDC in the prompt and reprocessing calibration and compare it to the MC expectations. We use full cosmic (PXD+SVD+CDC) tracks from the beginning of this chapter for evaluation. It is important to remark that CDC momentum resolution depends on the gas composition (and resulting gain), which is time-dependent (due to the addition of water vapor to prevent discharge in a high radiation environment) and not reflected in our MC simulation for comparison. In fact, the momentum resolution was best at the beginning of the experiment, before the water started to be added to the gas. The momentum resolution seems almost independent of the alignment improvements, as can be seen in Fig. 17.30. The extracted values of the relative momentum resolution $\sigma(p_t)/p_t$, where $\sigma(p_t)$ is sigma of a Gaussian fit to $\Delta p_t = (p_t^{\mathrm{up}} - p_t^{\mathrm{down}})/\sqrt{2}$, for data just agree with the expectations [10], see Sec. 3.2 and differ from the ideal simulation by less than 20%. After reprocessing, we get

$$\frac{\sigma(p_t)}{p_t} = (0.099 \pm 0.001)\% \, p_t \oplus (0.307 \pm 0.003)\%, \qquad (17.11)$$

where p_t is in units of GeV/c.

More interesting are charge-dependent relative momentum biases $\Delta p_t/p_t$ as a function of p_t, shown in Fig. 17.31. We can see a momentum bias also in the MC simulation, but charge-independent. The biases in the prompt alignment are reduced by the `proc12` reprocessing, but not entirely. Also, it looks like only the momentum bias for positively charged tracks changed. In the past, these biases used to be much larger. We managed to successfully reduce them by excluding cosmic ray tracks passing the poorly described magnetic field regions from the alignment procedure. We suspect the remaining biases to be related to inconsistencies in the geometrical description of the CDC; see also Sec. 16.4. Nevertheless, the relative size of the remaining effects is only at 10^{-4} level or smaller.
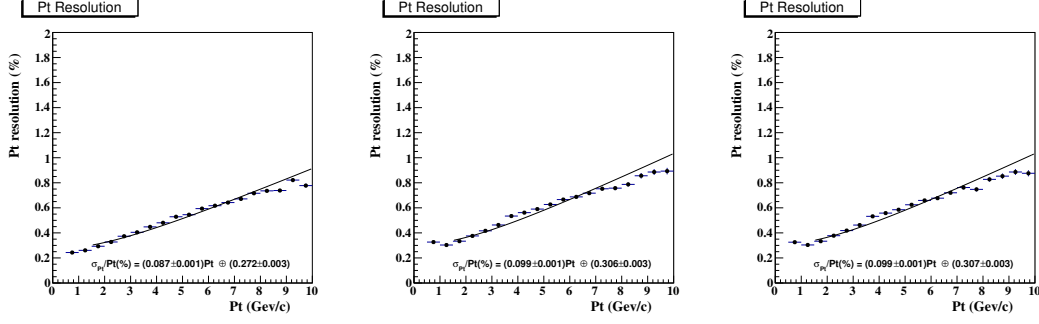
Figure 17.30: Relative momentum resolution of cosmic ray tracks in MC (left) and data of experiment 12 for prompt (middle) and `proc12` reprocessing (right).
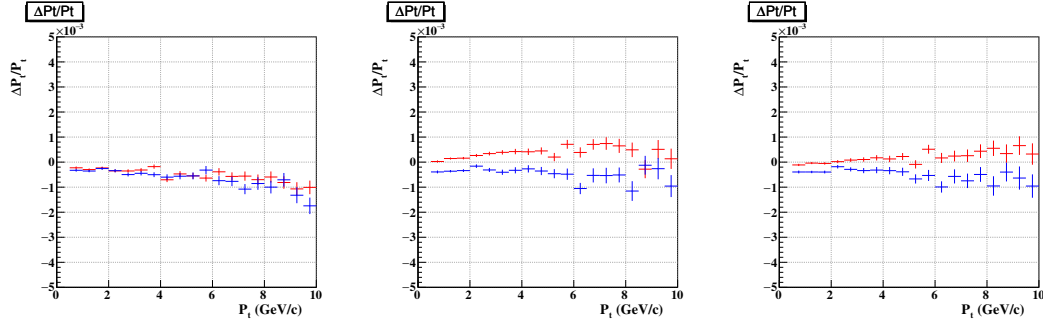
Figure 17.31: Relative momentum bias for positively (red) and negatively (blue) charged cosmic ray tracks in MC (left) and data of experiment 12 for prompt (middle) and `proc12` reprocessing (right).

17.6   Summary

In this chapter, we have demonstrated a very good alignment performance after the reprocessing. We have also identified some problems in the preliminary prompt alignment and verified that those had been largely mitigated. Beyond this, we identified some remaining issues. These are mostly related to the remaining systematic effects stemming from uncorrected time-dependent deformations of the SVD, mainly affecting the z-coordinate. Residual statistical misalignment is dominated by time instabilities, which are suppressed to an acceptable level by the reprocessing. The most significant remaining instabilities are observed in the second PXD layer for the z-coordinate. In general, negligible or only minimal degradation of physics performance is expected as the result of any residual misalignment based on the validation results.

This is not true for the prompt alignment, where significant degradation is observed both in systematic deviations and reduced statistical precision of the alignment due to the ignored time dependence. While the prompt alignment has a good performance in general, it is desirable to perform physics analyses that are particularly dependent on the vertex detector using the reprocessed data. As the delay of the reprocessing from the prompt alignment and thus data-taking can be close to one year, we aim to move the workflow to the prompt alignment loop. If this effort is successful, publication-level precision alignment can be achieved shortly (in a couple of weeks) after the data-taking.

CHAPTER 18

Possible Extensions

This chapter will give an overview of possible extensions to this work. Technically, almost all discussed features are already implemented and at various testing levels. However, these extensions (except stand-alone KLM alignment) are not used in the actual production of calibrations constants for physics analysis yet. Thus we only summarize the various ways of extensions and often also briefly discuss what still needs to be done if we want to use them. The primary challenges are lack of workforce (primarily for validation), lack of data, and a number of practical aspects arising when moving from perfect MC simulation to the real detector output.

We have envisioned the full simultaneous alignment and calibration of the VXD, CDC, and KLM, corresponding to more than 150×10^3 global parameters. In principle, everything is in place to exercise this approach, and, of course, many partial tests have been performed in the past. However, the general alignment philosophy has evolved since the first data from the completed Belle II detector arrived. As the alignment is a calibration on which most others depend, any recalculation due to possible errors requires a complete data re-processing. Thus the preferred alignment philosophy is very conservative: any update of the official alignment must be extensively validated, such that the performance is not degraded in any aspect (ideally is improved in some validations). For example, as the existing CDC calibration is working well, there is no urgent need to replace, e.g., T_0 or $x - t$ relation calibrations by other methods. Furthermore, any such replacement will require additional validation work. Nevertheless, several of these extensions were used in (usually unsuccessful) attempts to improve some aspects of alignment performance on the real data.

18.1 Data Samples

The alignment software allows for a wide range of different input data and usage of some advanced constraints beyond the IP constraint for di-muon events. However, these advanced data samples are not yet used for real constant production for various reasons.

In the future, with higher luminosities and dedicated calibration skims, we could utilize $J/\psi \rightarrow \mu^+\mu^-$ decays with an invariant mass constraint. Possibly, also other two-body decays could be used. Already at current luminosities, $e^+e^- \rightarrow \mu^+\mu^-$ events can be used with the full four-momentum constraint. The beam energy and the boost are also time-dependent, and thus an established calibration of these quantities is needed first. This calibration also has to derive the corresponding covariance matrices for the parameters of beam kinematics,

which are inputs for the additional external measurement at the common vertex of a GBL trajectory, combined from the two muons. Such calibration is currently being deployed to the automated prompt calibration.

Another interesting data source are multi-body vertex-constrained decays from channels like $D^{\star+} \to \pi^+_{\mathrm{slow}} D^0 (\to K3\pi)$. Here $D^{\star+}$ is produced by the $e^+ e^-$ collision and the slow pion π_{slow} is used for its identification and reconstruction. The true interest is in the four-body decay $D^0 \to K^- \pi^- \pi^+ \pi^+$. In this case, the tracks connect potentially four different parts of the detector, in contrast to (merged) cosmic rays and di-muon events, possibly further improving alignment systematics. Also, in this case, probably higher luminosity is needed, as well as dedicated alignment skims to start with tests using real data. Of course, as always, a dedicated MC study should be available before updates of the official alignment procedures.

One of the potential drawbacks of such advanced data samples is background contamination, which must be minimal for the use in alignment. This then naturally limits the available statistics. Another challenge might arise from ISR and FSR for kinematic constraint when the mass constraint is incorrect as some four-momentum is lost due to the radiated photon. This is yet another reason for dedicated MC studies to eliminate such kinds of events, e.g., by using ECL information as a veto.

18.2   Calibration Constants

18.2.1   KLM Alignment

The initial implementation of the alignment of the BKLM modules was enabled by extending the `RecoTrack` capabilities to include KLM hits into the track fit. Several issues had to be solved to provide unbiased residuals [119] for the GBL trajectory. The BKLM is separated into backward and forward parts, each with eight segments. There are 15 layers of detection modules, of which the innermost two use scintillators, while the rest is equipped with RPCs. This gives 240 ($2 \times 8 \times 15$) modules with six rigid-body parameters, i.e. 1440 alignment parameters.

The EKLM alignment was implemented later. The EKLM has 14 detection layers in the forward and 12 in the backward end-cap. It is divided into four sectors in $R - \phi$. Each layer is composed of two orthogonal planes with strips. This gives 104 modules, for which three rigid-body alignment parameters are considered: u, v shifts and rotation γ in the module plane, giving 312 alignment parameters.

The (simultaneous) alignment of BKLM and EKLM with Millepede II is implemented in the CAF [120] and included in the regular automated Belle II prompt calibration [121]. It runs after the completed VXD+CDC alignment and uses single muons from di-muon events and (unmerged) cosmic ray tracks. It is not executed simultaneously with VXD+CDC alignment due to issues with backward extrapolation for cosmic ray tracks to outer detectors. The hits from outer detectors are attached to the tracks using a GEANT4-based extrapolation module. Unlike un-merged cosmic tracks, the standard merged cosmic tracks, which are essential for a good alignment, have assigned KLM hits only on the bottom arm. This could introduce new unwanted asymmetries mainly to the CDC alignment,
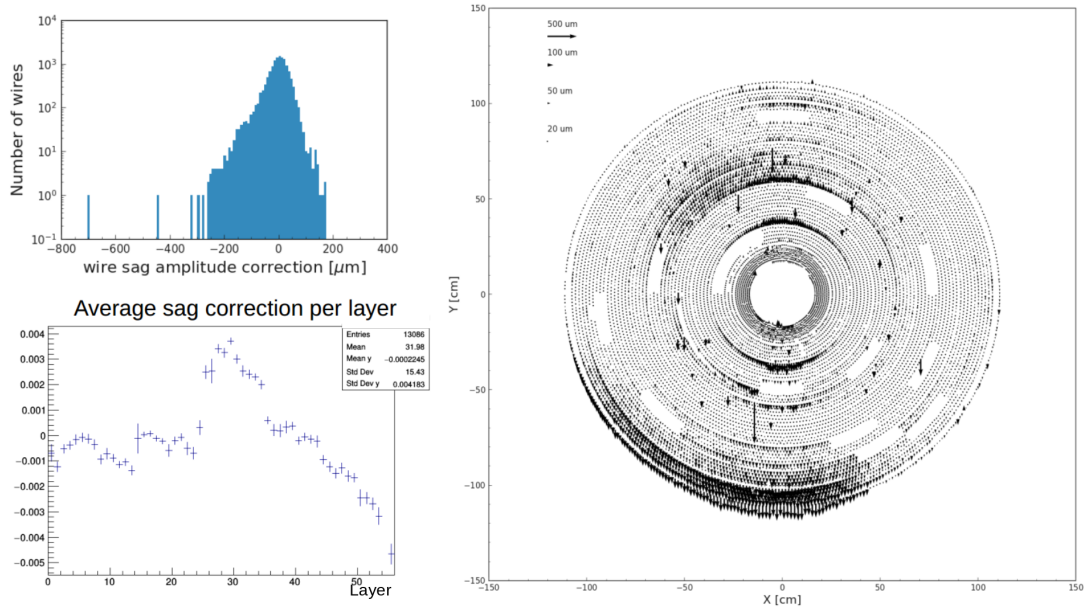
Figure 18.1: Determined amplitudes of the CDC wire sagging corrections (top left), average correction for wires per layer (bottom left) and visualization of the corrections for each wire (right).

and thus we postponed full KLM integration into the global alignment.

18.2.2 Alignment of CDC Wire Sagging

Due to gravitational and electrostatic forces, the wires are not exactly straight. The nominal tension for each sense wire is 50 g. The total tension from sense and field wires is 4.1 tonnes, and thus CDC had to be pre-stressed before wire-stringing [10]. Calculations show that the total expected deformation of the inner end-plate due to the tension is about 2 mm [21]. This deformation is taken into account in the wire z alignment parameters at both end-plates (which cannot be aligned from tracks). The tensions for the wires were measured after wire-stringing, and the results are used to correct the nominal tensions. The total tension for each wire is used to calculate its sagging due to gravitational force, using an approximate parabolic model. The total amplitude of the sagging applied in the reconstruction is up to 200 μm [10].

These values may also be updated using track-based alignment. One of the complications seems to be a non-linear relationship between the sagging amplitude and the wire tension, which caused drifts to non-physical regions in preliminary alignment tests. Another attempt was to align the corrections to the sagging amplitude, which could be, in principle, also extended to corrections to the wire shape in the horizontal direction due to electrostatic forces. These are currently entirely neglected in the reconstruction. The application of any such corrections will require some additional modifications of codes for the CDC parametrization and reconstruction. Nevertheless, a number of unofficial tests were performed, and the wire-sagging estimates were calculated by Millepede II using the MINRES-QLP solution method before Cholesky decomposition became available.

As we did not observe any improvement in the alignment performance, we have not proceeded with this feature yet. Also, a full set of simulation studies still needs to be performed to validate the method. The determined gravitational wire sagging amplitude corrections are shown in Fig. 18.1. These results are only very preliminary but suggest that significant corrections should be applied for some wires, especially in the bottom part of the outer layers. For those, the physics impact is relatively small, which could explain why basic validations did not show any improvements, and dedicated studies might be required before we can include wire-sagging alignment into the full automated global method and perform it simultaneously with the wire-by-wire alignment.

18.2.3 CDC T_0 Calibration

The T_0 calibration constant needs to be determined per each wire, giving about 14×10^3 global parameters. The CDC calibration can determine these numbers from scratch. We foresee that the Millepede alignment framework will only determine minor corrections to the previous values, ideally simultaneously with the rest of the alignment and calibration constants.

The derivative of the local u residual at each CDC hit with respect to T_0 is

$$\frac{\partial r_u}{\partial T_0} = v_{\text{drift}} \cdot s, \tag{18.1}$$

where v_{drift} is the estimate of the drift velocity for the given hit, and s is a sign factor to differentiate left/right track passage. The total sign might still need synchronization with other global and calibration parameter sets. This calibration constant is closely related to the additional track time correction we include into each track fit with CDC hits (fulfilling given requirements, see Sec. 11.8). Together, these would result in an undefined degree of freedom, so one needs to introduce one additional constraint to fix the average of all T_0 corrections to zero. Such an approach assumes that the global T_0 offset (to shift average event time to zero) is already performed by other means before or after the Millepede II calibration.

Calibration of T_0's was already exercised. The method performs well and is under integration into the global alignment.

18.2.4 CDC Time Walk Calibration

The time-walk effect introduces corrections depending on the collected charge in units of ADC. A simple parametrization $T_{\text{tw}} \propto 1/\sqrt{ADC}$ was recently replaced by $T_{\text{tw}} = p_{\text{tw}}^0 \exp(-p_{\text{tw}}^1 \text{ADC})$, where two parameters have to be calibrated for each of the 299 front-end read-out boards. For each hit, two global derivatives are added to each GBL measurement representing a CDC hit:

$$\begin{aligned}
\frac{\partial r_u}{\partial p_{\text{tw}}^0} &= -v_{\text{drift}} \cdot s \cdot \exp(-p_{\text{tw}}^1 \cdot \text{ADC}) \\
\frac{\partial r_u}{\partial p_{\text{tw}}^1} &= +v_{\text{drift}} \cdot s \cdot \text{ADC} \cdot p_{\text{tw}}^0 \cdot \exp(-p_{\text{tw}}^1 \cdot \text{ADC}),
\end{aligned} \tag{18.2}$$

where the total sign might still need synchronization with other global and calibration parameter sets. Time-walk calibration still needs more studies, including

validation on MC, especially after the recent change in the parametrization, which introduces non-linearities for which we previously observed issues in tests with the alignment of wire tensions. A reasonable preliminary calibration will be necessary for the global approach if we want to avoid additional iterations, which are generally not needed for the current bucket-to-bucket alignment updates.

18.2.5   Calibration of CDC $x - t$ Relations

The $x - t$ relations parametrize the function $l_{\mathrm{drift}}(T_{\mathrm{drift}})$, which translates time information into position measurement. The function is different for left/right passage (due to asymmetries resulting from Lorentz deflection for drifting electrons in the 1.5 T magnetic field), for every layer, each of 7 bins in the incident angle θ (17, 40, 60, 90, 120, 130 and 150°) and 18 bins of angle α (from $-90°$ to $+80°$ in 10° steps). The fit function is 5th order Chebyshev polynomial plus a linear function for larger drift times, giving nine calibration parameters for each $x - t$ function, where one parameter determines the position T_b of the Chebyshev-to-linear transition[1]. This gives 112 286 additional global parameters ($2 \times 56 \times 7 \times 18 \times 8$ excluding the boundary parameter, which cannot be determined in Millepede calibration). Therefore six global derivatives shall be attached to each CDC hit for $T_{\mathrm{drift}} < T_b$:

$$\frac{\partial r_u}{\partial p_{x-t}^i} = s \cdot C_i(T_{\mathrm{drift}}) \tag{18.3}$$

for $i = 0 \ldots 5$, where C_i is Chebyshev polynomial (of the first kind) of i-th order; and two global derivatives

$$\begin{aligned} \frac{\partial r_u}{\partial p_{x-t}^7} &= s \\ \frac{\partial r_u}{\partial p_{x-t}^8} &= s \cdot (T_{\mathrm{drift}} - T_b) \end{aligned} \tag{18.4}$$

for $T_{\mathrm{drift}} \geq T_b$. It should be noted that for some angles, the statistics is limited, and thus not all $x - t$ relations can be calibrated and must be taken from simulation estimates. In practice, about 80% of the $x - t$ relations can be calibrated from the data with reasonable precision. However, the standard CDC calibration takes care of all such corner cases, which is not possible in the one-time simultaneous global solution. Thus one first needs to determine which parameters can be reasonably included in the minimization with Millepede II and start the studies ideally on MC with a smaller subset of constants. We expect this task to be the most challenging if indeed pursued up to a point where it can be deployed for regular data calibration and alignment. So far, we have only implemented this calibration in the common Millepede method.

18.2.6   Extension to other Sub-Detectors

With the recent speed-ups in Millepede II, even more parameters, e.g., from the remaining outer sub-detectors, could be added if all the studies are performed.

---

[1]T_b is used as 6th parameter p_{x-t}^6.

However, additional work is needed for these sub-detectors to fully integrate their information into the track fit with GENFIT2 and thus the GBL trajectory construction. TOP, ARICH, and ECL already have their corresponding stand-alone alignment algorithms, which naturally rely on a good alignment of the tracking system.

18.3 Optimization of Time Dependence and Online Alignment

In Chapter 17, we have identified the most significant alignment issue to be the residual time dependence. In particular, we observed some uncorrected evolving deformation of the SVD. The time dependence for the PXD sensors, VXD half-shells, and CDC layers is defined by simply requiring enough events per each alignment block in the reprocessing. The alignment for this restricted set of parameters is computed for each block separately, on top of the previous full global alignment without any time dependence. Consecutive runs may be merged to make a block with enough events. However, the blocks might contain incompatible data in the sense that alignment changed significantly within the block.

As mentioned in Sec. 13.4.2, we have implemented the possibility of almost arbitrary time-dependence, which can be resolved in the global fit without the need to align each block independently. In addition, in this global approach, maximum information is used, and the final stage of run-dependent alignment could be merged into a single iteration with the wire alignment. Moreover, one could define different time dependence for the PXD and SVD sensors and possibly optimize them to describe the constants' evolution better.

However, the configuration for the time dependence must be defined beforehand. For that, we would need to develop specialized validation methods, which could provide optimized splitting of the data automatically. Such an approach is under consideration.

Currently, the selection of ROIs in the PXD selects the entire sensor for readout, as the data rate is still low. However, at higher luminosities, we will have to enable ROI selection and limit the area of the PXD sensors to be read out and whose digits are passed to clustering and tracking. Once this feature is enabled, the quality of the alignment at the moment of data-taking can have a direct and significant impact on the PXD hit efficiency.

Therefore enabled ROIs will possibly require more frequent alignment updates in the online global tag. The online global tag is currently updated only sporadically, at the request of calibration experts. That means the alignment used for data-taking can be even one year old in the current practice. Even more regular updates might not be sufficient, and some implementation of online alignment, possibly running on the express reconstruction, might be needed.

We have envisioned online alignment from the beginning, and some important features have been implemented to make it possible. In particular, the possibility to stream binary data for Millepede II as calibration objects instead of writing them into binary files directly. In such a way, several worker nodes could perform the reconstruction, filling of DQM histograms, and in addition, collecting GBL data from reconstructed tracks. The data from worker nodes can be collected and

merged in the express reconstruction similarly to DQM histograms and written to binary files, which could be used to compute the alignment, e.g., after each run finishes.

This would probably require performing full global alignment soon after the start of each data-taking period, which could serve as a baseline, on top of which the online alignment will be derived. After this alignment is added to the online global tag, regular realignments after each (long-enough) run could compare the results to the previous run and decide if a new alignment is needed in case of a significant difference. This information could also be useful to define the time-dependent configuration for the full global alignment, which would be performed later off-line in the prompt calibration loop.

To even start implementing such procedures, a discussion with DAQ and IT experts is needed. In fact, the main issue with implementing updates of the alignment online is a stringent security protocol. The DAQ infrastructure operates separately from the rest of the KEK network and the Internet. So far, we have not found a way to establish such an automated back-propagation of constants from the express reconstruction to the HLT, where extrapolation to PXD to define ROIs is performed. Whether this will require substantial changes to the DAQ remains to be understood in the future when the need for online alignment can arise. It should not be difficult to adjust the existing alignment framework for such a task.

CHAPTER 19

Discussion

In this chapter, we will summarize the alignment performance and connect some findings from different chapters to critically assess the results of this part of the thesis. We will start from the simulation studies, continue with findings from real data experience, and finish by discussing physics impact in the first precision measurements.

Simulation studies suggest an excellent control over systematic effects: coherent distortions of the VXD do not exceed 4 μm in amplitude[1], roughly the largest observed misalignment of any VXD sensor. Typically, the sensors in outer layers, particularly the forward slanted SVD sensors, suffer from the most significant weak mode contributions to their alignment precision. Precision for individual sensor parameters is $1 - 2$ μm, including sensor surface deformations. The half-shells of the PXD and SVD can be aligned with similar precision. The most significant systematic effects can be observed in the residual misalignment of the global VXD position, stemming from systematic distortions of the CDC. However, even the largest observed residual misalignment of 24 μm in the z-coordinate does not have any practical physics effects. The systematic global distortions of the CDC layers are limited to about 10 μm, where vertical deformations are the most significant due to the absence of horizontal cosmic ray tracks. Regarding global distortions at the wire level, average effects are limited to roughly 20 μm in magnitude, where radial distortions are more pronounced than $R - \phi$ deformations. Systematic effects naturally increase towards the outer edge of the CDC, where the scale of the effects is roughly twice larger than average. The alignment precision for the CDC wires is about 14 μm, slightly better at the backward end-plate.

Validation with reconstructed tracks is essential to assess alignment quality on real data. We have used multiple validation methods, which revealed issues with the baseline alignment. This alignment version does not include any wire alignment, does not use IP-constrained di-muons, and in particular, ignores any time-dependence over typically two to three weeks of data-taking. The most dangerous for physics are systematic biases, but we also observed significant instabilities in the PXD, particularly its second, incomplete layer, which can visibly degrade its resolution. We have verified that the new reprocessing alignment, which combines the full global alignment with wires followed by run-dependent alignment, fixes most of the observed issues. This last alignment stage involves a regular determination of positions of the CDC layers, VXD half-shells, and individual PXD sensors (including cubic sensor surface deformation parameters).

---

[1]To be really conservative, one should include one additional safety factor of two for a different definition of the amplitude, e.g., as a difference of the most displaced sensors in *both directions.*

We estimate that any residual misalignment contributes negligibly or only little to the overall physics performance using this approach. Given that some residual systematic effects related to the z_0 measurements are still partially present but reduced at least three times in scale by the reprocessing, we can include a safety factor of two to the maximum scale of some systematic distortions. This estimate is possible thanks to our ability to faithfully reproduce the observed biases in MC simulations with known misalignment. Namely, the z-expansion is likely only controlled to better than $\sim 10~\mu$m. This is still well below our requirements (50 μm over the entire VXD), and thus the total size of the VXD should be safely known to a precision better than 10^{-4} (also in the radial direction) after the reprocessing. This estimate includes thermal effects and fabrication precisions of the sensors, which are at most at a $\sim 10^{-5}$ level.

The data/MC correspondence of impact parameter resolutions is better than 20%, see Sec. 17.1.3. After the reprocessing, we obtain in the limit of infinite momentum (the a term):

$$\begin{aligned}
\sigma_{d_0} &= (9.8 \pm 0.7)~\mu\text{m} \\
\sigma_{z_0} &= (13.8 \pm 0.4)~\mu\text{m},
\end{aligned} \tag{19.1}$$

while for ideal MC simulation, we get

$$\begin{aligned}
\sigma_{d_0} &= (8.4 \pm 0.5)~\mu\text{m} \\
\sigma_{z_0} &= (12.9 \pm 0.3)~\mu\text{m}.
\end{aligned} \tag{19.2}$$

The difference is thus only roughly 1 μm, and in addition, part of the worse d_0 resolution is caused by too optimistic SVD simulation. Given this fact, the data/MC agreement is very good. Such an agreement was never achieved at Belle, where the data/MC difference was around $\sim 50\%$[16] (for SVD2). The resolution is slightly worse than for the older design simulations, as those include the complete second PXD layer. The degradation due to the missing PXD sensors is visible primarily in z_0. Nevertheless, the goal of roughly twice better resolution than Belle was achieved even in this limited PXD configuration (see also Fig. 20.1). Even if we attribute the total data/MC difference in quadrature to the misalignment (which is very unlikely), we obtain the corresponding contribution to the resolution to be $\sim 5~\mu$m, which was in the older MC misalignment studies still considered as an acceptably small misalignment with an almost negligible physics impact.

Regarding the CDC alignment performance, we have observed some remaining charge-dependent biases (at a smaller scale than 10^{-4}) even after the reprocessing. However, these biases were reduced considerably in the past by excluding cosmic ray tracks passing regions with a poorly described magnetic field. Nevertheless, the alignment is somehow not able to fully remove the charge asymmetry. We strongly suspect a relation of the remaining vertexing bias in z_0 versus $\tan \lambda$ and a known weak mode with possible inconsistency in the geometrical description of the CDC. As we have seen in Sec. 15.4.2, the z-biases can be coupled with z-expansion effects in the VXD at a much smaller scale. A possible deformation of the CDC end-plates could be responsible for tensions forcing the solution to settle in a wrong configuration. Most likely, the IP-constrained di-muons force the alignment solution to optimize for consistency of tracks originating from the IP, possibly by inducing some artificial alignment correction to the CDC. We

have also seen a slight bias in $\tan \lambda$ measurements introduced by the reprocessing alignment. These effects may be all connected, and a possible CDC end-plate deformation is currently the only hypothesis.

To test such a hypothesis, we will need to re-run the alignment followed by validations using various artificial deformations at the end-plates. For such studies, input from, e.g., finite element analysis of the CDC suggesting a plausible deformation model would be beneficial. The success of such a try-and-fail study is not guaranteed, and we could be easily wrong in our assumptions about the origin of the remaining systematic issues. Nevertheless, all these effects are not large. Most of the (charge-independent) momentum biases in CDC are, in fact, caused by an imperfect description of the magnetic field, and analysis-level corrections are already being applied to physics data. These corrections can also include the effect of the charge asymmetry and thus almost entirely eliminate the residual effects even when these will not be resolved at the alignment level.

Nonetheless, the remaining systematic effects are generally so small that finding their origin is difficult. In particular, one will need high statistics. For example, the z_0 versus $\tan \lambda$ bias needs more than a week worth of cosmic ray data to measure the linear slope as significant after the reprocessing.

The time-dependence is an entirely different story. In general, the detector is substantially less mechanically stable than expected. We can observe deformations of the CDC and correlate the most significant of them to relaxations and restorations of magnetic forces in the super-conducting optics. As the VXD seems to follow the CDC center movement, we believe the whole structure gets into mechanical contact with the accelerator structures through densely packed cables from the VXD. Furthermore, significant instabilities have been observed for the PXD sensors. While these are greatly reduced by the run-dependent alignment during the reprocessing, the current automated method is far from optimal. After the reprocessing, the residual instabilities are, in general, smaller than 3 μm and 5 μm in the $R - \phi$ and z-direction, respectively. As this constitutes roughly 30% of the corresponding resolutions, the impact of the alignment should be at most at $\sim 5\%$ level and thus only small if not negligible for most sensors. However, the residual instabilities in the second PXD layer are close to our upper limits in this case and have not fully recovered to the level observed in other sensors. Thus we are convinced that the current time resolution and, in particular, the simple method of run merging is not optimal and could be significantly improved, albeit the margin of the overall improvement in terms of physics performance will be relatively small. This will, however, become more and more important at higher luminosities, which should also automatically lead to more frequent re-alignments.

We leave the optimized treatment of the time-dependence and further extensions and improvements of the procedure for future studies and devote Chapter 18 to the discussion of the various ways of such endeavors. Before such optimizations, we should understand the origins of some of the instabilities much better—either to mitigate them or to give hints to the future algorithms optimizing the time-dependence. For that, extensive regular validations over long time periods correlated with various detector operation conditions will need to be monitored. For example, we see residual deformation at the sensor level in the SVD, which is supposed to be stable in each bucket. While the physics effect will be small, the

resulting evolving biases are clearly visible in the validations. These particular remaining deformations should be the main focus of future studies. At this point, we should admit that currently, the validations are mostly run by experts and not yet automated and tuned to the extent where such studies are easily feasible.

After the reprocessing, the VXD sensor instabilities look rather Gaussian and thus will not produce any significant biases when averaged over reasonably large data sets. However, the standard misalignment technique introduced in Sec 13.6 cannot reproduce the time instabilities in MC simulations. To simulate them, we have to smear every single hit. We have also tested this approach, getting an even better data/MC agreement, but there is currently no official procedure for this. On the other hand, the total effect should result in just a slightly worse resolution (and no biases), which is scalable in practically all physics analyses, typically as a parameter multiplying the width of a Gaussian resolution function (see, e.g., the resolution function $\mathcal{R}_{q\bar{q}}$ in Eq. 7.10 in part II).

What the analyses cannot (usually) absorb is a residual systematic misalignment. Therefore we have provided several of the residual misalignments and special misalignments derived from data to the analysts. The level of these misalignments has only a minor impact on the overall physics performance, like vertex resolution, but they emulate some possible remaining residual distortions. In particular, we have derived the model of prompt misalignment, reproducing the major observed vertexing biases.

Since the introduction of the misalignment scenarios, several physics analyses have investigated their impact. In particular, signal MC fits for the lifetimes of D^0 and D^+ mesons, the Λ_c^+ baryon, and the τ lepton have been examined by their respective analysis groups. This allows us to directly give concrete estimates of the alignment precision at the physics level. In general, in these analyses, the estimated alignment systematics is at a sub-micron level, i.e., less than ~ 1 fs in terms of absolute lifetimes (0.72 fs for D^0 and 1.70 fs for D^+ lifetime [5] in $D^0 \to K^-\pi^+$ and $D^+ \to K^-\pi^+\pi^-$ decays). The measured lifetimes are $\tau_{D^0} = (410.5 \pm 1.1 \pm 0.8)$ fs and $\tau_{D^+} = (1030.4 \pm 4.7 \pm 3.1)$ fs. The alignment-related systematic uncertainty corresponds to a relative precision of about 10^{-3}, which is ten times larger than our conservative estimate for absolute scale determination in the VXD. This suggests that naive estimates are not robust and explicit tests with a fully developed physics analysis workflow are necessary to assess alignment systematics at such a precision level with reasonable confidence.

So far, only the τ lifetime measurement (at roughly 200 fb^{-1}, still not unblinded) appears to be sensitive enough that the new reprocessing alignment will be absolutely essential to reach a world-leading precision measurement. This means the measurement has to be delayed until the data (mostly from the first part of the year 2021) is reprocessed. After that, the estimated τ lifetime uncertainty due to the alignment should be reduced from ~ 0.4 fs (for prompt misalignment model) to ~ 0.25 fs (for day-to-day misalignment model). Given that the current world average is $\tau_\tau = (290.3 \pm 0.5)$ fs, the efforts put into the alignment improvements finally pay off and demonstrate the importance of alignment performance in such precision measurements.

CHAPTER 20

Conclusion

The Belle II experiment is entering a phase where it can deliver the first precision measurements. In terms of detector performance, the most significant difference from the previous generation of *B*-Factories is its vertex detector that includes pixel sensors. It offers approximately twice better impact parameter resolution than available previously. Such a level of precision is only possible with an advanced alignment procedure.

This thesis presents the alignment strategy at Belle II, developed by the author from the first lines of code to the present status. This method is based on experiment-independent packages, Millepede II and General Broken Lines, integrated into the Belle II software framework. A brief historical excursion and the two following chapters took us from the first beam tests with tens of parameters of a couple of VXD sensors towards simultaneous determination of roughly sixty thousand parameters of the vertex detector and the drift chamber. To our best knowledge, using an exact solution method for an alignment problem of such a scale has no equivalent counterpart in any other HEP experiment. Belle II was the first detector to use this new Millepede II feature in actual data production. The scale of the Belle II alignment problem allows for reasonable solution times with (highly optimized) exact methods. This is a fortunate coincidence, as, for example, CMS with more than three times more parameters (due to a higher number of sensors) would require more than 3^3-times longer solution time. We have also demonstrated that approximate methods heavily utilized, e.g., at CMS, are unsuitable for the presented problem. The specific needs of the Belle II experiment have been one of the main driving factors for Millepede II development in recent years. The concept of simultaneous global and local alignment of multiple very different sub-detectors was never explored in such a setup and at a similar scale with real data. Thus it deserved extensive validation.

We have validated the procedure on simulations and evaluated its performance using some of the first data recorded by the Belle II experiment. The complete reprocessing alignment almost entirely fixed issues identified in the baseline method. The alignment meets all initial requirements and results in an overall good data/MC agreement regarding VXD (as well as CDC[1]) performance. The detector's final measured impact parameter resolutions are about twice better than what was achieved at Belle, see Fig. 20.1, and differ from the ideal MC simulation only by roughly 1 μm, see Sec. 17.1.3, demonstrating excellent alignment performance. The presented alignment procedure is fully automated and regularly performed for all newly collected collision data.

However, the ultimate precision of our vertex detector allows us to observe

---

[1]But for CDC, the calibration is the most essential for a good performance.
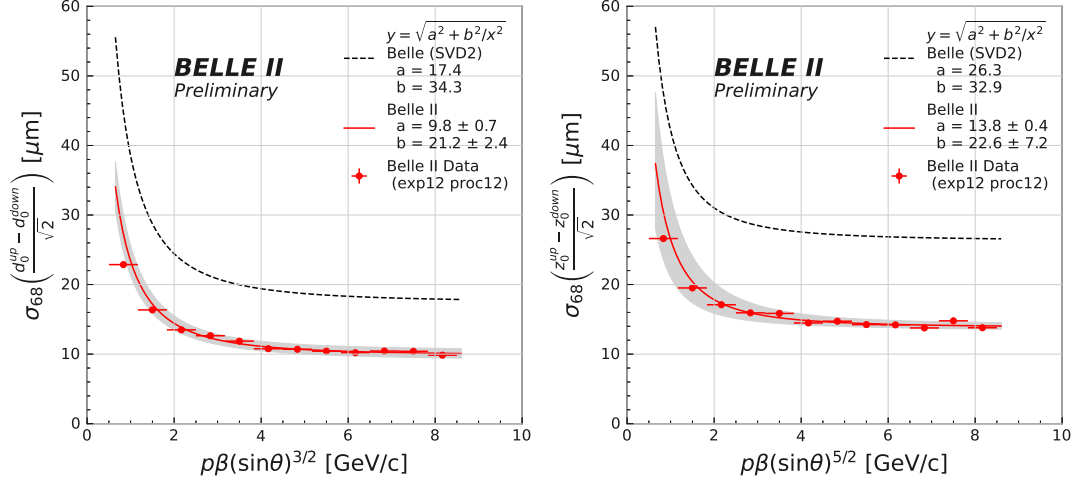
Figure 20.1: Comparison of impact parameter resolution with cosmic rays for Belle (SVD2 experiment) and Belle II (experiment 12 and `proc12` reprocessing). The quoted parameter uncertainties, also reflected in the grey $\pm 1\sigma$ error bands (obtained by sampling the estimated parameters using a normal distribution and their total estimated uncertainty), are combinations of (symmetrized) statistical and systematic errors.

some remaining inconsistencies. To assess the physics impact of the residual misalignment, we provided dedicated misalignment payloads to analysts. These misalignments include estimates derived from real data as well as several MC studies. Apart from the measurements mentioned in the Discussion, analysis of B^0 meson lifetime and mixing frequency Δm_d recently evaluated alignment uncertainties using the provided misalignments, observing smaller effects than in equivalent analyses at Belle, even with the prompt misalignment model.

The recently published D^0 and D^+ lifetime measurements [5], with a significant contribution from the author, are the most precise to date and consistent with the world average. They also demonstrate the alignment and vertex detector performance. Contrary to LHCb, which measures lifetimes relative to τ_{D^0}, the small IP at Belle II and clean collision environment allow for absolute lifetime measurements at unprecedented precision in some cases. Similar measurements of Λ_c^+ baryon or τ lepton lifetimes should follow soon. Thus the presented alignment procedure is already of great physics relevance even at these preliminary stages of the experiment with a limited data sample and (probably) long before the advertised high-precision time-dependent CP-violation measurements can be performed.

Epilogue

The measurement in the first part of this thesis is in agreement with the SM predictions and previous results. The combination of our and BaBar measurement [48] improves the consistency of CP-violation parameters between $B^0 \to \eta_c K_S^0$ and other charmonia, confirming the SM. This should not come as a disappointment, quite the opposite. The smallness of possible NP deviations is the primary motivation for the new generation of precision experiments, and the Belle II experiment has just entered the game, with the most exciting times still ahead. In the author's humble opinion, before any definitive sign of NP is confirmed, one should enjoy the tremendous success of the Standard Model. In addition, one should appreciate the ingenuity with which experiments sensitive to the subtle effects of CP-violation were designed, as well as the complexity of the analyses needed to extract these fundamental SM parameters and the enormous amount of collaborative effort essential to make such measurements possible.

The second part of this thesis is the main contribution to the search for NP. Precise and unbiased vertex measurements highly depend on the quality of the vertex detector alignment. While high statistics measurements of time-dependent CP-violation are still ahead, the vertexing capabilities of Belle II could be already demonstrated on the world's most precise measurement of the D^0 and D^+ meson lifetimes [5]. This result was achieved with just a preliminary prompt alignment, which can be further improved, as we demonstrated. The improved alignment quality is now provided automatically, already in the prompt calibration loop. We are confident that already at this initial stage of the experiment, the alignment quality is good enough to perform future CP-violation measurements with the final precision. The first world-leading physics results come as a great satisfaction for all the effort and time spent on many unsuccessful experiments, especially with real data, while trying to understand the issues and improve the performance.

The alignment efforts also resulted in the further spread of the knowledge among not only our Prague Belle II members and led to several academic degrees awarded to the author's younger colleagues for alignment-related studies. These studies would be impossible without all the developments presented in this thesis.

Finally, let us note that the alignment of HEP detectors is a never-ending story and can always be improved. This should go hand-in-hand with a better understanding of the detector, and we have numerous ideas on what to do next. This effort, which will surely produce many interesting results and insights, is left for further research.

Bibliography

[1] G. Aad et al. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". *Physic Letters B* 716.1 (2012), pp. 1–29. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2012.08.020.

[2] A. Abashian et al. "The Belle detector". *Nucl. Instrum. Methods Phys. Res. Sect. A* 479 (Feb. 2002), pp. 117–232. DOI: 10.1016/S0168-9002(01)02013-7.

[3] B. Aubert et al. "The BaBar detector". *Nucl. Instrum. Methods Phys. Res. Sect. A* 479.1 (2002). Detectors for Asymmetric B-Factories, pp. 1–116. ISSN: 0168-9002. DOI: 10.1016/S0168-9002(01)02012-5.

[4] A. Canepa. "Searches for supersymmetry at the Large Hadron Collider". *Reviews in Physics* 4 (2019), p. 100033. ISSN: 2405-4283. DOI: 10.1016/j.revip.2019.100033.

[5] F. Abudinén et al. "Precise measurement of the D^0 and D^+ lifetimes at Belle II". *Phys. Rev. Lett.* 127 (21 Nov. 2021), p. 211801. DOI: 10.1103/PhysRevLett.127.211801.

[6] T. Hurth and F. Mahmoudi. "Colloquium: New physics search with flavor in the LHC era". *Rev. Mod. Phys.* 85 (2 May 2013), pp. 795–808. DOI: 10.1103/RevModPhys.85.795.

[7] Belle II Collaboration. *Belle II Luminosity.* [online] Accessed 26 Jun. 2022. URL: https://confluence.desy.de/display/BI/Belle+II+Luminosity.

[8] A. J. Bevan et al. "The Physics of the B Factories". *The European Physical Journal C* 74.11 (Nov. 2014). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-014-3026-9.

[9] M. Baszczyk et al. *SuperB Technical Design Report.* 2013. arXiv: 1306.5655 [physics.ins-det].

[10] T. Abe et al. *Belle II Technical Design Report.* KEK Report 2010-1. 2010. arXiv: 1011.0352 [physics.ins-det].

[11] E. Kou et al. "The Belle II Physics Book". *Progress of Theoretical and Experimental Physics* 2019.12 (Dec. 2019). DOI: 10.1093/ptep/ptz106.

[12] R. Aaij et al. "Test of lepton universality in beauty-quark decays". *Nature Physics* 18.3 (Mar. 2022), pp. 277–282. ISSN: 1745-2481. DOI: 10.1038/s41567-021-01478-8.

[13] Belle II Collaboration. *SuperKEKB and Belle II.* [online] Accessed 20 Jun. 2022. URL: https://www.belle2.org.

[14] M. Ritter. "Measurement of the branching fraction and time-dependent CP asymmetry in $B^0 \to D^{*-}D^{*+}K_S^0$ decays at the Belle experiment". PhD thesis. Munich: Ludwig-Maximilians-Universitat Munchen, Feb. 2014.

[15] S. Hashimoto et al. *Letter of intent for KEK Super B factory*. KEK Report 04-4. June 2004. URL: http://superb.kek.jp/documents/loi/img/LoI_detector.pdf.

[16] H. Ishino et al. *Alignment method for the SVD2 and Its Performance*. Belle internal note BN715. 2004.

[17] C. Irmler et al. "Origami chip-on-sensor design: Progress and new developments". *Journal of Instrumentation* 8 (Jan. 2013), p. C01014. DOI: 10.1088/1748-0221/8/01/C01014.

[18] J. Kemmer and G. Lutz. "New detector concepts". *Nucl. Instrum. Methods Phys. Res. Sect. A* 253.3 (1987), pp. 365–377. ISSN: 0168-9002. DOI: doi.org/10.1016/0168-9002(87)90518-3.

[19] Max-Planck Institute of Physics. *Belle II: A pixel vertex detector for an updated accelerator*. [online] Accessed 20 Jun. 2022. URL: https://www.mpp.mpg.de.

[20] H. Ye et al. "Commissioning and performance of the Belle II pixel detector". *Nucl. Instrum. Methods Phys. Res. Sect. A* 987 (2021), p. 164875. ISSN: 0168-9002. DOI: 0.1016/j.nima.2020.164875.

[21] D. V. Thanh et al. "Calibration for Belle II Central Drift Chamber". *2017 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*. 2017, pp. 1–6. DOI: 10.1109/NSSMIC.2017.8532695.

[22] M. Milesi. *Particle identification at Belle II with the electromagnetic calorimeter*. Belle II StarterKit Workshop (Jan. 2020). [online] Accessed 13 May 2022. URL: https://indico.belle2.org/event/1307/sessions/378/attachments/3070/5670/mmilesi_ECLPid_B2SKW_2020.pdf.

[23] R. Itoh. *The Performance of Belle II High Level Trigger in the First Physics Run*. CHEP2019 (Nov. 2019) [online] Accessed 14 June 2022. URL: https://indico.cern.ch/event/773049/contributions/3474295/attachments/1937594/3212655/Belle2-HLT-Itoh.pdf.

[24] P. Dirac. "The quantum theory of the electron". *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 117.778 (1928), pp. 610–624. DOI: 10.1098/rspa.1928.0023.

[25] G. C. Branco, L. Lavoura, and J. P. Silva. *CP Violation*. Vol. 103. International series of monographs on physics. London: Clarendon Press, 1999. ISBN: 9780198503996.

[26] T. D. Lee and C.-N. Yang. "Question of Parity Conservation in Weak Interactions". *Phys. Rev.* 104 (1956), pp. 254–258. DOI: 10.1103/PhysRev.104.254.

[27] C. S. Wu et al. "Experimental Test of Parity Conservation in β Decay". *Phys. Rev.* 105 (1957), pp. 1413–1414. DOI: 10.1103/PhysRev.105.1413.

[28]   A. D. Sakharov. "Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe". *Pisma Zh. Eksp. Teor. Fiz.* 5 (1967), pp. 32–35. DOI: `10.1070/PU1991v034n05ABEH002497`.

[29]   I. I. Bigi. *CP Violation*. Ed. by A. I. Sanda. 2nd ed. Cambridge monographs on particle physics, nuclear physics, and cosmology. Cambridge: Cambridge University Press, 2009. ISBN: 978-0-511-58069-7.

[30]   F. Abe et al. "Observation of Top Quark Production in $p\bar{p}$ Collisions with the Collider Detector at Fermilab". *Physical Review Letters* 74.14 (Apr. 1995), pp. 2626–2631. ISSN: 1079-7114. DOI: `10.1103/physrevlett.74.2626`.

[31]   Wikimedia Commons. *Standard Model of Elementary Particles*. [online] Accessed 04 Jun. 2022. 2019. URL: `https://en.wikipedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg`.

[32]   J. Hořejší. *Fundamentals of Electroweak Theory*. Prague: Karolinum Press, 2002. ISBN: 9788024606392.

[33]   M. Kobayashi and T. Maskawa. "CP-Violation in the Renormalizable Theory of Weak Interaction". *Progress of Theoretical Physics* 49.2 (Feb. 1973), pp. 652–657. ISSN: 0033-068X. DOI: `10.1143/PTP.49.652`.

[34]   P. A. Zyla et al. "Review of Particle Physics". *Progress of Theoretical and Experimental Physics* 2020.8 (Aug. 2020). 083C01. ISSN: 2050-3911. DOI: `10.1093/ptep/ptaa104`.

[35]   A. J. Bevan et al. "The Physics of the B Factories". *The European Physical Journal C* 74.11 (Nov. 2014). ISSN: 1434-6052. DOI: `10.1140/epjc/s10052-014-3026-9`.

[36]   M. Tanabashi et al. "Review of Particle Physics". *Phys. Rev. D* 98 (3 Aug. 2018), p. 030001. DOI: `10.1103/PhysRevD.98.030001`.

[37]   Y. Grossman, A. L. Kagan, and Z. Ligeti. "Can the CP asymmetries in $B \to \Psi K_S$ and $B \to \Psi K_L$ differ?" *Physics Letters B* 538.3 (2002), pp. 327–334. ISSN: 0370-2693. DOI: `10.1016/S0370-2693(02)02027-0`.

[38]   J. Silva. "Use of the reciprocal basis in neutral meson mixing". *Physical Review D* 62.11 (Nov. 2000). ISSN: 1089-4918. DOI: `10.1103/physrevd.62.116008`.

[39]   W. Kwong, J. L. Rosner, and C. Quigg. "Heavy-Quark Systems". *Annual Review of Nuclear and Particle Science* 37.1 (1987), pp. 325–382. DOI: `10.1146/annurev.ns.37.120187.001545`.

[40]   Zbyněk Drásal. "Measurement of Time-Dependent CP Violation in $B^0 \to \eta_c K_S^0$ at Belle Experiment, Optimization Studies of the Belle II Vertex Detector". PhD thesis. Prague: Charles University, 2014.

[41]   P. Frings, U. Nierste, and M. Wiebusch. "Penguin Contributions to CP Phases in $B_{d,s}$ Decays to Charmonium". *Physical Review Letters* 115.6 (Aug. 2015). ISSN: 1079-7114. DOI: `10.1103/physrevlett.115.061802`.

[42]   D. Besson and T. Skwarnicki. "Upsilon Spectroscopy: Transitions in the Bottomonium System". *Annual Review of Nuclear and Particle Science* 43.1 (1993), pp. 333–378. DOI: `10.1146/annurev.ns.43.120193.002001`.

[43] A. Einstein, B. Podolsky, and N. Rosen. "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?" *Phys. Rev.* 47 (10 May 1935), pp. 777–780. DOI: 10.1103/PhysRev.47.777.

[44] J. E. Augustin, A. M. Boyarski, and M. Breidenbach. "Discovery of a Narrow Resonance in e^+e^- Annihilation". *Phys. Rev. Lett.* 33 (23 Dec. 1974), pp. 1406–1408. DOI: 10.1103/PhysRevLett.33.1406.

[45] A. Hosaka et al. "Exotic Hadrons with Heavy Flavors - X, Y, Z and Related States". *Progress of Theoretical and Experimental Physics* 2016 (Mar. 2016). DOI: 10.1093/ptep/ptw045.

[46] F. Jackson. "A measurement of the $B \to \eta_c K$ branching fractions using the BaBar detector". PhD thesis. Manchester: University of Manchester, 2003.

[47] A. Vinokurova et al. "Study of $B^\pm \to K^\pm (K_S K\pi)^0$ decay and determination of η_c and $\eta_c(2S)$ parameters". *Phys. Lett. B* 706 (2011), pp. 139–149. DOI: 10.1016/j.physletb.2011.11.014.

[48] B. Aubert et al. "Measurement of time-dependent CP asymmetry in $B^0 \to c\bar{c}K^{(*)0}$ decays". *Phys. Rev. D* 79 (7 Apr. 2009), p. 072009. DOI: 10.1103/PhysRevD.79.072009.

[49] Y. S. Amhis et al. "Averages of b-hadron, c-hadron, and τ-lepton properties as of 2018". *Eur. Phys. J.* C81 (2021), p. 226. DOI: 10.1140/epjc/s10052-020-8156-7. URL: https://www.slac.stanford.edu/xorg/hflav/triangle/latest/.

[50] D. Atwood and G. Hiller. "Implications of nonstandard CP violation in hadronic B decays" (July 2003). arXiv: hep-ph/0307251.

[51] B. Casey. *HadronB*. Belle internal note BN390. 2001.

[52] R. Brun et al. *GEANT 3: user's guide Geant 3.10, Geant 3.11; rev. version.* Geneva: CERN, 1987. URL: https://cds.cern.ch/record/1119728.

[53] F. Fang. *Study of $K_S^0 \to \pi^+\pi^-$ Selection*. Belle internal note BN323. 2000.

[54] H. Kakuno et al. "Neutral B flavor tagging for the measurement of mixing-induced CP violation at Belle". *Nucl. Instrum. Methods Phys. Res. Sect. A* 533.3 (Nov. 2004), pp. 516–531. ISSN: 0168-9002. DOI: 10.1016/j.nima.2004.06.159.

[55] I. Adachi et al. "Precise Measurement of the CP Violation Parameter $\sin 2\phi_1$ in $B^0 \to (c\bar{c})K^0$ Decays". *Physical Review Letters* 108.17 (Apr. 2012). ISSN: 1079-7114. DOI: 10.1103/physrevlett.108.171802.

[56] G. C. Fox and S. Wolfram. "Observables for the Analysis of Event Shapes in e^+e^- Annihilation and Other Processes". *Phys. Rev. Lett.* 41 (23 Dec. 1978), pp. 1581–1585. DOI: 10.1103/PhysRevLett.41.1581.

[57] J. Beringer et al. "Review of Particle Physics". *Phys. Rev. D* 86 (1 July 2012), p. 010001. DOI: 10.1103/PhysRevD.86.010001.

[58] T. M. Karbach and M. Schlupp. *Constraints on Yield Parameters in Extended Maximum Likelihood Fits*. 2012. arXiv: 1210.7141 [physics.data-an].

[59] A. N. Vinokurova, A. S. Kuzmin, and S. I. Eidelman. *Estimation of the interference between signal and non-resonant component in the $B^\pm \to K^\pm (K_S K \pi)^0$ decay.* Belle internal note BN1082. 2010.

[60] J. P. Lees et al. "Measurement of the $I = 1/2$ $K\pi$ $\mathcal{S}$-wave amplitude from Dalitz plot analyses of $\eta_c \to K\overline{K}\pi$ in two-photon interactions". *Phys. Rev. D* 93 (1 Jan. 2016), p. 012005. DOI: 10.1103/PhysRevD.93.012005.

[61] Belle Collaboration. *Resolution and Convolution package (tatami).* Belle internal web page. [online] Accessed 05 Jun. 2022. 2002. URL: https://belle.kek.jp/~nakadair/private/libcnvl.html.

[62] H. Tajima et al. "Proper-time resolution function for measurement of time evolution of B mesons at the KEK B-Factory". *Nucl. Instrum. Methods Phys. Res. Sect. A* 533.3 (2004), pp. 370–386. ISSN: 0168-9002. DOI: 10.1016/j.nima.2004.07.199.

[63] G. Punzi. "Comments on Likelihood fits with variable resolution". *arXiv e-prints* (Jan. 2004). arXiv: physics/0401045 [physics.data-an].

[64] O. Long et al. "Impact of tag-side interference on time dependent CP asymmetry measurements using coherent B^0 anti-B^0 pairs". *Phys. Rev. D* 68 (2003), p. 034010. DOI: 10.1103/PhysRevD.68.034010.

[65] K. Hara. *Study of tag-side interference effect in time-dependent CP violating asymmetry measurements.* Belle internal note BN872. 2006.

[66] Y. Yusa and K. Miyabayashi. *Measurement of time-dependent CP violation in $B^0 \to (c\bar{c})K^0$ decays with 772 M $B\overline{B}$.* Belle internal note BN1149. 2011.

[67] K. Nishimura. *Number of B events in Hadron B(J).* Belle internal web page. [online] Accessed 05 Jun. 2022. 2009. URL: https://belle.kek.jp/secured/nbb/nbb.html.

[68] M. Niiyama et al. *KID.* Belle internal web page. [online] Accessed 05 Jun. 2022. 2009. URL: https://belle.kek.jp/group/pid_joint/kid/kid.html.

[69] B. Bhuyan. *High P_T Tracking Efficiency Using Partially Reconstructed D^* Decays.* Belle internal note BN1165. 2010.

[70] D. N. Brown et al. "Local alignment of the BaBar Silicon Vertex Tracking detector". *Nucl. Instrum. Methods Phys. Res. Sect. A* 603.3 (May 2009), pp. 467–484. ISSN: 0168-9002. DOI: 10.1016/j.nima.2009.02.001.

[71] J. Kandra. "Simulation of Belle II physics events and performance tests of reconstruction software". MA thesis. Prague: Charles University, 2016.

[72] J. Kandra and T. Bilka. "Alignment and physics performance of the Belle II vertex detector". *PoS* FPCP2017 (2017), p. 053. DOI: 10.22323/1.304.0053.

[73] C. Kleinwort. "H1 alignment experience". *Proceedings of the first LHC Detector Alignment Workshop* (2007), pp. 41–49. DOI: 10.5170/CERN-2007-004.41.

[74] The CMS collaboration. "Alignment of the CMS tracker with LHC and cosmic ray data". *Journal of Instrumentation* 9.06 (June 2014), P06009. DOI: 10.1088/1748-0221/9/06/p06009.

[75] R. Frühwirth, T. Todorov, and M. Winkler. "Estimation of detector alignment parameters using the Kalman filter with annealing". *Journal of Physics G: Nuclear and Particle Physics* 29.3 (Feb. 2003), pp. 561–574. DOI: 10.1088/0954-3899/29/3/309.

[76] C. Kleinwort. Personal communication. Jan. 13, 2022.

[77] T. Bilka. "Simulation and analysis of tests of Belle II vertex detector sensors". MA thesis. Prague: Charles University, 2014.

[78] V. Blobel and C. Kleinwort. "A New Method for the High-Precision Alignment of Track Detectors". *Proceedings of the Conference on Advanced Statistical Techniques in Particle Physics, Durham (March 2002)* (Sept. 2002). DOI: 10.48550/ARXIV.HEP-EX/0208021.

[79] V. Blobel. "Software alignment for tracking detectors". *Nucl. Instrum. Methods Phys. Res. Sect. A* 566.1 (2006). TIME 2005, pp. 5–13. ISSN: 0168-9002. DOI: 10.1016/j.nima.2006.05.157.

[80] C. Kleinwort. *Millepede-II*. [online] Accessed 02 Jan. 2022. 2022. URL: https://www.desy.de/~kleinwrt/MP2/doc/html/index.html.

[81] M. Stoye. "Calibration and Alignment of the CMS Silicon Tracking Detector". PhD thesis. Hamburg: University of Hamburg, July 2007.

[82] C. Kleinwort. *Millepede-II*. [online] Accessed 30 Jun. 2022. 2022. URL: https://gitlab.desy.de/claus.kleinwort/millepede-ii.git.

[83] E. Anderson et al. *LAPACK Users' Guide*. Third edition. Philadelphia, PA: Society for Industrial and Applied Mathematics, 1999. ISBN: 0-89871-447-8.

[84] Intel Corporation. *Intel oneAPI Math Kernel Library*. [online] Accessed 30 Jun. 2022. 2022. URL: https://www.intel.com.

[85] P. Connor. "Tracker alignment of the CMS detector". *PoS* ICHEP2020 (2021), p. 729. DOI: 10.22323/1.390.0729.

[86] J. Rauch and T. Schlüter. "GENFIT - A generic track-fitting toolkit". *Journal of Physics: Conference Series* 608 (Oct. 2014). DOI: 10.1088/1742-6596/608/1/012042.

[87] T. Bilka et al. *Implementation of GENFIT2 as an experiment independent track-fitting framework*. 2019. arXiv: 1902.04405 [physics.data-an].

[88] V. Bertacchi et al. "Track finding at Belle II". *Computer Physics Communications* 259 (Feb. 2021), p. 107610. ISSN: 0010-4655. DOI: 10.1016/j.cpc.2020.107610.

[89] V. Blobel. "A new fast track-fit algorithm based on broken lines". *Nucl. Instrum. Methods Phys. Res. Sect. A* 566.1 (2006). TIME 2005, pp. 14–17. ISSN: 0168-9002. DOI: 10.1016/j.nima.2006.05.156.

[90] C. Kleinwort. "General broken lines as advanced track fitting method". *Nucl. Instrum. Methods Phys. Res. Sect. A* 673 (2012), pp. 107–110. ISSN: 0168-9002. DOI: 10.1016/j.nima.2012.01.024.

[91] C. Kleinwort. *General Broken Lines*. [online] Accessed 02 Jan. 2022. URL: https://www.desy.de/~kleinwrt/GBL/doc/cpp/html/.

[92] C. Kleinwort. *General Broken Lines as advanced track model. Draft Manual.* [online] Accessed 02 Jan. 2022. 2013. URL: `https://www.terascale.de/sites/site_terascale/content/e1443/e295960/e296478/Gbl_man.pdf`.

[93] G. R. Lynch and O. I. Dahl. "Approximations to multiple Coulomb scattering". *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* 58.1 (1991), pp. 6–10. ISSN: 0168-583X. DOI: `10.1016/0168-583X(91)95671-Y`.

[94] E. Widl and R. Frühwirth. *Representation and Estimation of Trajectories from Two-body Decays.* Tech. rep. Geneva: CERN, Oct. 2007. URL: `http://cds.cern.ch/record/1073690`.

[95] C. Hagmann, D. Lange, and D. Wright. "Cosmic-ray shower generator (CRY) for Monte Carlo transport codes". *2007 IEEE Nuclear Science Symposium Conference Record.* Vol. 2. 2007, pp. 1143–1146. DOI: `10.1109/NSSMIC.2007.4437209`.

[96] G. D. Marino and K. Trabelsi. *HLT skims.* Belle II internal note. BELLE2-NOTE-TE-2020-018. Oct. 2020.

[97] V. Karimaki et al. *Sensor Alignment by Tracks.* Tech. rep. Geneva: CERN, June 2003. URL: `https://cds.cern.ch/record/619975`.

[98] R. Brun and F. Rademakers. "ROOT: An object oriented data analysis framework". *Nucl. Instrum. Meth. A* 389 (1997). Ed. by M. Werlen and D. Perret-Gallix, pp. 81–86. DOI: `10.1016/S0168-9002(97)00048-X`.

[99] S. Agostinelli, J. Allison, and K. Amako. "GEANT4-a simulation toolkit". *Nucl. Instrum. Methods Phys. Res. Sect. A* 506 (July 2003), p. 250. DOI: `10.1016/S0168-9002(03)01368-8`.

[100] R. Brun. *ROOT TGeoMatrix.* [online] Accessed 07 Nov. 2021. URL: `https://root.cern/doc/master/classTGeoMatrix.html`.

[101] T. Kuhr et al. "The Belle II Core Software". *Computing and Software for Big Science* 3.1 (Nov. 2018). ISSN: 2510-2044. DOI: `10.1007/s41781-018-0017-9`.

[102] J.-F. Krohn et al. "Global decay chain vertex fitting at Belle II". *Nucl. Instrum. Methods Phys. Res. Sect. A* 976 (Oct. 2020), p. 164269. ISSN: 0168-9002. DOI: `10.1016/j.nima.2020.164269`.

[103] W. Waltenberger. "RAVE – A Detector-Independent Toolkit to Reconstruct Vertices". *Nuclear Science, IEEE Transactions on Nuclear Science* 58 (May 2011), pp. 434–444. DOI: `10.1109/TNS.2011.2119492`.

[104] Apache Software Foundation. *Apache Airflow.* [online] Accessed 17 Dec. 2021. URL: `https://airflow.apache.org/`.

[105] J. Kandra et al. *Calibration and alignment of the Belle II tracker.* 2019. arXiv: `1910.06289 [physics.ins-det]`.

[106] J. Kandra and T. Bilka. *B lifetime and $B^0 - \bar{B}^0$ mixing results from early Belle II data.* 2019. arXiv: `1906.08940 [hep-ex]`.

[107]   A. Tumasyan et al. "Strategies and performance of the CMS silicon tracker alignment during LHC Run 2". *Nucl. Instrum. Meth. A* 1037 (2022), p. 166795. DOI: `10.1016/j.nima.2022.166795`.

[108]   T. Bilka. *Alignment with Millepede II using General Broken Lines for the Belle II Vertex Detector*. [online] Accessed 30 Jun. 2022. 2013. URL: `https://www.desy.de/f/students/2013/reports/bilka.pdf.gz`.

[109]   T. Bilka et al. "Demonstrator of the Belle II Online Tracking and Pixel Data Reduction on the High Level Trigger System". *IEEE Transactions on Nuclear Science* 62.3 (June 2015), pp. 1155–1161. ISSN: 1558-1578. DOI: `10.1109/tns.2015.2419879`.

[110]   T. Bilka. "The beam test measurements of the Belle II vertex detector modules". *JINST* 12.03 (2017). Ed. by Pietro Govoni et al., p. C03002. DOI: `10.1088/1748-0221/12/03/C03002`.

[111]   T. Bilka and J. Kandra. "Millepede alignment of the Belle 2 sub-detectors after first collisions". *38th International Symposium on Physics in Collision*. Dec. 2018. arXiv: `1812.05340 [physics.ins-det]`.

[112]   G. de Marino. "D^0 Lifetime Measurement with Belle II Early Data". MA thesis. Pisa: Universita di Pisa, 2019.

[113]   S. Kurz. "Performance of Belle II tracking on collision data". Connecting The Dots 2020. Apr. 2020. URL: `https://bib-pubdb1.desy.de/record/451990`.

[114]   J. Kandra and T. Bilka. "Alignment Studies at Belle II Vertex Detector". *Acta Physica Polonica B*. Vol. 51. 6. 2020. URL: `https://www.actaphys.uj.edu.pl/fulltext?series=Reg&vol=51&page=1385`.

[115]   T. Bilka. "Belle II Vertex Detector Performance". *Proceedings of The 28th International Workshop on Vertex Detectors*. Vol. 373. 2020. DOI: `10.22323/1.373.0001`.

[116]   V. Blobel. *Millepede II in 2009*. [online] Accessed 30 Jun. 2022. June 2009. URL: `https://indico.cern.ch/event/50502/contributions/1183071/attachments/964111/1368903/cernali.pdf`.

[117]   T. Bilka et al. "Simultaneous Global and Local Alignment of the Belle II Tracking Detectors". *EPJ Web Conf.* 251 (2021), p. 03028. DOI: `10.1051/epjconf/202125103028`.

[118]   G. Aad et al. "Alignment of the ATLAS Inner Detector in Run 2". *The European Physical Journal C* 80 (Dec. 2020). DOI: `10.1140/epjc/s10052-020-08700-6`.

[119]   Y. Guan et al. *BKLM Alignment*. Belle II internal note BELLE2-NOTE-TE-2015-035. 2015.

[120]   T. Bilka and J. Kandra. "Calibration and Alignment Framework for the Belle II detector". *Journal of Physics: Conference Series* 1085 (Sept. 2018), p. 032023. DOI: `10.1088/1742-6596/1085/3/032023`.

[121]   T. Bilka et al. "Alignment for the first precision measurements at Belle II". *EPJ Web Conf.* 245 (2020), p. 02023. DOI: `10.1051/epjconf/202024502023`.

List of Figures

List of Tables

List of Abbreviations

ACC	Aerogel Cherenkov Counter
ARICH	Aerogel Ring Imaging Cherenkov detector
basf	Belle Analysis Software Framework
basf2	Belle II Analysis Software Framework
BKLM	Barrel K_L and Muon detector
CAF	Calibration and Alignment Framework
CDC	Central Drift Chamber
CKF	Combinatorial Kalman Filter
CKM	Cabbibo-Kobayashi-Maskawa (matrix)
CM	Center-of-Mass (frame)
CMS	Compact Muon Solenoid (detector at the LHC)
***CP*V**	*CP* Violation
DAF	Deterministic Annealing Filter
DAQ	Data Acquisition System
DEPFET	DEPleted Field Effect Transistor
DSSD	Double-sided Silicon Strip Detector
DST	Data Summary Tables
ECL	Electromagnetic Calorimeter
EKLM	End-cap K_L and Muon detector
FET	Field Effect Transistor
FOM	Figure Of Merit
FSR	Final State Radiation
GBL	General Broken Lines
GDL	Global Decision Logic
GENFIT	Generic Track Fitting Toolkit
GSIM	Geant(3) Simulation (in basf at Belle)
GT	Global Tag
HEP	High Energy Physics
HER	High Energy Ring
HLT	High Level Trigger
IG	Internal Gate
IOV	Interval Of Validity
IP	Interaction Point
ISR	Initial State Radiation
KEK	High Energy Accelerator Research Organization

KLM	K_L and Muon detector
KM	Kobayashi-Maskawa (mechanism of CP violation)
LER	Low Energy Ring
LHC	Large Hadron Collider
MC	Monte Carlo
MINRES	MINimization of RESiduals
MS	Multiple Scattering
NDF	Number of Degrees of Freedom
NP	New Physics
PDF	Probability Density Function
PDG	Particle Data Group (The Review of Particle Physics)
PID	Particle Identification (detector)
POCA	Point Of Closest Approach
PXD	PiXel Detector
QCD	Quantum Chromodynamics
QCS	Superconducting optics system at SuperKEKB
QED	Quantum Electrodynamics
RMS	Root Mean Square
ROI	Range Of Interest
RPC	Resistive Plate Chamber
SM	Standard Model
SVD	Strip Vertex Detector
TDCPV	Time-Dependent CP Violation
TOF	Time-Of-Flight detector
TOP	Time-Of-Propagation detector
UT	Unitary Triangle
VXD	VerteX Detector

List of Publications

The list of the main publications (ordered chronologically) written exclusively by the author or with a direct and significant author's contribution follows. Only publications directly related to the thesis topic and referenced in the main text are listed as per formal requirements.

- T. Bilka et al. "Demonstrator of the Belle II Online Tracking and Pixel Data Reduction on the High Level Trigger System". *IEEE Transactions on Nuclear Science* 62.3 (June 2015), pp. 1155–1161. ISSN: 1558-1578. DOI: `10.1109/tns.2015.2419879`

- T. Bilka. "The beam test measurements of the Belle II vertex detector modules". *JINST* 12.03 (2017). Ed. by Pietro Govoni et al., p. C03002. DOI: `10.1088/1748-0221/12/03/C03002`

- T. Bilka and J. Kandra. "Calibration and Alignment Framework for the Belle II detector". *Journal of Physics: Conference Series* 1085 (Sept. 2018), p. 032023. DOI: `10.1088/1742-6596/1085/3/032023`

- T. Bilka and J. Kandra. "Millepede alignment of the Belle 2 sub-detectors after first collisions". *38th International Symposium on Physics in Collision.* Dec. 2018. arXiv: `1812.05340 [physics.ins-det]`

- E. Kou et al. "The Belle II Physics Book". *Progress of Theoretical and Experimental Physics* 2019.12 (Dec. 2019). DOI: `10.1093/ptep/ptz106`

- T. Bilka. "Belle II Vertex Detector Performance". *Proceedings of The 28th International Workshop on Vertex Detectors.* Vol. 373. 2020. DOI: `10.22323/1.373.0001`

- T. Bilka et al. "Alignment for the first precision measurements at Belle II". *EPJ Web Conf.* 245 (2020), p. 02023. DOI: `10.1051/epjconf/202024502023`

- T. Bilka et al. "Simultaneous Global and Local Alignment of the Belle II Tracking Detectors". *EPJ Web Conf.* 251 (2021), p. 03028. DOI: `10.1051/epjconf/202125103028`

- V. Bertacchi et al. "Track finding at Belle II". *Computer Physics Communications* 259 (Feb. 2021), p. 107610. ISSN: 0010-4655. DOI: `10.1016/j.cpc.2020.107610`

- F. Abudinén et al. "Precise measurement of the D^0 and D^+ lifetimes at Belle II". *Phys. Rev. Lett.* 127 (21 Nov. 2021), p. 211801. DOI: `10.1103/PhysRevLett.127.211801`

APPENDIX A

Summary of Model Parameters and Fit Results

This appendix summarizes the fit results and parameters of the final 2D and 4D fits to the Belle data. The 2D fit parameters are shown in Tables A.2 and A.3 for the control and signal mode, respectively. The final 4D fit parameters are shown in Tables A.4 and A.5 for the control and signal mode, respectively. Whether a parameter is determined in the fit or fixed from some previous (MC or data) fit or PDG is denoted by notation explained in Table A.1.

Source	Description
F	Floating fit parameter
PDG	Physics parameter fixed from PDG
MC Truth	Parameter fixed from MC truth information (not fitted)
2D MC	Signal MC fit in $\Delta E \times M(K_S^0 K^+ \pi^-)$
2D M SB	Data fit in $\Delta E \times M(K_S^0 K^+ \pi^-)$ in the $M(K_S^0 K^+ \pi^-)$ sideband
4D MC	Signal MC fit in $\Delta E \times M(K_S^0 K^+ \pi^-) \times \Delta t \times q \times l$
4D M_{bc} SB	Data fit in $\Delta E \times M(K_S^0 K^+ \pi^-) \times \Delta t \times q \times l$ in M_{bc} sideband
4D $B\overline{B}$ MC	Fit in $\Delta E \times M(K_S^0 K^+ \pi^-) \times \Delta t \times q \times l$ to generic $B\overline{B}$ MC ($\star$)
2D CS	Full 2D fit in $\Delta E \times M(K_S^0 K^+ \pi^-)$ of the control sample
4D CS	Full 4D fit in $\Delta E \times M \times \Delta t \times q \times l$ of the control sample

Table A.1: Description of sources of values for floating/fixed parameters of the data model. ($\star$)—with removed signal and peaking background contributions.

Table A.2: 2D model parameters and fit results for the control mode $B^+ \to \eta_c(K_S^0 K^+ \pi^-)K^+$.

Parameter (dim) [unit]	Source	SVD1 Value	SVD2 Value
$\mu_{\mathrm{main}}(\Delta E)$ [GeV]	2D MC	0.001260 ± 0.000060	0.000922 ± 0.000049
$\mu_{\mathrm{main}}^{\mathrm{CF}}(\Delta E)$ [GeV]	F	-0.0002 ± 0.0011	-0.00061 ± 0.00040
$\sigma_{\mathrm{main}}(\Delta E)$ [GeV]	2D MC	0.009242 ± 0.000057	0.009284 ± 0.000045
$\sigma_{\mathrm{main}}^{\mathrm{CF}}(\Delta E)$	F	1.26 ± 0.10	1.111 ± 0.039
$\mu_{\mathrm{tail}}^{1}(\Delta E)$ [GeV]	2D MC	-0.00248 ± 0.00030	-0.00050 ± 0.00029
$\sigma_{\mathrm{tail}}^{1}(\Delta E)$	2D MC	1.557 ± 0.033	1.520 ± 0.030
$f_1(\Delta E)$	2D MC	0.1767 ± 0.0047	0.1349 ± 0.0034
$\mu_{\mathrm{tail}}^{2}(\Delta E)$ [GeV]	2D MC	-0.0080 ± 0.0010	-0.00389 ± 0.00085
$\sigma_{\mathrm{tail}}^{2}(\Delta E)$	2D MC	4.72 ± 0.13	4.42 ± 0.10
$f_2(\Delta E)$	2D MC	0.0660 ± 0.0030	0.0606 ± 0.0024
$\mu^{\mathrm{Voigt}}(M)$ [GeV/c^2]	2D MC	0.000065 ± 0.000073	0.000088 ± 0.000087
$\sigma_{\mathrm{main}}^{\mathrm{Voigt}}(M)$ [GeV/c^2]	2D MC	0.00666 ± 0.00034	0.00636 ± 0.00033
$\sigma_{\mathrm{tail}}^{\mathrm{Voigt}}(M)$	2D MC	5.48 ± 0.51	5.54 ± 0.77
$f(M)$	2D MC	0.102 ± 0.017	0.075 ± 0.017
$k_1^{\mathrm{corr}}(\Delta E \times M)$	2D MC	0.0387 ± 0.0021	0.0404 ± 0.0017
$k_2^{\mathrm{corr}}(\Delta E \times M)$	2D MC	0.728 ± 0.012	0.821 ± 0.011
$c_{\mathrm{mis}}^{1}(M)$	2D MC	0.290 ± 0.074	0.261 ± 0.067
$f_{\mathrm{mis}}(M)$	MC Truth	0.01067 ± 0.00052	0.00990 ± 0.00043
$c_1^{\Delta E}(\Delta E)$	F	-0.279 ± 0.055	-0.154 ± 0.021
$c_1^{M}(M)$	F	0.004 ± 0.059	0.063 ± 0.022
$\mu_{\mathrm{pb}}^{\Delta E}(\Delta E)$ [GeV]	2D M SB	-0.0029 ± 0.0021	-0.0000 ± 0.0020
$\sigma_{\mathrm{pb}}^{\Delta E}(\Delta E)$ [GeV]	2D M SB	0.0100 ± 0.0031	0.0116 ± 0.0026
$f_{\frac{\mathrm{pb}}{\mathrm{sig}}}$	F	0.193 ± 0.080	0.075 ± 0.036
N_{sig}	F	359.5 ± 31.1	1978 ± 75
N_{bkg}	F	1038 ± 41	7216 ± 104

Parameter [unit]	Source	Global value	
m_{η_c} [GeV/c^2]	PDG	2.9839 ± 0.0005	
Γ_{η_c} [GeV]	PDG	0.0320 ± 0.0007	

Table A.3: 2D model parameters and fit results for the signal mode $B^0 \rightarrow \eta_c(K_S^0 K^+ \pi^-) K_S^0$. Source of parameters denoted with $(\star)$ is rare-B MC for the SVD1 experiment.

Parameter (dim) [unit]	Source	SVD1 Value	SVD2 Value
$\mu_{\mathrm{main}}(\Delta E)$ [GeV]	2D MC	0.001485 ± 0.000080	0.001221 ± 0.000060
$\mu_{\mathrm{main}}^{\mathrm{CF}}(\Delta E)$ [GeV]	2D CS	-0.0002 ± 0.0011	-0.00061 ± 0.00040
$\sigma_{\mathrm{main}}(\Delta E)$ [GeV]	2D MC	0.009337 ± 0.000079	0.009063 ± 0.000059
$\sigma_{\mathrm{main}}^{\mathrm{CF}}(\Delta E)$	2D CS	1.26 ± 0.10	1.111 ± 0.039
$\mu_{\mathrm{tail}}^1(\Delta E)$ [GeV]	2D MC	-0.00206 ± 0.00043	0.00055 ± 0.00038
$\sigma_{\mathrm{tail}}^1(\Delta E)$	2D MC	1.684 ± 0.046	1.723 ± 0.040
$f_1(\Delta E)$	2D MC	0.1833 ± 0.0069	0.1479 ± 0.0047
$\mu_{\mathrm{tail}}^2(\Delta E)$ [GeV]	2D MC	-0.0029 ± 0.0013	0.00245 ± 0.00089
$\sigma_{\mathrm{tail}}^2(\Delta E)$	2D MC	4.40 ± 0.18	4.11 ± 0.11
$f_2(\Delta E)$	2D MC	0.0614 ± 0.0047	0.0685 ± 0.0038
$\mu^{\mathrm{Voigt}}(M)[\mathrm{GeV}/c^2]$	2D MC	0.00014 ± 0.00014	0.00009 ± 0.00011
$\sigma_{\mathrm{main}}^{\mathrm{Voigt}}(M)[\mathrm{GeV}/c^2]$	2D MC	0.00668 ± 0.00038	0.00669 ± 0.00037
$\sigma_{\mathrm{tail}}^{\mathrm{Voigt}}(M)$	2D MC	5.76 ± 0.58	5.684 ± 1.072
$f(M)$	2D MC	0.103 ± 0.017	0.061 ± 0.018
$k_1^{\mathrm{corr}}(\Delta E \times M)$	2D MC	0.0388 ± 0.0030	0.0445 ± 0.0022
$k_2^{\mathrm{corr}}(\Delta E \times M)$	2D MC	0.719 ± 0.017	0.793 ± 0.014
$c_{\mathrm{mis}}^1(M)$	2D MC	0.294 ± 0.059	0.164 ± 0.052
$f_{\mathrm{mis}}(M)$	MC Truth	0.0303 ± 0.0012	0.02574 ± 0.00090
$c_1^{\Delta E}(\Delta E)$	F	-0.33 ± 0.12	-0.279 ± 0.043
$c_1^M(M)$	F	-0.05 ± 0.12	0.096 ± 0.045
$\mu_{\mathrm{pb}}^{\Delta E}(\Delta E)$ [GeV]	2D M SB	0.0011 ± 0.0038	-0.0018 ± 0.0015
$\sigma_{\mathrm{pb}}^{\Delta E}(\Delta E)[\mathrm{GeV}]$	2D M SB $(\star)$	0.01092 ± 0.00027	0.0114 ± 0.0015
$f_{\frac{\mathrm{pb}}{\mathrm{sig}}}$	F	0.42 ± 0.28	0.244 ± 0.077
N_{sig}	F	64 ± 15	548 ± 39
N_{bkg}	F	251 ± 20	1707 ± 52

Parameter [unit]	Source	Global value
$m_{\eta_c}[\mathrm{GeV}/c^2]$	PDG	2.9839 ± 0.0005
$\Gamma_{\eta_c}[\mathrm{GeV}]$	PDG	0.0320 ± 0.0007

Table A.4: Final 4D model parameters and fit results for
the control mode $B^+ \rightarrow \eta_c(K_S^0 K^+ \pi^-)K^+$.

Parameter (dim) [unit]	Source	SVD1 Value	SVD2 Value
$\mu_{\mathrm{main}}(\Delta E)$ [GeV]	4D MC	0.001241 ± 0.000060	0.000926 ± 0.000049
$\mu_{\mathrm{main}}^{\mathrm{CF}}(\Delta E)$ [GeV]	F	-0.0004 ± 0.0010	-0.00049 ± 0.00037
$\sigma_{\mathrm{main}}(\Delta E)$ [GeV]	4D MC	0.009275 ± 0.000057	0.009306 ± 0.000045
$\sigma_{\mathrm{main}}^{\mathrm{CF}}(\Delta E)$	F	1.233 ± 0.092	1.106 ± 0.035
$\mu_{\mathrm{tail}}^1(\Delta E)$ [GeV]	4D MC	-0.00265 ± 0.00031	-0.00061 ± 0.00029
$\sigma_{\mathrm{tail}}^1(\Delta E)$	4D MC	1.538 ± 0.032	1.504 ± 0.030
$f_1(\Delta E)$	4D MC	0.1715 ± 0.0045	0.1320 ± 0.0033
$\mu_{\mathrm{tail}}^2(\Delta E)$ [GeV]	4D MC	-0.0079 ± 0.0010	-0.00373 ± 0.00083
$\sigma_{\mathrm{tail}}^2(\Delta E)$	4D MC	4.65 ± 0.13	4.381 ± 0.097
$f_2(\Delta E)$	4D MC	0.0675 ± 0.0030	0.0615 ± 0.0024
$\mu^{\mathrm{Voigt}}(M)$ [GeV/c^2]	4D MC	-0.00022 ± 0.00010	0.000083 ± 0.000087
$\sigma_{\mathrm{main}}^{\mathrm{Voigt}}(M)$ [GeV/c^2]	4D MC	0.00670 ± 0.00028	0.00644 ± 0.00032
$\sigma_{\mathrm{tail}}^{\mathrm{Voigt}}(M)$	4D MC	5.54 ± 0.51	5.92 ± 0.79
$f(M)$	4D MC	0.100 ± 0.014	0.070 ± 0.015
$k_1^{\mathrm{corr}}(\Delta E \times M)$	4D MC	0.0354 ± 0.0019	0.0368 ± 0.0016
$k_2^{\mathrm{corr}}(\Delta E \times M)$	4D MC	0.737 ± 0.012	0.828 ± 0.010
$c_{\mathrm{mis}}^1(M)$	4D MC	0.290 ± 0.074	0.261 ± 0.067
$f_{\mathrm{mis}}(M)$	MC Truth	0.01067 ± 0.00052	0.00990 ± 0.00043
$c_1^{\Delta E}(\Delta E)$	F	-0.274 ± 0.055	-0.154 ± 0.021
$c_1^M(M)$	F	0.012 ± 0.058	0.061 ± 0.022
$\mu_{\mathrm{pb}}^{\Delta E}(\Delta E)$ [GeV]	2D M SB	-0.0029 ± 0.0021	-0.0000 ± 0.0020
$\sigma_{\mathrm{pb}}^{\Delta E}(\Delta E)$ [GeV]	2D M SB	0.0100 ± 0.0031	0.0116 ± 0.0026
$\tau_{\mathrm{bkg}}(\Delta t)$ [ps]	4D M_{bc} SB	0.967 ± 0.078	0.767 ± 0.022
$\mu_\delta(\Delta t)$ [ps]	4D M_{bc} SB	-0.024 ± 0.022	0.0039 ± 0.0075
$f_\delta(\Delta t)$	4D M_{bc} SB	0.559 ± 0.045	0.485 ± 0.017
$\mu_{\mathrm{bkg\ sng}}(\Delta t)$ [ps]	4D M_{bc} SB	-0.094 ± 0.046	-0.020 ± 0.012
$f_{\mathrm{bkg\ sng}}^{\mathrm{tail}}(\Delta t)$	4D M_{bc} SB	0.032 ± 0.012	0.0629 ± 0.0058
$s_{\mathrm{bkg\ sng}}^{\mathrm{main}}(\Delta t)$	4D M_{bc} SB	1.170 ± 0.034	1.305 ± 0.015
$s_{\mathrm{bkg\ sng}}^{\mathrm{tail}}(\Delta t)$	4D M_{bc} SB	5.51 ± 0.74	4.72 ± 0.15
$\mu_{\mathrm{bkg\ mlt}}(\Delta t)$ [ps]	4D M_{bc} SB	0.08 ± 0.10	0.003 ± 0.034
$f_{\mathrm{bkg\ mlt}}^{\mathrm{tail}}(\Delta t)$	4D M_{bc} SB	0.146 ± 0.041	0.134 ± 0.013
$s_{\mathrm{bkg\ mlt}}^{\mathrm{main}}(\Delta t)$	4D M_{bc} SB	0.954 ± 0.050	1.029 ± 0.023
$s_{\mathrm{bkg\ mlt}}^{\mathrm{tail}}(\Delta t)$	4D M_{bc} SB	3.02 ± 0.35	4.18 ± 0.21
$f_{q\bar{q}}^{\Delta t}(\Delta t)$	F	0.818 ± 0.058	0.840 ± 0.018

Continued on next page

Parameter (dim) [unit]	Source	SVD1 Value	SVD2 Value
$\tilde{f}_{\text{sig}}^0$ (r-bin)	4D MC	0.4279 ± 0.0037	0.4068 ± 0.0033
$\tilde{f}_{\text{sig}}^1$ (r-bin)	4D MC	0.2874 ± 0.0029	0.3317 ± 0.0026
$\tilde{f}_{\text{sig}}^2$ (r-bin)	4D MC	0.1744 ± 0.0022	0.1877 ± 0.0019
$\tilde{f}_{\text{sig}}^3$ (r-bin)	4D MC	0.1685 ± 0.0020	0.1416 ± 0.0016
$\tilde{f}_{\text{sig}}^4$ (r-bin)	4D MC	0.1151 ± 0.0016	0.1147 ± 0.0014
$\tilde{f}_{\text{sig}}^5$ (r-bin)	4D MC	0.1503 ± 0.0016	0.1572 ± 0.0014
$\tilde{\eta}_0^{\text{CF}}$ (r-bin)	F	1.29 ± 0.11	1.199 ± 0.062
$\tilde{\eta}_1^{\text{CF}}$ (r-bin)	F	0.81 ± 0.12	0.918 ± 0.058
$\tilde{\eta}_2^{\text{CF}}$ (r-bin)	F	0.85 ± 0.16	1.119 ± 0.076
$\tilde{\eta}_3^{\text{CF}}$ (r-bin)	F	0.75 ± 0.14	1.044 ± 0.082
$\tilde{\eta}_4^{\text{CF}}$ (r-bin)	F	0.65 ± 0.16	1.041 ± 0.083
$\tilde{\eta}_5^{\text{CF}}$ (r-bin)	F	1.02 ± 0.13	1.037 ± 0.060
$\tilde{f}_{\text{bkg}}^0$ (r-bin)	F	0.411 ± 0.026	0.4052 ± 0.0090
$\tilde{f}_{\text{bkg}}^1$ (r-bin)	F	0.218 ± 0.019	0.2818 ± 0.0070
$\tilde{f}_{\text{bkg}}^2$ (r-bin)	F	0.181 ± 0.016	0.1390 ± 0.0050
$\tilde{f}_{\text{bkg}}^3$ (r-bin)	F	0.134 ± 0.013	0.1083 ± 0.0043
$\tilde{f}_{\text{bkg}}^4$ (r-bin)	F	0.080 ± 0.010	0.0725 ± 0.0034
$\tilde{f}_{\text{bkg}}^5$ (r-bin)	F	0.0438 ± 0.0078	0.0366 ± 0.0026
$f_{\frac{\text{pb}}{\text{sig}}}$	F	0.198 ± 0.076	0.054 ± 0.028
N_{sig}	F	358 ± 30	1990 ± 70
N_{bkg}	F	1039 ± 40	7205 ± 100

Parameter [unit]	Source	Global value
$m_{\eta_c}[\text{GeV}/c^2]$	PDG	2.9839 ± 0.0005
$\Gamma_{\eta_c}[\text{GeV}]$	PDG	0.0320 ± 0.0007
τ_{eff} [ps]	4D $B\bar{B}$ MC	1.152 ± 0.017
τ [ps]	PDG	1.641 ± 0.008
$\mathcal{S}_{CP}$	F	-0.027 ± 0.080
$\mathcal{A}_{CP}$	F	-0.052 ± 0.059

Table A.5: Final 4D model parameters and fit results for the signal mode $B^0 \to \eta_c(K_S^0 K^+ \pi^-)K_S^0$. Source of parameters denoted with $(\star)$ is rare-B MC for the SVD1 experiment.

Parameter (dim) [unit]	Source	SVD1 Value	SVD2 Value
$\mu_{\mathrm{main}}(\Delta E)$ [GeV]	4D MC	0.001473 ± 0.000079	0.001239 ± 0.000060
$\mu_{\mathrm{main}}^{\mathrm{CF}}(\Delta E)$ [GeV]	4D CS	-0.0004 ± 0.0010	-0.00049 ± 0.00037
$\sigma_{\mathrm{main}}(\Delta E)$ [GeV]	4D MC	0.009429 ± 0.000078	0.009125 ± 0.000059
$\sigma_{\mathrm{main}}^{\mathrm{CF}}(\Delta E)$	4D CS	1.233 ± 0.092	1.106 ± 0.035
$\mu_{\mathrm{tail}}^{1}(\Delta E)$ [GeV]	4D MC	-0.00203 ± 0.00043	0.00031 ± 0.00038
$\sigma_{\mathrm{tail}}^{1}(\Delta E)$	4D MC	1.630 ± 0.044	1.687 ± 0.039
$f_1(\Delta E)$	4D MC	0.1672 ± 0.0060	0.1381 ± 0.0042
$\mu_{\mathrm{tail}}^{2}(\Delta E)$ [GeV]	4D MC	-0.0029 ± 0.0012	0.00260 ± 0.00084
$\sigma_{\mathrm{tail}}^{2}(\Delta E)$	4D MC	4.20 ± 0.14	4.000 ± 0.096
$f_2(\Delta E)$	4D MC	0.0671 ± 0.0047	0.0724 ± 0.0037
$\mu^{\mathrm{Voigt}}(M)$[GeV/c^2]	4D MC	0.00015 ± 0.00014	0.00008 ± 0.00011
$\sigma_{\mathrm{main}}^{\mathrm{Voigt}}(M)$[GeV/$c^2$]	4D MC	0.00674 ± 0.00038	0.00674 ± 0.00028
$\sigma_{\mathrm{tail}}^{\mathrm{Voigt}}(M)$	4D MC	6.13 ± 0.60	6.32 ± 0.91
$f(M)$	4D MC	0.101 ± 0.015	0.059 ± 0.011
$k_1^{\mathrm{corr}}(\Delta E \times M)$	4D MC	0.0316 ± 0.0025	0.0364 ± 0.0019
$k_2^{\mathrm{corr}}(\Delta E \times M)$	4D MC	0.742 ± 0.017	0.812 ± 0.013
$c_{\mathrm{mis}}^{1}(M)$	2D MC	0.294 ± 0.059	0.164 ± 0.052
$f_{\mathrm{mis}}(M)$	MC Truth	0.0303 ± 0.0012	0.02574 ± 0.00090
$c_1^{\Delta E}(\Delta E)$	F	-0.32 ± 0.12	-0.276 ± 0.043
$c_1^{M}(M)$	F	-0.07 ± 0.12	0.096 ± 0.045
$\mu_{\mathrm{pb}}^{\Delta E}(\Delta E)$ [GeV]	2D M SB $(\star)$	0.0011 ± 0.0038	-0.0018 ± 0.0015
$\sigma_{\mathrm{pb}}^{\Delta E}(\Delta E)$[GeV]	2D M SB	0.01092 ± 0.00027	0.0114 ± 0.0015
$\tau_{\mathrm{bkg}}(\Delta t)$ [ps]	4D M_{bc} SB	1.11 ± 0.44	0.951 ± 0.068
$\mu_\delta(\Delta t)$ [ps]	4D M_{bc} SB	-0.011 ± 0.062	-0.017 ± 0.017
$f_\delta(\Delta t)$	4D M_{bc} SB	0.61 ± 0.16	0.548 ± 0.039
$\mu_{\mathrm{bkg\ sng}}(\Delta t)$ [ps]	4D M_{bc} SB	-0.02 ± 0.16	-0.052 ± 0.036
$f_{\mathrm{bkg\ sng}}^{\mathrm{tail}}(\Delta t)$	4D M_{bc} SB	0.050 ± 0.033	0.064 ± 0.016
$s_{\mathrm{bkg\ sng}}^{\mathrm{main}}(\Delta t)$	4D M_{bc} SB	1.176 ± 0.094	1.302 ± 0.030
$s_{\mathrm{bkg\ sng}}^{\mathrm{tail}}(\Delta t)$	4D M_{bc} SB	5.96 ± 1.15	4.33 ± 0.40
$\mu_{\mathrm{bkg\ mlt}}(\Delta t)$ [ps]	4D M_{bc} SB	-0.03 ± 0.26	0.055 ± 0.083
$f_{\mathrm{bkg\ mlt}}^{\mathrm{tail}}(\Delta t)$	4D M_{bc} SB	0.26 ± 0.32	0.136 ± 0.037
$s_{\mathrm{bkg\ mlt}}^{\mathrm{main}}(\Delta t)$	4D M_{bc} SB	1.03 ± 0.28	1.095 ± 0.051

Continued on next page

Parameter (dim) [unit]	Source	SVD1 Value	SVD2 Value
$s_{\mathrm{bkg\ mlt}}^{\mathrm{tail}}(\Delta t)$	4D M_{bc} SB	2.315 ± 1.068	3.55 ± 0.57
$f_{q\bar{q}}^{\Delta t}(\Delta t)$	4D M SB	0.83 ± 0.11	0.803 ± 0.030
$\tilde{f}_{\mathrm{sig}}^{0}$ (r-bin)	4D MC	0.4293 ± 0.0048	0.4115 ± 0.0040
$\tilde{f}_{\mathrm{sig}}^{1}$ (r-bin)	4D MC	0.2993 ± 0.0037	0.3423 ± 0.0032
$\tilde{f}_{\mathrm{sig}}^{2}$ (r-bin)	4D MC	0.1628 ± 0.0027	0.1904 ± 0.0023
$\tilde{f}_{\mathrm{sig}}^{3}$ (r-bin)	4D MC	0.1523 ± 0.0025	0.1285 ± 0.0018
$\tilde{f}_{\mathrm{sig}}^{4}$ (r-bin)	4D MC	0.1172 ± 0.0021	0.1116 ± 0.0017
$\tilde{f}_{\mathrm{sig}}^{5}$ (r-bin)	4D MC	0.1537 ± 0.0021	0.1650 ± 0.0018
$\tilde{\eta}_{0}^{\mathrm{CF}}$ (r-bin)	4D CS	1.29 ± 0.11	1.199 ± 0.062
$\tilde{\eta}_{1}^{\mathrm{CF}}$ (r-bin)	4D CS	0.81 ± 0.12	0.918 ± 0.058
$\tilde{\eta}_{2}^{\mathrm{CF}}$ (r-bin)	4D CS	0.85 ± 0.16	1.119 ± 0.076
$\tilde{\eta}_{3}^{\mathrm{CF}}$ (r-bin)	4D CS	0.75 ± 0.14	1.044 ± 0.082
$\tilde{\eta}_{4}^{\mathrm{CF}}$ (r-bin)	4D CS	0.65 ± 0.16	1.041 ± 0.083
$\tilde{\eta}_{5}^{\mathrm{CF}}$ (r-bin)	4D CS	1.02 ± 0.13	1.037 ± 0.060
$\tilde{f}_{\mathrm{bkg}}^{0}$ (r-bin)	4D M_{bc} SB	0.410 ± 0.017	0.3963 ± 0.0061
$\tilde{f}_{\mathrm{bkg}}^{1}$ (r-bin)	4D M_{bc} SB	0.238 ± 0.013	0.2897 ± 0.0048
$\tilde{f}_{\mathrm{bkg}}^{2}$ (r-bin)	4D M_{bc} SB	0.1226 ± 0.0094	0.1385 ± 0.0034
$\tilde{f}_{\mathrm{bkg}}^{3}$ (r-bin)	4D M_{bc} SB	0.1132 ± 0.0086	0.1120 ± 0.0029
$\tilde{f}_{\mathrm{bkg}}^{4}$ (r-bin)	4D M_{bc} SB	0.0866 ± 0.0073	0.0738 ± 0.0023
$\tilde{f}_{\mathrm{bkg}}^{5}$ (r-bin)	4D M_{bc} SB	0.0318 ± 0.0045	0.0293 ± 0.0015
$f_{\frac{\mathrm{pb}}{\mathrm{sig}}}$	F	0.20 ± 0.20	0.214 ± 0.069
N_{sig}	F	68 ± 14	548 ± 37
N_{bkg}	F	247 ± 19	1707 ± 51

Parameter [unit]	Source	Global value
m_{η_c}[GeV/c^2]	PDG	2.9839 ± 0.0005
Γ_{η_c}[GeV]	PDG	0.0320 ± 0.0007
τ_{eff} [ps]	4D $B\bar{B}$ MC	1.111 ± 0.033
τ [ps]	PDG	1.519 ± 0.007
$\mathcal{S}_{CP}$	F	0.588 ± 0.165
$\mathcal{A}_{CP}$	F	0.161 ± 0.116

APPENDIX B

SVD1 Experiment Projections

In this appendix, 4D fit model and data projections of the signal MC sample (Fig. B.1) and M_{bc} sideband data (Fig. B.2) for the SVD1 experiment are shown. The corresponding projections for the SVD2 experiment are included in the main text.
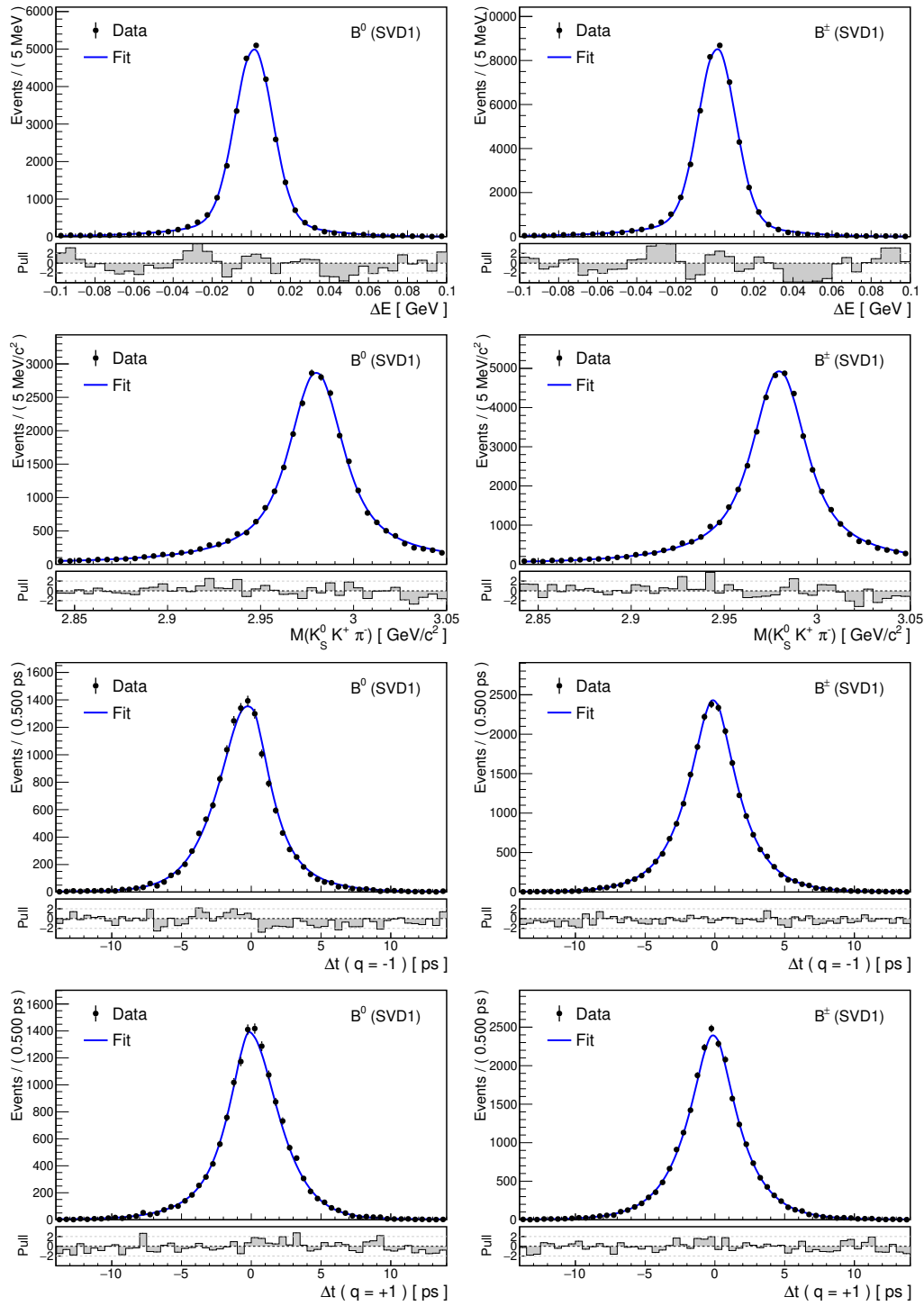
Figure B.1: Fit model and data projections for the signal MC sample of the SVD1 experiment for the signal (left) and control mode (right).
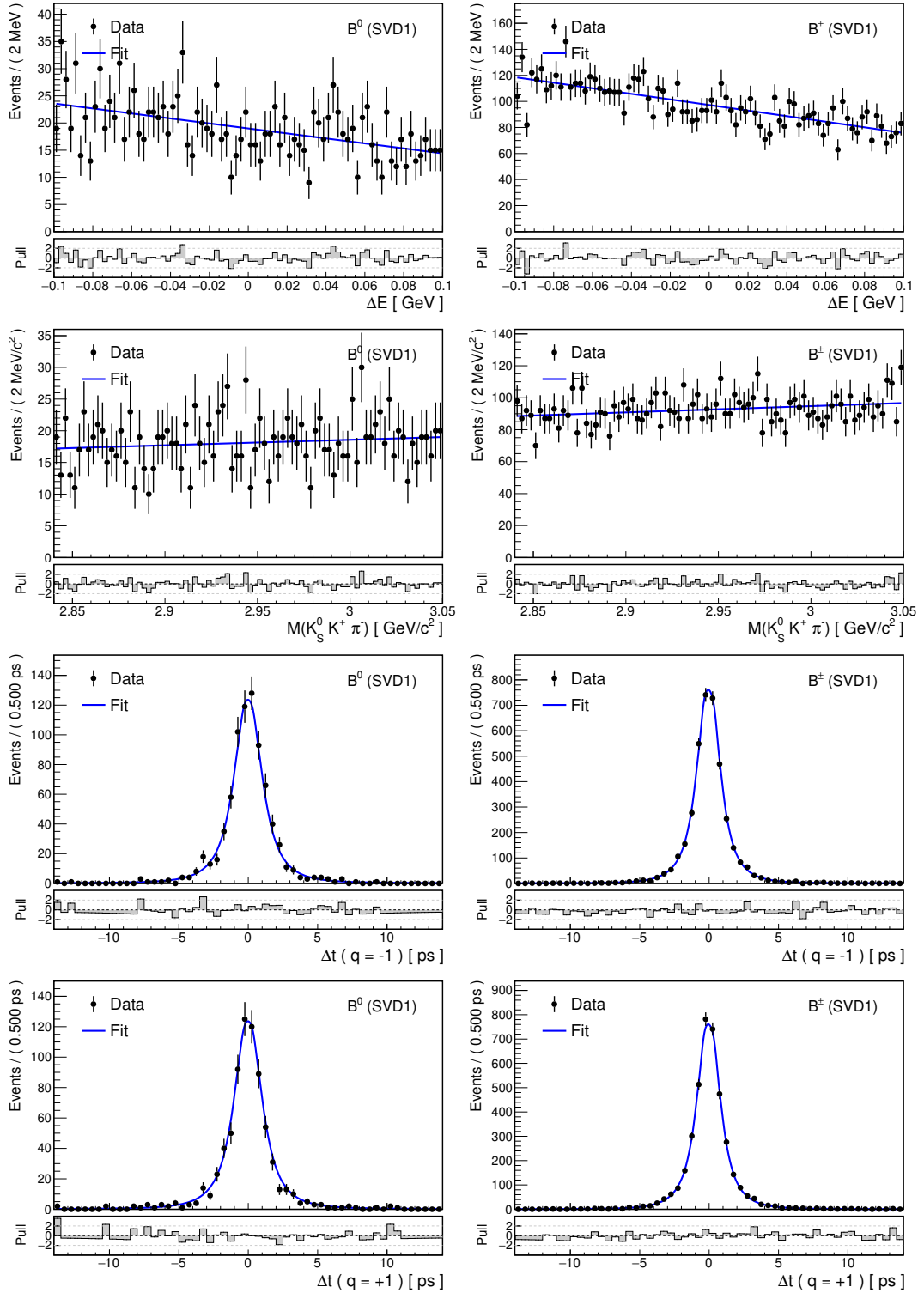
Figure B.2: Fit model and data projections for the M_{bc} sideband data of the SVD1 experiment for the signal (left) and control mode (right).

APPENDIX C

Final Result Projections

In this appendix, we show data and fit projection in the full signal window, while single-enhanced region is used for projections in the main text. The control mode is shown in Fig. C.1 and the signal mode in Fig. C.2. Note that only data with $|\Delta t| < 7$ ps are projected (such that the same events are present in all shown histograms).
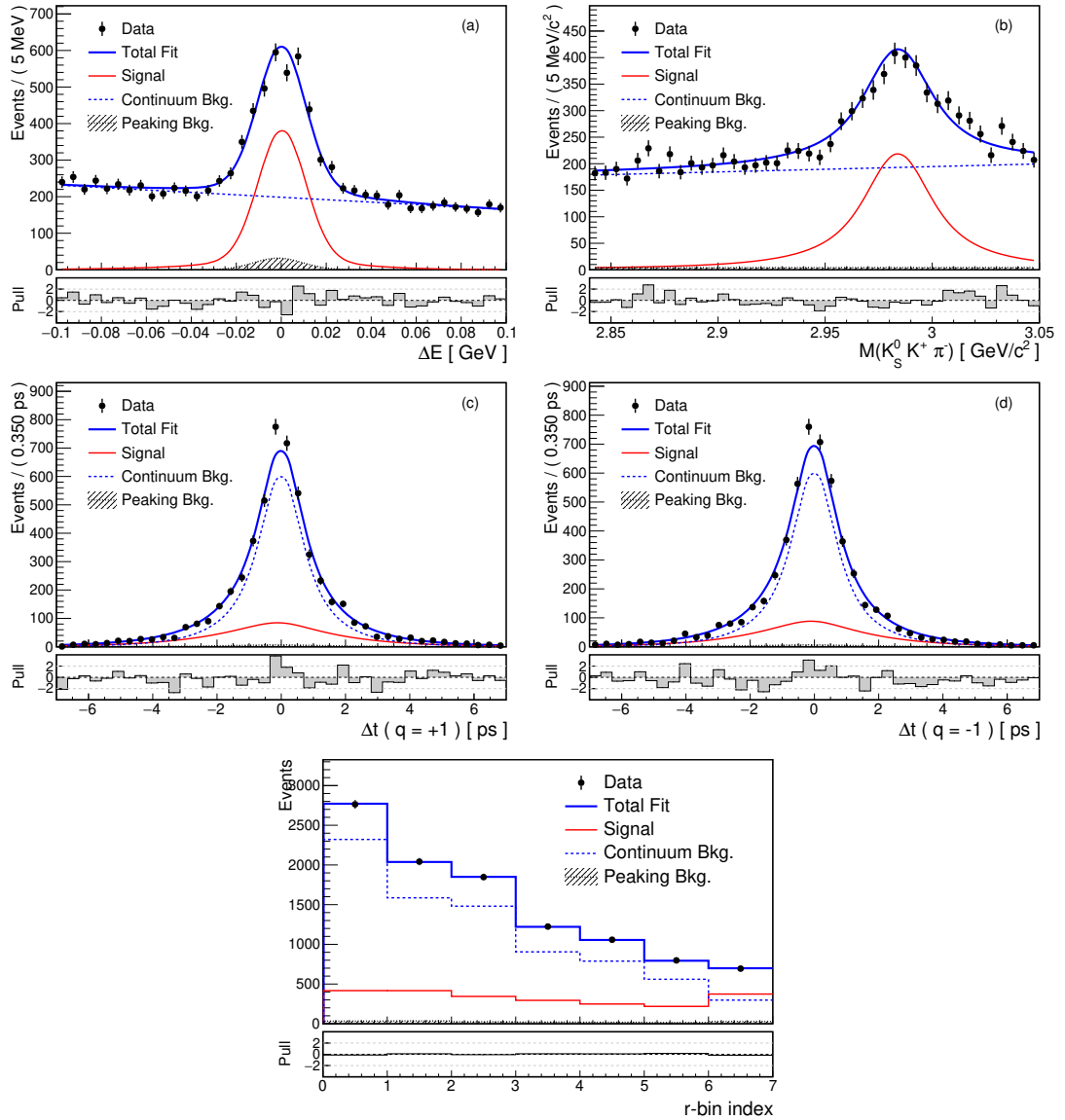

Figure C.1: Control mode data and fit projections in the full signal window.
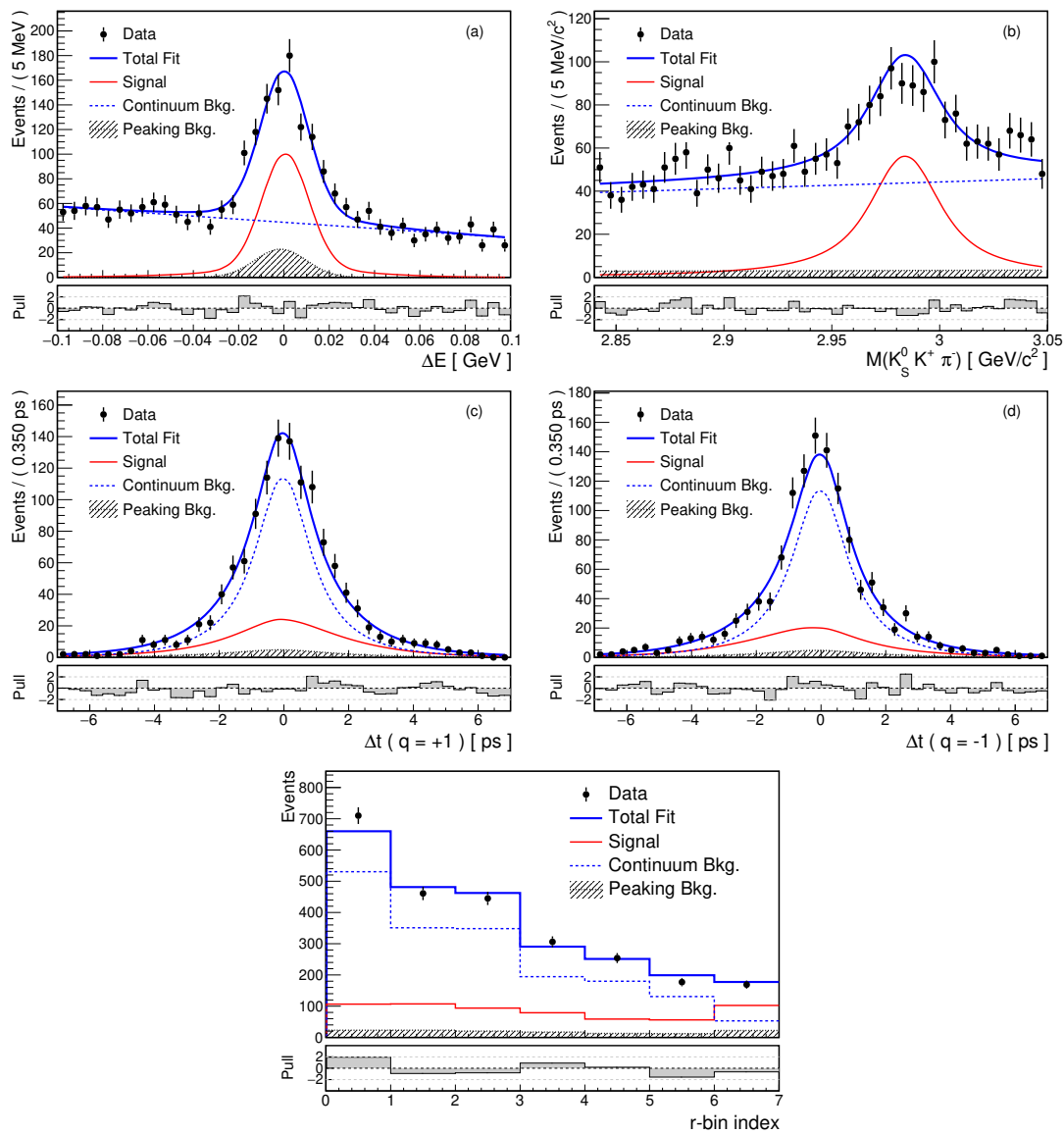
Figure C.2: Signal mode data and fit projections in the full signal window.

APPENDIX D

Rigid Body Parameter Transformations

In Sec. 12.2.3, we discuss alignment in different local hierarchy frames defined by their coordinate systems. The 6×6 transformation matrices between rigid body parameters of two coordinate frames can be decomposed into four 3×3 blocks

$$\frac{\partial \boldsymbol{a}_A}{\partial \boldsymbol{a}_B} = \begin{pmatrix} \boldsymbol{M_1} & \boldsymbol{M_2} \\ 0 & \boldsymbol{M_3} \end{pmatrix}, \tag{D.1}$$

where the lower left block corresponds to $\frac{\partial(\Delta\alpha,\Delta\beta,\Delta\gamma)_A^{\mathsf{T}}}{\partial(\Delta u,\Delta v,\Delta w)_B^{\mathsf{T}}}$ and is zero as angles do not change under translations. Block $\boldsymbol{M_1}$ corresponds to $\frac{\partial(\Delta u,\Delta v,\Delta w)_A^{\mathsf{T}}}{\partial(\Delta u,\Delta v,\Delta w)_B^{\mathsf{T}}}$. If the origin of the local frame B in the A-frame is[1] $\boldsymbol{r_A}$, one can write

$$\boldsymbol{r_A} = \boldsymbol{R}\boldsymbol{R}^{\mathsf{T}}\boldsymbol{r_A} = \boldsymbol{R}\,\boldsymbol{r_B}, \tag{D.2}$$

where $\boldsymbol{r_B}$ is the same vector as $\boldsymbol{r_A}$, but now expressed in coordinates of the lower frame B. A small change of the daughter's frame origin in its local coordinates $(\Delta u, \Delta v, \Delta w)_B^{\mathsf{T}}$ will then translate to a change in mother frame $(\Delta u, \Delta v, \Delta w)_A^{\mathsf{T}}$ as

$$(\Delta u, \Delta v, \Delta w)_A^{\mathsf{T}} = \boldsymbol{R}\,(\Delta u, \Delta v, \Delta w)_B^{\mathsf{T}}. \tag{D.3}$$

Therefore block $\boldsymbol{M_1} = \boldsymbol{R}$.

Lemma: For an infinitesimal rotation matrix $\boldsymbol{\Delta R}$ of the form

$$\boldsymbol{\Delta R} = \begin{pmatrix} 1 & -\Delta\gamma & \Delta\beta \\ \Delta\gamma & 1 & -\Delta\alpha \\ -\Delta\beta & \Delta\alpha & 1 \end{pmatrix}, \tag{D.4}$$

the vector

$$\boldsymbol{\alpha} = (\Delta\alpha, \Delta\beta, \Delta\gamma)^{\mathsf{T}} \tag{D.5}$$

is (up to a multiplication factor) the axis of the rotation $\boldsymbol{\Delta R}$.
Proof:

$$\boldsymbol{\Delta R}\,\boldsymbol{\alpha} = \boldsymbol{\alpha} \tag{D.6}$$

and thus $\boldsymbol{\alpha}$ is unchanged by the rotation $\boldsymbol{R}$ ($\boldsymbol{\alpha}$ is an eigenvector of $\boldsymbol{\Delta R}$ with eigenvalue of one). Therefore $\boldsymbol{\alpha}$ must be the axis of the rotation.

---

[1] We have previously denoted $\boldsymbol{r_A}$ as $\boldsymbol{r_0}$, for the case of A being the global frame. The convention for the rotation matrix is the same as defined before, so $\boldsymbol{R}$ transforms vectors from lower to upper frame.

Block $M_3 = \frac{\partial(\Delta\alpha, \Delta\beta, \Delta\gamma)_A^\intercal}{\partial(\Delta\alpha, \Delta\beta, \Delta\gamma)_B^\intercal}$. In the daughter frame, an infinitesimal rotation ΔR will have axis $\alpha_B = (\Delta\alpha, \Delta\beta, \Delta\gamma)_B^\intercal$. But any direction vector, in particular the rotation axis, transforms from daughter to mother frame using R as $\alpha_A = R\,\alpha_B$ and therefore

$$(\Delta\alpha, \Delta\beta, \Delta\gamma)_A^\intercal = R\,(\Delta\alpha, \Delta\beta, \Delta\gamma)_B^\intercal \tag{D.7}$$

and block $M_3 = R$.

In block M_2, we are interested in the partial derivatives $\frac{\partial(\Delta u, \Delta v, \Delta w)_A^\intercal}{\partial(\Delta\alpha, \Delta\beta, \Delta\gamma)_B^\intercal}$. By expanding $(\Delta u, \Delta v, \Delta w)_A^\intercal$ to first order in $(\Delta\alpha, \Delta\beta, \Delta\gamma)_A^\intercal$, one gets

$$(\Delta u, \Delta v, \Delta w)_A^\intercal = \frac{\partial(\Delta u, \Delta v, \Delta w)_A^\intercal}{\partial(\Delta\alpha, \Delta\beta, \Delta\gamma)_A^\intercal}\, R\,(\Delta\alpha, \Delta\beta, \Delta\gamma)_B^\intercal, \tag{D.8}$$

where we have already used Eq. D.7. In the mother frame, the daughter frame origin is $r_A = (u, v, w)_A^\intercal$ and rotations in mother frame propagate to small shifts of r_A as

$$\begin{pmatrix} \Delta u \\ \Delta v \\ \Delta w \end{pmatrix}_A = \begin{pmatrix} 1 & -\Delta\gamma & \Delta\beta \\ \Delta\gamma & 1 & -\Delta\alpha \\ -\Delta\beta & \Delta\alpha & 1 \end{pmatrix}_A \begin{pmatrix} u \\ v \\ w \end{pmatrix}_A. \tag{D.9}$$

The desired matrix of partial derivatives evaluates to

$$D = \frac{\partial(\Delta u, \Delta v, \Delta w)_A^\intercal}{\partial(\Delta\alpha, \Delta\beta, \Delta\gamma)_A^\intercal} = \begin{pmatrix} 0 & w & -v \\ -w & 0 & u \\ y & -u & 0 \end{pmatrix}_A \tag{D.10}$$

and block $M_2 = DR$. In total, the transformation from daughter to mother frame reads

$$\frac{\partial a_A}{\partial a_B} = \begin{pmatrix} R & DR \\ 0 & R \end{pmatrix} \tag{D.11}$$

and for the inverted relation, one gets

$$\frac{\partial a_B}{\partial a_A} = \begin{pmatrix} R^\intercal & -R^\intercal D \\ 0 & R^\intercal \end{pmatrix}, \tag{D.12}$$

as can be checked by direct multiplication: $\frac{\partial a_B}{\partial a_A}\frac{\partial a_A}{\partial a_B} = \frac{\partial a_A}{\partial a_B}\frac{\partial a_B}{\partial a_A} = \mathbf{1}_{6\times 6}$.