UNIVERSITÀ DI PISA

Dipartimento di Fisica
Corso di Laurea Magistrale in Fisica

TESI DI LAUREA MAGISTRALE IN FISICA

Measurement of the branching fraction of the baryonic decay mode $B^0(\bar{B}^0)\rightarrow pp\bar{p}\bar{p}$ at *BABAR* and future prospects at Belle II

CANDIDATA
Laura Zani

RELATORE
Prof. Francesco Forti

ANNO ACCADEMICO 2015/2016

Contents

Introduction

The goal of this thesis is the measurement of the branching fraction of the baryonic channel $B^0(\bar{B}^0)\to pp\bar{p}\bar{p}$ using the full data set collected by the High Energy Physics experiment *BABAR*, situated at PEP-II e^+e^- collider, at SLAC National Laboratory (California), and its extrapolation at Belle II, operating at SuperKEKB at KEK, in Tsukuba (Japan). This thesis also includes the experimental activity performed at the High Technology Laboratories of INFN, Sezione di Pisa, with the Pisa Silicon Vertex Detector (SVD) group of the Belle II collaboration.

In the field of the *beauty* physics, baryonic B meson decays are an important element for better understanding the process of hadronization into baryons. Up to now, the discrepancy between the inclusive branching fraction of all the B meson decay modes with at least a couple of baryons in the final state, measured by ARGUS [1] in 1992 to be $6.8 \pm 0.6\%$, and the sum of exclusive baryonic channels, averaged on neutral and positive B mesons at less than 1%, represents an open issue.

The analysis presented in this thesis studies the decays of 471 milions of B meson pairs collected at the $\Upsilon(4S)$ peak during the decade of operation of *BABAR*. It aims at reconstructing the neutral B meson decays to a pair of protons and a pair of antiprotons. The best candidate reconstruction and the event selection have been optimized by studying the background and signal distributions from official *BABAR* Monte Carlo samples. Different selection techniques have been tested, based on cuts on the event variable distributions and on multivariate analysis methods. To estimate the branching fraction, I start from the upper limit on $\mathrm{BF}(\bar{B}^0 \to \Lambda_c^+ p\bar{p}\bar{p})$ measured at *BABAR* to be 2.8×10^{-6} at a 90% confidence level. Scaling the result with the CKM factor $\frac{|V_{ub}|^2}{|V_{cb}|^2}$ due to the $b \to u$ transition, and including the phase space enlargement factor because of the lower proton mass with respect to the Λ_c mass, the branching fraction estimate for the decay mode $B^0(\bar{B}^0)\to pp\bar{p}\bar{p}$ is of 10^{-7}.

This channel has a high reconstruction efficiency due to its distinct signature of four charged tracks coming from the same vertex. Despite the limited contribution coming from this mode to the inclusive baryonic branching fraction, it is helpuful for clarifying the strong mechanism of baryon production and for providing experimental support to theoretical models. In the event reconstruction, two kinematic variables are used: the beam energy substituted mass m_{ES} and the energy difference ΔE. The former is the B meson invariant mass where the reconstructed B meson energy is replaced by the beam energy and the latter corresponds to the difference between the B meson energy and the beam energy. These variables quantify the kinematic constraints

---

[1] ARGUS was an experiment ran at the electron-positron collider ring DORIS II at DESY, Hamburg, Germany.

imposed by the experimental setup and they are used for identifying B meson candidates coming from $\Upsilon(4S)$ decays. For a signal event, the m_{ES} distribution is expected to peak at the B meson invariant mass (5.279 GeV/c^2), while ΔE peaks at zero.

The event selection has been finalized in two phases: a cut-based preselection, using kinematic cuts, the particle identification reponse for the tracks detected in the event, and the probability of the common vertex fit, removes most part of combinatorial backgrounds. The suppression of the remaining background is optimized by a Boosted Decision Tree (BDT) selector. It exploits as input variables ΔE, the angular distribution of the B meson candidate in the $\Upsilon(4S)$ rest frame and event shape variables, which are powerful in the rejection of the backgrounds coming from continuum events $e^+e^- \to q\bar{q}$ with q=u,d,s and $e^+e^- \to c\bar{c}$, characterized by a *jet-like* behaviour.

I compared Monte Carlo (MC) distributions with those obtained from the side band region of on-peak data (m_{ES}< 5.27) and with the off-peak data collected at 40 MeV below the $\Upsilon(4S)$ energy threshold and therefore not containing any signal events. The extraction of the signal yield consists of an unbinned extended maximum likelihood fit to the m_{ES} distribution of on-peak data, in the whole reconstruction range $5.2 < m_{ES} < 5.3$ GeV/c^2, after the event selection application. I implemented Monte Carlo studies for modeling the background and signal MC distributions of m_{ES}. I set up toy MC studies in order to evaluate the bias on the signal yield introduced by the fitting procedure.

From MC studies, the expected significance on the extracted signal yield (10 events with the assumed branching fraction of 10^{-7}) is around 2σ and the branching fraction upper limit at a 90% confidence level will be estimated.

In this thesis, no result from the fit to $\Upsilon(4S)$ on-peak data is going to be presented, since the *unblinding* has not been approved yet. The signal window will remain blind until the end of the analysis, in order to avoid any experimenter's bias.

The measurement is mainly statistically limited at the *BABAR* data integrated luminosity. In the last part of this thesis, the prospects for this measurement at Belle II will be extrapolated, comparing the *BABAR* performance with the new detector which is being installed at SuperKEKB factory. A factor 10 improvement in the statistical precision is expected at the Belle II design integrated luminosity $\int \mathcal{L}dt = 50$ ab^{-1}.

My involvement in Belle II experiment concerns also the detector construction activity performed as a member of the SVD Pisa group, regarding the assembling and the electrical testing of the *forward* and *backward* Double-sided Silicon Strip Detector modules used in the three outer layers of the SVD. The SVD is an essential part of the tracking system at Belle II and its efficiency and resolution determines the quality of the charged tracks reconstruction. I have personally contributed to the assembly and testing of the forward and backward modules performed in the Pisa laboratory.

Chapter 1

Physics Motivation

The Standard Model (SM) of particle physics has been developed starting from 1970s as result of both experimental discoveries and theories of the previous decades and indicates that matter is composed of structureless units (quarks and leptons) known as fundamental particles, whose interactions are governed by four fundamental forces (electromagnetism, weak and strong interactions, gravity). As a unified description of all those theories, it provides the best understanding of these particles and how three of the four fundamental forces (gravity excluded) act on them. Up to now, it has explained almost all experimental results and predicted a large variety of phenomena.

1.1 Brief summary of the Standard Model

Matter is composed of two basic blocks, both consisting in $\frac{1}{2}$-spin particles (in units of $\hbar$) and hence fermions obeying to the Fermi-Dirac statistics: leptons and quarks. Leptons denote particles that, if charged, interact both electromagnetically and weakly, if neutral only weakly. Quarks are the fermionic constituents of hadrons — and then nuclei — and they interact strongly, weakly and electromagnetically.

Both leptons and quarks consist of six types of particles gathered into doublets, which give rise to the three mass generations. The same structure is repeated for anti-particles. All stable matter in the universe belongs to the first and lightest generation, while the heavier doublets correspond to unstable particles which necessarily decay to the lighter generations. All doublets have two components differing for one unit charge.

As regards the six types of quarks, known as *flavours*, they also present another property corresponding to the strong charge, the *colour*. Three colours exist in the theory of strong interaction, as described in the *quantum chromodynamics* (QCD). When the quarks combine to form a meson (bound state of quark-antiquark) or a baryon (bound state of three quarks), they must produce a colourless object, due to the colour-confinement principle.

In addition, the SM provides, for each force, carrier bosons responsible for the interactions between matter particles. The interactions are interpreted in terms of the exchange of the corrisponding boson, which for the three fundamental forces unified in the SM are:

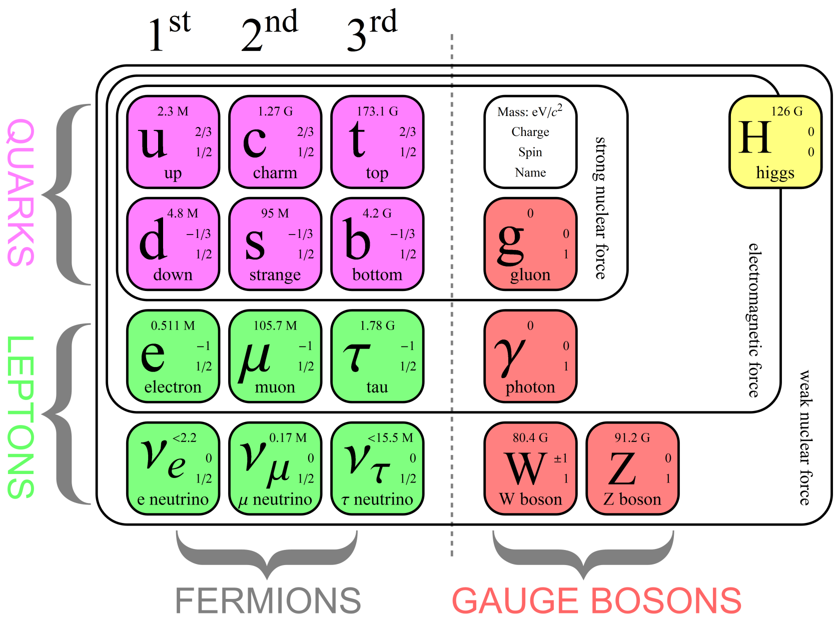


Figure 1.1: The Standard Model particles are shown. For each particle, mass, charge, spin and name are given. All the fermions are gathered in the first, second and third column, representing the three mass generations. Quarks are grouped in the first two rows. Their masses varies from a few MeV/c^2 (u, d) to a hundred GeV/c^2 (t). In the third and fourth rows, the leptons doublets are shown, with masses going from fractions of eV/c^2 to almost 2 GeV/c^2 (τ). The fourth column contains the gauge bosons, the force carrying particles of fundamental interactions. The Higgs boson is represented on the top right corner of this chart.

- the gluons for strong interaction;
- the $W^{+/-}, Z^0$ for weak interaction;
- the photons for the electromagnetic interaction.

Finally, the graviton, a massless boson of spin $s = 2$ with no electrical charge, is predicted to be the force-carrying particle for gravity, although it has not yet been found experimentally.

Two further ingredients are needed to describe the nature of the interactions and how particles acquire mass: local gauge invariance and the spontaneously symmetry breaking, whose combined application requires the gauge vector bosons (W^+, W^-, Z^0) to acquire mass. This phenomenon is known as the *Higgs mechanism* and the Higgs boson is the scalar field that causes spontaneous breaking of the gauge symmetry[1].

The mathematical description of the SM is given by the gauge symmetry $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$, where the three terms correspond respectively to the *colour* symmetry for the QCD, the weak *isospin* symmetry and the *hypercharge* symmetry for electromagnetism.

Figure 1.1 shows all the particles, fermions and bosons, mentioned above and reports also their main properties (mass, charge, spin).

1.2 The Cabibbo-Kobayashi-Maskawa matrix and its parameterization

The flavour transition between quarks are mediated by the weak interaction charged currents. Transitions are allowed between up and down-type quarks, not only within the same doublets, but also between different generations, even though the latter are suppressed. This suppression was first explained by Cabibbo [2] with the principle of the quark mixing matrix describing the rotation of the mass quark eigenstates (d, s) to the weak interaction eigenstates (d', s'):

$$|d'\rangle = \cos\theta_C|d\rangle + \sin\theta_C|s\rangle \ , \tag{1.1}$$

$$|s'\rangle = -\sin\theta_C|d\rangle + \cos\theta_C|s\rangle \ . \tag{1.2}$$

Experimental observations allowed Cabibbo to estimate $\sin\theta_C \simeq 0.23$, which explains the mixing in the scenario of 2 quark families. The extension to the three quark generations is provided by the Kobayashi-Maskawa mechanism [3] which explains the source of CP violation in the SM. The mixing of the six known quarks is therefore described by a 3×3 unitary matrix:

$$\hat{V}_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{1.3}$$

The indices of each element V_{q1q2} refer to the transition $q_1 \to q_2$. The Cabibbo-Kobayashi-Maskawa matrix (CKM) is described by four free parameters: 3 mixing angles which are real parameters and one irreducibile complex phase from which the CP violation arises. In the standard parameterization the CKM matrix is:

$$V_{CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} . \tag{1.4}$$

It is written as the composition of three rotations and its elements are given in terms of sine (s_{ij}) and cosine (c_{ij}) of the three mixing angles (θ_{ij}) and the complex phase δ. The Cabibbo 2×2 matrix is embodied in the CKM extension and $|V_{us}| = s_{12} = sin\theta_C$. Experimental observations establish $|V_{ub}| = s_{13} \sim 10^{-3}$, from which derives that c_{13} is close to one. The four independent parameters are δ and $|V_{ub}| = s_{13}$,$|V_{cb}| \approx s_{23}$,$|V_{us}| \approx s_{12}$. The hierarchical trend of these angles $s_{13} \ll s_{23} \ll s_{12} \ll 1$ is highlighted by the Wolfenstein parameterization, which is an expansion in terms of the small parameter $\lambda = 0.2272 \pm 0.0010 = s_{12}$. The CKM mixing matrix, in the basis of the independent parameters (A, λ, ρ, η), becomes:

$$V_{CKM} = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4) \ . \tag{1.5}$$

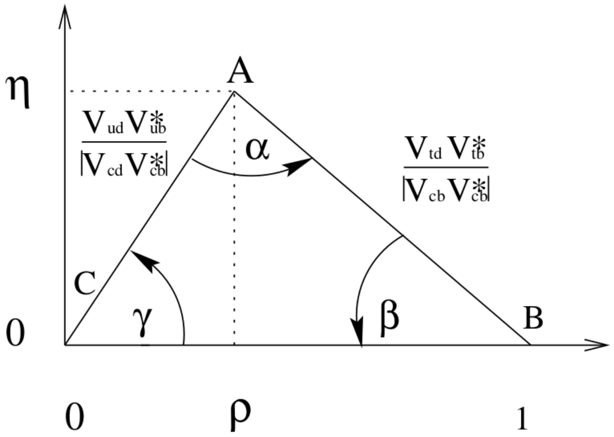


Figure 1.2: The normalized unitarity triangle in the (ρ, η) plane. The vertex A has coordinates $(\bar{\rho}, \bar{\eta})$.

In this parameterization the CP violation arises from the imaginary part η and the connection between the two formalisms is given by the transformation rules:

$$s_{12} = \lambda = \frac{|V_{us}|}{\sqrt{|V_{ud}|^2 + |V_{us}|^2}} \ , \quad s_{23} = A\lambda^2 = \lambda \frac{|V_{cb}|}{|V_{us}|} \ , \quad s_{13}e^{i\delta} = A\lambda^3(\rho + i\eta) = V^*_{ub} \ , \tag{1.6}$$

and finally:

$$\rho = \frac{s_{13}}{s_{23}s_{12}} \cos\delta \qquad\qquad \eta = \frac{s_{13}}{s_{23}s_{12}} \sin\delta$$

The Wolfenstein parameterization underlines that the diagonal elements are close to one, while mixing strenght is reduced for off-diagonal elements. These couplings are the explanation for the favoured transitions within the same quarks doublet. The expansion until $\mathcal{O}(\lambda^4)$ for the matrix element is meaningful, since the first corrections only occurs at $\mathcal{O}(\lambda^7) - \mathcal{O}(\lambda^8)$ for V_{us} ,V_{cb}.

1.2.1 The unitarity triangle

The unitarity of CKM matrix is translated into relations between the rows and columns of the matrix itself. In particular, for $j \neq k$ the relations $\sum_i V_{ij}V^*_{ik} = 0$ can be regarded as triangles in the complex plane (ρ, η). They all have the same area which represents a measure of the occuring CP violation.

The most studied triangle is the one given by:

$$V_{ud}V^*_{ub} + V_{cd}V^*_{cb} + V_{td}V^*_{tb} = 0 \ . \tag{1.7}$$

Dividing the above equation by $V_{cd}V^*_{cb}$ one gets:

$$\frac{V_{ud}V^*_{ub}}{V_{cd}V^*_{cb}} + 1 + \frac{V_{td}V^*_{tb}}{V_{cd}V^*_{cb}} \tag{1.8}$$

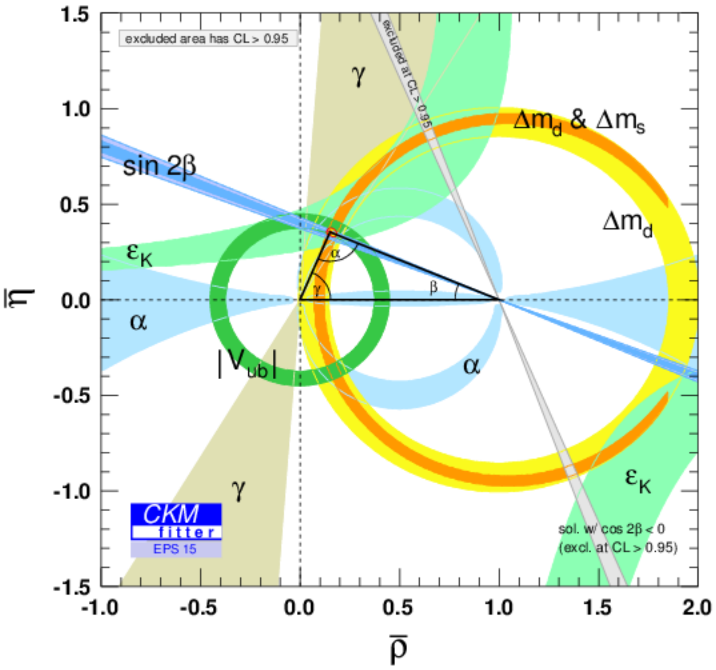


Figure 1.3: Constraints on the CKM parameterization in the $\bar{\rho}, \bar{\eta}$ plane; the results of most recent extrapolations are shown.

known as the normalized *unitarity triangle*. Its representation is given in Figure 1.2.

The normalized side of the triangle has vertices (0, 0) and (0, 1) in the (ρ, η) plane. The remaining vertex has coordinates $(\bar{\rho}, \bar{\eta})$, with $\bar{\rho} = \rho(1 - \lambda^2/2)$ and $\bar{\eta} = \eta(1 - \lambda^2/2)$. The three internal angles can also be defined as function of the CKM matrix elements:

$$\alpha = \arg\left[\frac{V_{ub}^* V_{ud}}{V_{tb}^* V_{td}}\right] \quad , \quad \beta = \arg\left[\frac{V_{tb}^* V_{td}}{V_{cb}^* V_{cd}}\right] \quad , \quad \gamma = \arg\left[\frac{V_{cb}^* V_{cd}}{V_{ub}^* V_{ud}}\right]. \tag{1.9}$$

The CP-violating condition corresponds to $\eta \neq 0$ or equivalently to a non-vanishing area of the unitarity triangle. An overview of the most recent experimental limits for $\bar{\rho}, \bar{\eta}$ is given in Figure 1.3.

1.3 Time-dependent CP asymmetries

In nature, pairs of neutral mesons with exactly the same quantum numbers except for the *flavour* are subject to oscillations of one particle in each other, due to the fact that mass eigenstates do not correspond to the *flavour* ones. The time evolution of quantum-coeherent pairs of neutral mesons, one the anti-particle of the other, induce the *mixing* phenomenon. Experimental evidence of mixing have been found in three types of neutral meson systems: K^0, D^0, B^0.

A deeper understanding of the mixing of the B mesons is fundamental for the study of the time-dependent CP asymmetries, whose measurement is one of the goal of the B Factories (see Chapter 2). For a full treatment of the topic, the reader may find interesting information elsewhere [4]. A brief introduction to the physics underlying the method of the time-dependent CP asymmetries measurement is given.

1.3.1 CP violation in B mesons phenomenology

The CP violation due to the CKM irreducible complex phase as established by the KM mechanism becomes observable through interference effects. The simplest case is given by an amplitude A_f consisting of two contributions:

$$A_f = \lambda_1 \langle f|O_1|B\rangle + \lambda_2 \langle f|O_2|B\rangle \tag{1.10}$$

where $\lambda_{1,2}$ are the complex coupling constants for the two processes which causes the intial B meson state $|B\rangle$ to decay to the final state $|f\rangle$. The interaction operators between the intial and final states are O_1, O_2. If one considers the CP conjugate process $\bar{B} \to \bar{f}$, the amplitude can be written as

$$\bar{A}_{\bar{f}} = \lambda_1^* \langle \bar{f}|O_1^\dagger|\bar{B}\rangle + \lambda_2^* \langle \bar{f}|O_2^\dagger|\bar{B}\rangle \tag{1.11}$$

Assuming CP-invariance for the strong interactions involved in the matrix element, one can write:

$$\langle \bar{f}|O_1^\dagger|\bar{B}\rangle = \langle f|O_1|B\rangle \qquad \langle \bar{f}|O_2^\dagger|\bar{B}\rangle = \langle f|O_2|B\rangle \tag{1.12}$$

end it can be shown that the CP asymmetry $\mathcal{A}_{CP}$ is proportional to both the *weak phase* from the CKM matrix included in the coupling constants product and to the *strong phase* appearing in the matrix element product, which derives form a difference in phase between the two contributions involved in the amplitude:

$$\mathcal{A}_{C\mathcal{P}}(B \to f) \equiv \frac{\Gamma(B \to f) - \Gamma(\bar{B} \to \bar{f})}{\Gamma(B \to f) + \Gamma(\bar{B} \to \bar{f})} \propto 2Im[\lambda_1 \lambda_2^*] Im[\langle f|O_1|B\rangle \langle f|O_2|B\rangle^*] \tag{1.13}$$

Considering decays into CP eigenstates for which is true that $f = \bar{f}$, if one assumes to have as intial state a quantum-coeherent pair of neutral B mesons, as the colour-singlet $B^0\bar{B}^0$ produced at the $\Upsilon(4S)$ resonance in a B Factory, their time evolution generates a phase difference $(\Delta m \Delta t)$ which acts in the same way as the above mentioned strong phase difference, allowing to observe the CP violation. In this case the time-dependent CP asymmetry results from the two contributions $B \to f$ and $B \to \bar{B} \to \bar{f}$.

One may separate three different sources of CP violation in the SM:

- the CP violation in the decays, also referred to as *direct* CP violation; it occurs when the rate for a process differs from the rate for its conjugate one, and therefore the amplitude ratio gives $|A_f/\bar{A}_{\bar{f}}| \neq 1$; it has a non-vanishing effect also in time-integrated measurement;
- the CP violation in the mixing, which happens when the coefficients of the two opposite

flavoured neutral meson states, describing the rotated mass eigenstate, are not equal, $|M_{mass}\rangle = p|M\rangle + q|\bar{M}\rangle$, and $|p/q| \neq 1$;

- the CP violation in the interference between mixing and decay, also known as the mixing-induced CP violation. This is the case described above and exploited in the *time-dependent* method for measuring CP asymmetries.

CP violation was first observed in the neutral K system in 1964 by the Cronin and Fitch experiment, which measured the CP violation in the $K^0 - \bar{K}^0$ mixing at the level 0.003. The CP asymmetry in the $B^0 - \bar{B}^0$ system was expected to be far larger, but the branching fractions of the sensitive channels were the main limit for its experimental observation and for this reason the B Factory concept was developed (Section 2). The measurements in the B sector, which confirm the KM model, have been carried out at the B Factories, using the time-dependent method: it is based on the measurement of the interference betwen conjugate decays to identical final CP-eigenstates. In the golden channel $B^0 \rightarrow J/\Psi K^0_s$ it is proportional to the product of CKM matrix element, $V^*_{cd}V_{cb}V_{td}V^*_{tb}$. The complex phase of this product, β, corresponds to one of the internal angles of the unitarity triangle and its measurement provides the CP violation experimental evidence. The phase difference is proportional to the decay time difference, which has to be carefully measured, since time-integrated asymmetry vanishes. For this kind of measurement, specific experimental requirements must be fulfilled, which are further discussed in the next chapter.

1.4 Baryonic decays of the B meson

The large amount of B mesons[1] produced at the B Factories (see Chapter 2) not only allows to study CP asymmetries and put constraints on the CP violation predicted in the Standard Model, but it also enables other kinds of New Physics searches, such as studies of the rare decays of the B mesons. Among them, the baryonic decays of the B meson became very attractive in the early 1990s, because of the *baryons puzzle*. It rose after the inclusive[2] measurement of the branching fraction of charged and neutral B meson decaying to baryons, provided by Argus in 1992 [5]. They measured that the probability for finding baryons as final state of B meson decays was $6.8 \pm 0.6\%$. This result has to be compared with the sum of all the exclusive baryonic modes that have been measured since then: the average on B^+, B^0, up to now, gives less then 1%.

A further motivation for studying baryonic decays comes from the large mass of the B mesons, which allows a varied spectrum of baryons to be produced. Baryonic B decays are the optimal place to search for exotic baryons production and at the same time investigate the mechanism of hadronisation into baryons. Similarly to what happens with the jet fragmentation, the production of a pair of baryon-antibaryon from a meson requires at least a pair of quark-antiquark produced out of vacuum. So far, this process is only qualitatively understood and

---

[1] A quark-antiquark buond state containing the quark $\bar{b}$ (or b for $\bar{B}$ meson) and a light down-type (anti)quark.

[2] In experimental particle physics, inclusive means that the hadronic final state from the B meson decay is not reconstructed in a specific channel.

theoretical models — the QCD sum rule [6], the pole model and the pQCD approach [7] — need validation and input from experimental results.

B mesons can only decay weakly and the size of the CKM matrix element V_{cb}, whose module is ten times the off-diagonal element V_{ub}, makes the $B \to X_c$ transition the dominant process. Most measured final states show at least a charmed meson among their final products. Moreover, charmless two-body baryonic decays haven't been observed yet, while the first charmless three-body decay was observed in $B^+ \to p\bar{p}K^+$ [8]. This rare process highlighted two peculiar features that have been repeatedly observed in several baryonic B decays: a peak in the baryon-antibaryon mass distribution at low values, nearly at threshold, and the increasing of the branching fraction with the final state multiplicity. In literature, they are usually referred to as the *threshold enhancement* and the *multiplicity effect* [4, 9]. These peculiar aspects are still something that has to be better understood and provide further motivations for the rare baryonic B decay searches.

1.4.1 Baryonic hadronization in B-meson decays

The B meson decays proceed through the spectator quark tree diagrams involving the $b \to c$ transition in almost 99% of the cases. The complexity of diagrams in the B meson decays into baryons increases with the multiplicity of the final state and a full discussion of all the contributions is beyond the scope of this thesis. The Feynman diagrams for the two most basic processes involved are shown in Figure 1.4: the b quark to $c(u)$ transition with a W-emission, and the quark-antiquark vacuum production. It can be shown that all contributing Feynman diagrams can be divided into two classes [9]:

1. the meson-meson class; the B meson decay into a meson pair and one of the two mesons decays again into a baryon-antibaryon pair, requiring the production of two pairs of quark-antiquark from the vacuum. There is no colour suppression, as no colour-matching is needed in this case to form a colourless baryon. Furthermore no hard gluons are required, since momentum is carried away by the *spectator* meson. This should be the kind of processes responsible for the threshold enhancement effect;

2. the diquark-antidiquark class; no intermediate meson state is produced and colour matching with the quark-antiquark additional pairs is required in order to form colour-singlet baryons. This configuration may explain the two-body final state suppression among the baryonic B decay modes.

Models for a four-body decay of the B-meson into baryons have not been specifically developed. For larger multiplicities, subsequent decays of pseudo-mesons into baryon-antibaryon pairs are considered, with additional quark-antiquark pairs produced by gluons annihilation out of the vacuum.

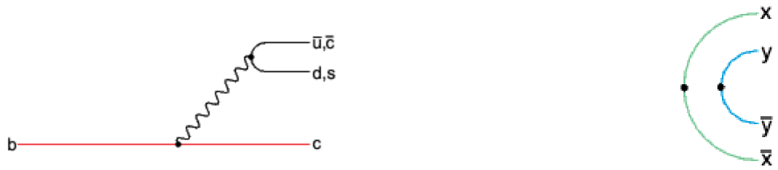


Figure 1.4: Main processes involved in B decays into baryons: the b quark transition to c, u with the W-emission on the left and the quark-antiquark pairs production out of the vacuum on the right.

1.4.2 Phenomenological interpretation

An experimental fact is that two-body baryonic decays are largely suppressed (more than one order of magnitude) with respect to three-body modes and observations confirm the general trend of the branching fraction increase with larger multiplicities of the final state [9].

Always from experimental results, the threshold effect is mainly observed in charmless three-body decays [8]. This effect might show a correlation with the $b \to u$ transition and it points out that favored decays of the B meson to charmless baryons are those where:

- the baryon-antibaryon pair peaks at low invariant mass;
- the dibaryon couple is accompanied by a fast recoiling meson.

Several approches have been suggested for explaining these observations, but a clear understanding of baryonic production from B mesons is still missing. Most of the proposed models rely on the *short-distance picture* and on the explanation for the two-body suppression in terms of hard and soft gluon production, as summarized in the two diagrams of Figure 1.5. Roughly, the interpretation for a baryon-antibaryon pair production in a two-body decay is that hard glouns (highly off mass-shell) are required and since the decay amplitude can be written proportionally to α_s/q^2, where α_s is the strong coupling constant and q^2 the transfered momentum carried by glouns, these modes are highly suppressed ($1/q^2$).

On the other hand, having at least one recoiling meson allows a different configuration for the decay topology (see diagram (b) in Figure 1.5), in which the baryon-antibaryon pair is emitted collinearly against the meson and the gluon is now moving in the same direction of the quark-antiquark pair, very close to its mass shell and therefore not suppressed by the short-distance term. The soft gluon emission implies also that the dibaryon pair is produced near its mass threshold. This is a possible explanation for the enhanced rate of higher multiplicity modes, with relative low mass final state.

1.5 The $B^0(\bar{B}^0)\to pp\bar{p}\bar{p}$ decay

As a first test for the above mentioned interpretation, a search for the the decay $\bar{B}^0 \to \Lambda_c^+ p\bar{p}\bar{p}$ was performed by *BABAR* [10]. The motivation was to check if the reduced phase space factor in the amplitude for 4 baryons production from a B meson decay could be compensated by the threshold enhancement. No event was detected, however the analysis put an upper limit at a 90% confidence level to the branching fraction for the mentioned mode, 2.8×10^{-6}.

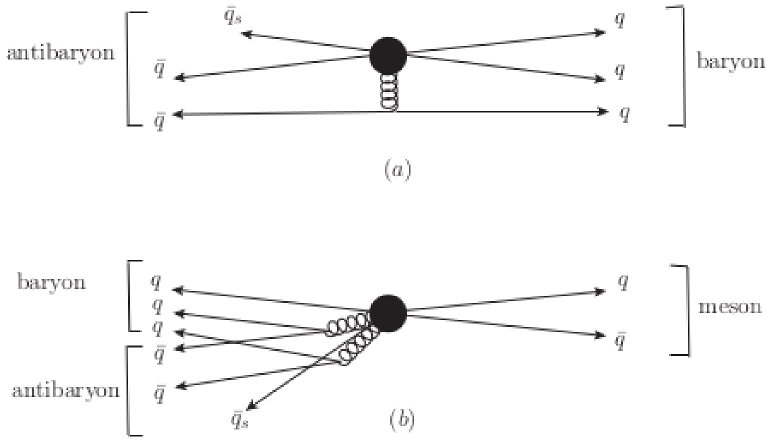


Figure 1.5: Figure taken from *The Physics at the B Factories* [4]. The short-distance picture for threshold enhancement and multiplicity effect explanation is shown in the two diagrams, respectively for a two-body (a) and a three-body (b) baryonic decay.

The observation of the thershold effect seems to be related to the energy available in the phase space for a given decay and experimental results might indicate that no enhancement has been found with less than 200 MeV available.

These considerations lead to the idea of the measurement of the branching fraction of $B^0(\bar{B^0})\rightarrow pp\bar{p}\bar{p}$ decay. The baryonic decay that I am going to study in this thesis will provide further experimental support for the interpretation of these peculiar features of the baryonic decays. In fact, a non-vanishing rate for this mode might help in validating the phenomenological explanation of the threshold enhancement and the related multiplicity effect.

In the remaining part of this chapter, starting from the previous analysis of the mode $\bar{B}^0 \rightarrow \Lambda_c^+ p\bar{p}\bar{p}$ at BABAR, some general expectations on $B^0(\bar{B^0})\rightarrow pp\bar{p}\bar{p}$ are provided.

A comparison of the two modes is summarized in Table 1.1. A rough estimate of the branching fraction expected for the channel $B^0(\bar{B^0})\rightarrow pp\bar{p}\bar{p}$ can be derived, starting from the assumptions:

- the decay amplitude is proportional to the squared module of the CKM matrix element involved in the flavour transition $b \rightarrow u$;
- the phase space factor enters the decay amplitude as a function of the transfered momentum, q;

Therefore the following contributions are evaluated:

- the ratio $\frac{|V_{ub}|^2}{|V_{cb}|^2} \sim 10^{-2}$;
- the Q-values ratio $\frac{Q_{pp\bar{p}\bar{p}}}{Q_{\Lambda_c^+ p\bar{p}\bar{p}}} \sim 8.2$;

where the Q-value of a decay is defined as the difference between the mother particle mass and the sum of the masses of the decay products, related to the maximum transfered momentum in the kinematic of the decay. In light of these considerations, the branching fraction

	Mode	
	$\bar{B}^0 \to \Lambda_c^+ p\bar{p}\bar{p}$	$B^0(\bar{B}^0)\to pp\bar{p}\bar{p}$
Weak Interaction [11]	$b \to c$ $V_{cb} = (41.1 \pm 1.3) \times 10^{-3}$	$b \to u$ $V_{ub} = (4.13 \pm 0.49) \times 10^{-3}$
Phase space (Q_{value})	$Q_{\Lambda_c^+ p\bar{p}\bar{p}} = 0.186$ GeV/c^2	$Q_{pp\bar{p}\bar{p}} = 1.52$ GeV/c^2
Reconstruction efficiency	$\epsilon = 3.5\%$	$\epsilon \sim 40\%$

Table 1.1: Comparison between $\bar{B}^0 \to \Lambda_c^+ p\bar{p}\bar{p}$ and $B^0(\bar{B}^0)\to pp\bar{p}\bar{p}$ decays. The weak coupling, the approximate phase space contribution and the reconstruction efficiency are listed for each decay mode.

$B^0(\bar{B}^0)\to pp\bar{p}\bar{p} \approx (\bar{B}^0 \to \Lambda_c^+ p\bar{p}\bar{p}) \cdot \frac{|V_{ub}|^2}{|V_{cb}|^2} \cdot \frac{Q_{pp\bar{p}\bar{p}}}{Q_{\Lambda_c^+ p\bar{p}\bar{p}}} \sim 10^{-7}$ can be estimated. The scaling factor $\frac{Q_{pp\bar{p}\bar{p}}}{Q_{\Lambda_c^+ p\bar{p}\bar{p}}}$ relies on the assumed linearity of the phase space element with respect to the available kinetic energy (quantified by the Q_{value}): this approximation is introduced to estimate the enlargement of the decay amplitude due to the lower masses of the final state particles ($m(\Lambda_c) = 2286.46$ MeV/c^2, $m(p) = 938.27$ MeV/c^2 [11]). This estimate is assumed as a working hypothesis throughout the rest of this thesis.

Moreover, from preliminary MC studies a factor 10 improvement is expected for the reconstruction efficiency of the $B^0(\bar{B}^0)\to pp\bar{p}\bar{p}$ decay with respect to the $\bar{B}^0 \to \Lambda_c^+ p\bar{p}\bar{p}$ mode, coming from a good effciency for four charged track reconstruction at high momentum (1 GeV/c per track), while the reconstruction of Λ_c into the channel π, p, K with BF of 5%, has an efficiency of 3.5% as shown in the last row of Table 1.1. As discussed in Chapter 7, after the final event selection 10 signal events are expected in the BABAR data sample.

The measurement of the $B^0(\bar{B}^0)\to pp\bar{p}\bar{p}$ decay will certainly not resolve the puzzle of the inclusive baryonic B decay branching fraction, but it can provide useful insight in the still obscure mechanisms of these transitions, and help improve the theoretical description. Moreover, it is the first search for a baryonic channel with a 4-baryon final state, apart from the $\bar{B}^0 \to \Lambda_c^+ p\bar{p}\bar{p}$ decay study, and it would be the first time this upper limit is measured.

Chapter 2

Experiments at the B Factories

Before 1964, it was known that the fundamental symmetries of charge coniugation (C) and parity (P) were individually violated by the weak interaction, but no evidence for a CP violation had been found. The revolutionary result of the experiment by Cronin and Fitch on the neutral K system provided the first experimental evidence that the laws of Nature are not the same for matter and antimatter and CP symmetry is violated. In 1973 Kobayashi and Maskawa showed in their paper that CP violation could naturally arise in Standard Model thanks to the irreducible complex phase in the weak interaction quark-flavor-mixing matrix, if only the quark flavors were six [3]. At that time, only the three lightest quarks—u, d, s—were experimentally known. By 1980, after the *charm* and *beauty* quark discovery, the Cabibbo-Kobayashi-Maskawa quark flavor-mixing matrix was embodied in the Standard Model, providing predictions for the size of CP violation in the charm and beauty sectors. The CLEO experiment, situated at Cornell (New York), performed measurements on data samples of a few tens of $B\overline{B}$ events; large CP violation was expected in neutral B meson decays to CP eigenstates (*e.g.*, $B^0 \to J/\Psi K_s^0$), but no evidence could be found due to a branching fraction lower than 0.1%.

Few years later, the measurement of the long B meson lifetime (1983, SLAC) and the first observation of B meson oscillations (1987, ARGUS) confirmed that CP violation could be successfully tested in the $B^0 - \bar{B}^0$ system. However, experiments could achieve such a goal only with a large amount of B mesons. With this purpose in mind, the B Factory concept was born and BABAR (1999-2008) and Belle (1999-2010) experiments were built [4]. Nowadays the SuperKEKB/Belle II upgrade program of the japanese B Factory represents the future of this kind of experiments.

The main features of those detectors were dominated by the goal, measuring the CP asymmetries, exploiting the time-dependent method (see Section 1.3). In addition, they also performed precise measurements of the CKM matrix elements and of several branching fractions of rare B meson decays, enabling constraints on the Standard Model parameters to be set. These experiments also allow the study of the charm and tau lepton physics, as well as two-photons physics.

The concept of a B Factory is explained in Section 2.1, where the asymmetric energies method for boosting the B mesons is described; general features of $B\bar{B}$ events are discussed and also kinematical variables implemented to reduce the backgrounds are defined.

Section 2.2 focuses on the performance of the accelerators and the samples produced at the B Factories. Integrated luminositites for each machine are given and the KEKB upgrade to a Super Flavour Factory is illustrated.

An overview of the BABAR detector, which collected the data set used for the analysis presented in this thesis, is provided in Section 2.3. The last part of this chapter (Section 2.4) illustrates the main picture for the future of the B Factories, represented by the Belle II experiment. A description of the detector is given. In the next chapter, my active contribution to the Belle II experiment construction, as member of the Silicon Vertex Detector group in Pisa, will be reported.

2.1 The B production at $\Upsilon(4S)$

B hadrons can be produced both in e^+e^- interactions and in hadronic interactions. While the b hadron production at high energy (for instance in $p\bar{p}$ collisions at 13 TeV at LHC) offers a large cross-section and a full spectrum for beauty mesons and baryons, electron-positron colliders provide an extremely clean environment and a well known and coherent intial state.

The BABAR and Belle experiments were installed on the high luminosity e^+e^- colliders PEP-II at SLAC, in California, and KEKB at KEK, in Tsukuba (Japan) and collected data for almost a decade, starting from 1999 until 2008. Today, the future of the high-luminosity experiments is represented by the upgraded accelerator SuperKEKB and the detector Belle II, installed at KEK.

In order to have the large amount of $B\overline{B}$ mesons, these colliders operate at a centre-of-mass (CM) energy of 10.58 GeV, corresponding to the invariant mass of the $\Upsilon(4S)$ resonance. This resonance is a bound state of the beauty quark b and its antiparticle, the antiquark $\bar{b}$ (*bottomonium*) with quantum number $J^{CP} = 1^{--}$. The $\Upsilon(4S)$ resonance is the first one of the bottonium family above the threshold for B meson pair production and mainly decays to neutral or charged $B\overline{B}$ pairs, roughly in the same proportion. Since the mass of the B meson is 5.279 GeV/c^2 [11], they are produced almost at rest in the CM frame and because of their short lifetime ($\tau = 1.519 \times 10^{-12}$ s [11]), they travel a very reduced distance in the laboratory. Therefore, the measurement of their decay vertices separation ($\sim$ 60 μm) would be hardly achivable, given the state-of-the-art of vertex detectors. The original asymmetric B Factory concept, which was proposed by P.Oddone in 1988 [12], uses different energies for the electron and positron beams so that the $\Upsilon(4S)$ acquires a boost in the laboratory frame and the decay products are pushed forward in the direction of the electron beam. From the now measurable decay lengths $\Delta z = \beta\gamma c\tau$, proper decay times can be inferred. These machines also required unprecedented luminosities, of order 10^{33} cm^{-2} s^{-1}: this is why they were called *factories*.

B meson production at in e^+e^- interactions at the $\Upsilon(4S)$ has several advantages: first of all, the high signal-to-background ratio, $\sigma_{b\bar{b}}/\sigma_{TOT} = 0.28$. In Table 2.1, the list of cross sections for different processes taking place at the CM energy of $\Upsilon(4S)$ resonance is shown. Second, clean events with a higher mean charged multiplicity ($\sim$ 11) with respect to the backgrounds are expected, as shown in Figure 2.1. The event rate (10 Hz) is also quite low and allows relative simple trigger strategies. Another advantage at a B Factory is the hermeticity of the detector and the feasibility of the neutral final state detection, where π^0 and photons are involved. The absence of fragmentation products and the kinematics of the $\Upsilon(4S)$ decay also reduces the combinatorial backgrounds. In fact, the initial state and also the momentum magnitudes of B mesons in the CM frame are completly known, allowing therefore to set kinematical constraints on reconstructed candidates, which results in a very powerful tool for suppressing backgrounds.

$e^+e^- \to$	Cross-section (nb)
$b\bar{b}$	1.05
$c\bar{c}$	1.30
$s\bar{s}$	0.35
$u\bar{u}$	1.39
$d\bar{d}$	0.35
$\tau^+\tau^-$	0.94
$\mu^+\mu^-$	1.16
e^+e^-	$\sim$ 40

Table 2.1: Production cross-sections at $\sqrt{s} = m(\Upsilon(4S)) = 10.58$ GeV/c^2.

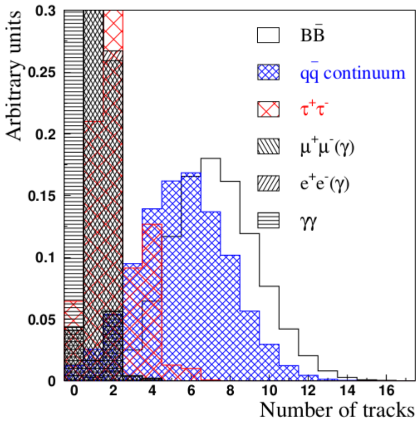


Figure 2.1: The number of charged tracks per event is shown, for different processes. Histogram taken from *The physics at B Factories*, Ch. 7, B-meson reconstruction [4].

2.1.1 Kinematical constraints on $B\overline{B}$ events

The B Factory experimental setup imposes constraints which improves both the knowledge of the B meson's momentum and the discrimination between signal and background. In the centre of mass (CM) frame, since the $\Upsilon(4S)$ resonance decays in two same-mass particles, the energy and the mass of a correctly reconstructed B meson candidate must satisfy:

$$E^*_{rec} = E^*_{beam} = \sqrt{s}/2 \qquad\qquad m_{rec} = m_B \tag{2.1}$$

where the * quantities refer to the CM frame, *i.e.* , the rest frame of $\Upsilon(4S)$, and $\sqrt{s}$ is the energy of the e^+e^- system. From this consideration, two variables are introduced, known as the energy difference ΔE and beam-energy substituted mass m_{ES}. In the $\Upsilon(4S)$ rest frame, the former is described by the following formula:

$$\Delta E = E^*_B - E^*_{\text{beam}} \tag{2.2}$$

E^*_B is the energy for the reconstructed B meson and E^*_{beam} is the beam energy. The resolution of ΔE receives two contributions from the spread of the reconstructed B-meson energy and from the the beam energy spread. Usually the first component, especially for those final states involving photons, is the dominant one and it is related to the detector energy resolution. This variable is very slightly correlated to the beam-energy substituted mass, whose expression evaluated in the $\Upsilon(4S)$ rest frame is:

$$m_{ES} = \sqrt{(E^{*}_{beam})^2 - |\vec{p_B}^{*}|^2} \tag{2.3}$$

where the $|\vec{p_B}^{*}|$ is the CM frame three-momentum of the B meson and the E^{*}_{beam} corresponds to the energy of beams in the CM frame, which replaces the B meson reconstructed energy. The Lorentz invariant expression of m_{ES} is given by:

$$m_{ES} = \sqrt{\left[(s/2 + \vec{p_i} \cdot \vec{p_B})^2 / E_i^2\right] - |\vec{p_B}|^2} \tag{2.4}$$

where $(E_i, \vec{p_i})$ and $(E_B, \vec{p_B})$ are the four-momenta of respectively the e^+e^- system and the reconstructed B candidate in the laboratory frame.

The uncertainty on this variable has contributions from $\sigma_{E^{*}_{beam}}$ and from $(\vec{p_B}^{*}/m_B)\,\sigma_{p^{*}_{B}}$. Since B mesons are nearly at rest in the CM frame ($\vec{p_B}^{*}/m_B \approx 0.06$), the uncertainty is dominated by the beam energy spread. The typical resolution is less than 4 MeV, when neutrals are not involved in the final state, whereas ΔE resolution depends on the analysed mode and it goes from 30 MeV for low mass particles to ~10 MeV for heavier particles.

The two variables exploit complementary constraints from the event kinematics: m_{ES} is independent on the mass hypothesis since it only implies three-momenta reconstruction, while ΔE relies on the energy of the reconstructed B meson, which needs both the momenta and the mass assignments to be calculated. Assigning the wrong particle identity to a reconstructed track (*mis-identification*) produces a shift in ΔE, but it should not affect m_{ES} distribution: further details on how they are used for rejecting backgrounds are provided in Chapter 4. Some examples of m_{ES} and ΔE distributions are shown in Figure 2.2.

Finally, these variables usually show no relevant correlation, except for some specific cases. One of them is represented by the B meson decays to heavy particles, *i.e.* , B meson baryonic decays, in which other pairs of variables are preferred for analysis based on the cut-and-count method in the ΔE-m_{ES} space, for instance the m_B-m_{ES} plane. Since the analysis here presented is not going to exploit that technique, even though it deals with a B meson baryonic decay, the kinematics variables used are m_{ES} and ΔE. In Figure 2.3, scatter plots in both ΔE-m_{ES} and m_B-m_{ES} planes are shown for the baryonic B meson decay $\bar{B}^0 \to \Lambda_c^+ \bar{p} \pi^+ \pi^-$.

2.2 The B Factories data samples

At both the accelerators, in order to achieve asymmetric energies, two dedicated rings were needed: the High Energy Ring (HER) for electrons, and the Low Energy Ring (LER) for positrons. A schematic view of PEP-II and KEKB colliders is shown in Figure 2.4. For luminosity optimization, both the facilities had a single Interaction Region (IR): unusual beam optics very close to the Interaction Point (IP) became necessary to accomodate for high luminosity requirements. In terms of the collider parameters, the luminosity is expressed by:

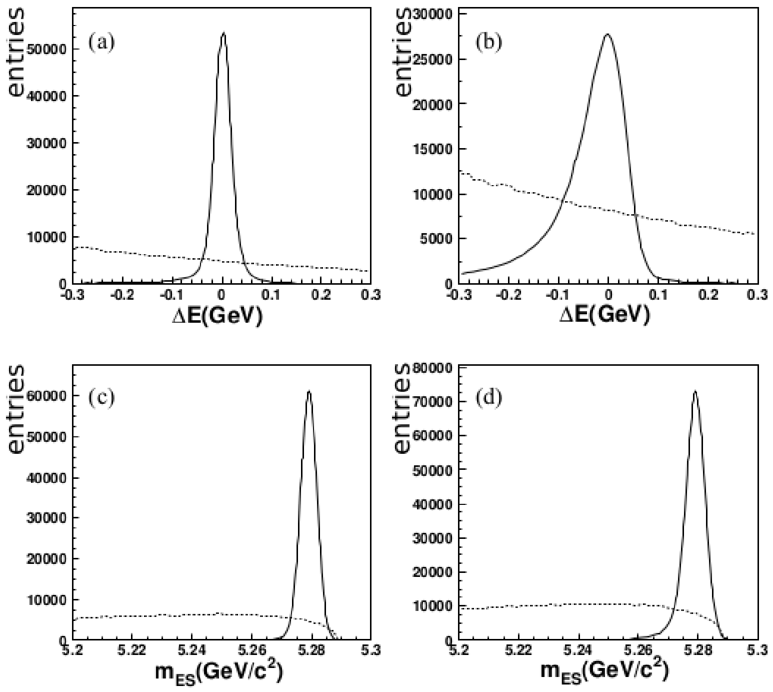


Figure 2.2: The ΔE and m_{ES} distributions for $B^{+} \rightarrow K^{0}_{s}\pi^{+}$ (left) and for the mode $B^{+} \rightarrow K^{+}\pi^{0}$ (right). Solid line histograms represent signal events from the official MC production, dotted lines stand for the continuum events. Plot (b) shows a worse resolution in ΔE compared to plot (a), due to the two photons in the final state (π^{0} decay). The m_{ES} distribution for the same mode, plot (d), shows a better resolution due to the substitution of the reconstructed B meson energy E^{*}_{B} with the beam energy of the event, as explained through this Section.

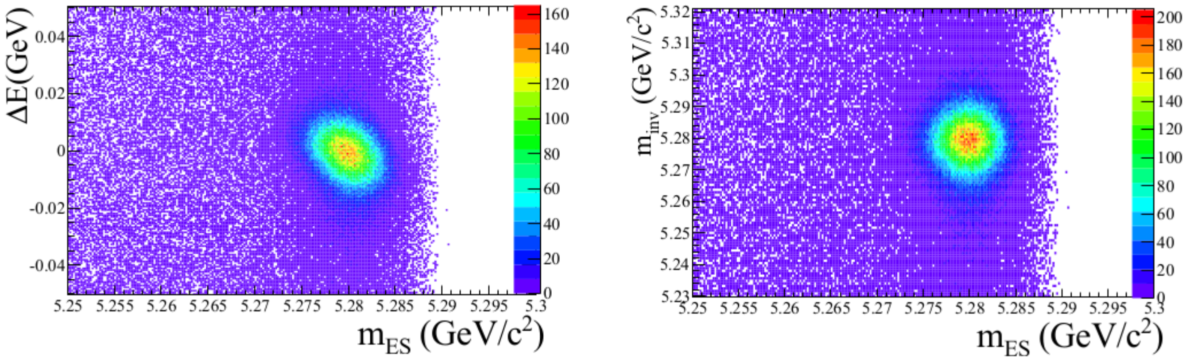


Figure 2.3: Scatter plots for $\Delta E(m_{B})$ versus m_{ES}. Left: ΔE vs m_{ES} is shown for the baryonic mode $\bar{B}^{0} \rightarrow \Lambda^{+}_{c}\bar{p}\pi^{+}\pi^{-}$, where the above discussed correlation is visible in the rotation of the ellipse representing the signal. Right: the invariant mass m_{B} vs m_{ES} plot for the same mode is shown; the variables are only weakly correlated through measurement errors.

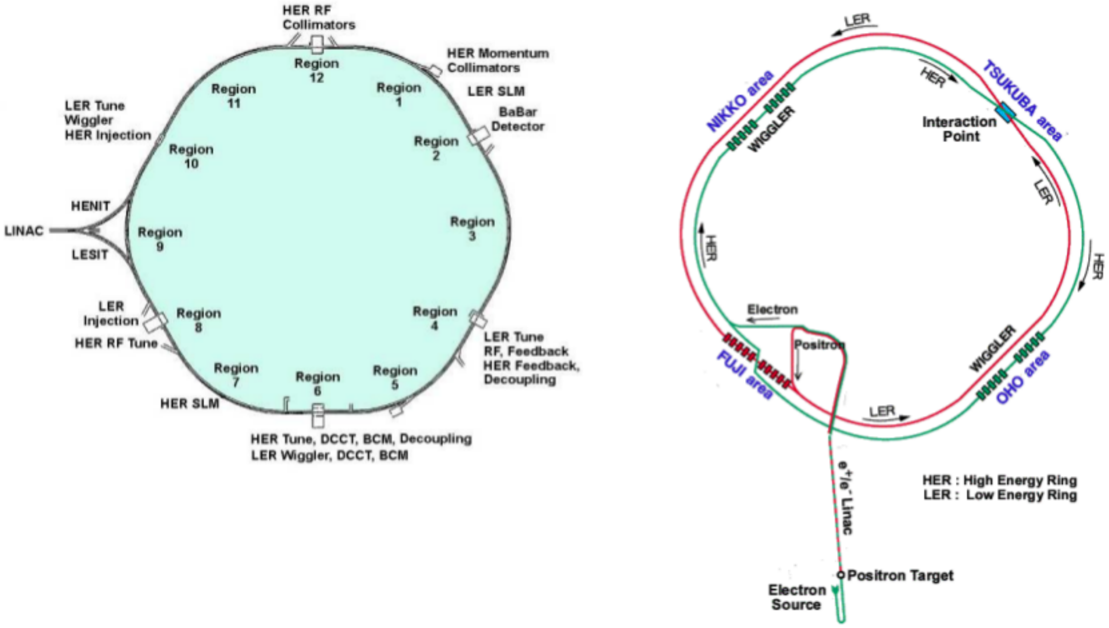


Figure 2.4: PEP-II (left) and KEKB (right) e^+e^- storage rings.

$$\mathcal{L} = \frac{N_b n_{e^-} n_{e^+} f}{A_{eff}} \tag{2.5}$$

where N_b is the number of bunches, n_{e^-} and n_{e^+} respectively the numbers of electrons and positrons per bunch, f the circulating frequency, and A_{eff} the effective overlapping in the transverse area of the two beams at the interaction point. To ensure good performances and limit the damaging effect of sinchrotron radiation and high beam currents, a high quality vacuum was needed throughout all the beam pipe.

In addition, it was fundamental to introduce the beam-separation scheme, in order to handle the short bunch spacings in a two ring machine. At PEP-II interaction point, after the head-on collision the separation was achieved thanks to the bending magnets, which diverted the electron and positron bunches soon after their intersection, protecting the beams from parasitic interaction, while KEKB introduced a small crossing angle ($\sim$ 11 mrad), which simultaneously allowed to have shorter bunch spacings and more space available for detector components closer to the IP.

This was not without risk, since the crossing angle induces coupling effect between transverse betatron oscillations and longitudinal sinchrotron oscillations, inflating instabilities which were finally cured allowing KEKB to reach an istantaneous luminosity even twice larger than the nominal originally designed, for a final peak of 2.1×10^{34} cm^{-2} s^{-1}.

In Figure 2.5 the luminosities achieved at PEP-II and at KEKB during different data-taking periods are shown and in Table 2.2 the major parameters of the machines for both the factories during their final stage of operation are given.

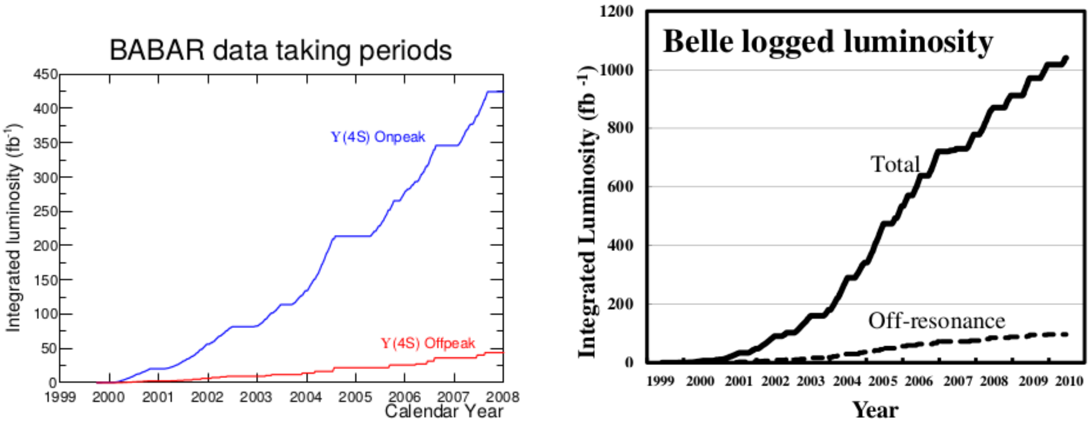


Figure 2.5: The *BaBar* (left) and Belle (right) luminosities integrated over the $\Upsilon(4S)$ running period, for both the on- and off-resonance operational modes, are shown. Pictures are taken from *The physics of the B Factories* [4].

Parameters	Units	PEP-II	KEKB	SuperKEKB
			[HER (e^-) / LER (e^+)]	
Beam Energy	(GeV)	9.0/3.1	8.0/3.5	7.0/4.0
Beam Current (I)	(A)	1.8/2.7	1.19/1.64	2.62/3.60
Beam Size at IP (x)	(μm)	140	80	11.2/10.2
Beam Size at IP (y)	(μm)	3	1	0.0618/0.0483
Beam Size at IP (z)	(mm)	8.5	5	5/6
Luminosity	(10^{34}cm^{-2}s^{-1})	1.2	2.1	80
ξ_y			0.090/0.129	0.088/0.090
β_y^*	(mm)		5.9/5.9	0.41/0.27
Lorentz factor ($\beta\gamma$)		0.56	0.43	0.28
Half-crossing angle	(mrad)	0	11	41.5

Table 2.2: Main machine parameters for PEP-II and KEKB during their final stage of operation; in the last column values from Belle II Technical Design Report [13] for SuperKEKB upgrade are listed.

		on-peak		off-peak
		Luminosity (fb^{-1})	$\Upsilon(4S)$	Luminosity (fb^{-1})
PEP-II		424.2	$(471.0 \pm 2.8) \times 10^6$	43.9
KEKB	(SVD-1)	140.0	$(152 \pm 1) \times 10^6$	15.6
	(SVD-2)	571.0	$(620 \pm 9) \times 10^6$	73.8

Table 2.3: Integrated luminosities on- and off-peak and the number of $\Upsilon(4S)$ particles, corresponding to *BABAR* and Belle data sets, are listed.SVD1 and SVD2 refer to the vertex detector upgrade perfromed at KEKB, which defined two different operational stages of the detector.

2.2.1 Integrated luminosity

PEP-II operated mainly at the CM energy of the mass of $\Upsilon(4S)$, collecting the *on-peak* events: the design luminosity was already reached at the end of the first data-taking period, begun in 1999 (Run 1). Afterwards, other six periods of data-taking followed, from Run 2 to the last Run 7. Initially the machine was operated in *coasting mode*, alternating data-taking periods, when the beams would gradually reduce their intensity, to injection periods, when the beam currents would be restored to their initial values. However, since the beginning of Run 4, significant improvements in accelerator stability and in a reduction of injection backgrounds enabled to adopt the *trickle injection* mode, whose aim was to keep LER and HER currents beam to a constant level by continuosly injecting a small amount of electron and positron to compensate for the losses, without interrupting the data-taking.

About 10% of data were also collected at a CM energy of $\sim$ 40 MeV below the $\Upsilon(4S)$ threshold (the *off-peak* mode): since the width of this resonance is 20 MeV, the CM energy shift ensured to collect the *continuum* events, which were not coming from $\Upsilon(4S)$ decays. The luminosity calculation for Run 1-6 was performed measuring the BhaBha scattering ($e^-e^+ \to e^-e^+$) and dimuon ($e^-e^+ \to \mu^-\mu^+$) events. The time-integrated luminosities at the $\Upsilon(4S)$ resonance (as well as the number of Υ particles produced) and for its corresponding continuum region, for both *BABAR* and Belle, are listed in Table 2.3.

2.2.2 KEKB upgrade to a SuperB Factory

The Belle/KEKB upgrade to Belle II/SuperKEKB started with the accelerator upgrade. The new collider will provide 40 times the old luminosity thanks to the technology first proposed for Super B in Italy, the so-called *Nano-Beam* scheme. It exploits the inverse proportionality that relates the luminosity to the vertical beta function β_y^{*}[1]: for increasing the luminosity, the β_y^* squeezing at the IP is provided. Valuable improvements in several machine parameters are needed to achieve this goal.

The main modification regards the minimization of the longitudinal size of the beams overlap d (schematic view of the beams intersection in the *Nano-Beam* scheme is reported in Figure 2.6),

---

[1] In Accelerator Physics, the betathrone function describes the displacement of a particle beam due to oscillations in the transverse plane xy with respect to its nominal trajectory in a given point s along the collision axis. β^* denotes the value of the β function at the interaction point.

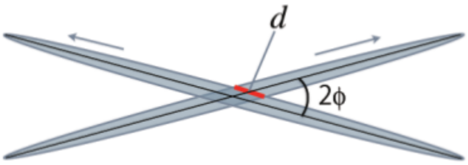


Figure 2.6: The beam collision in the Nano-Beam scheme: the effective beam size d and the half-crossing angle ϕ are illustrated. Picture taken from the Technical Design Report [13].

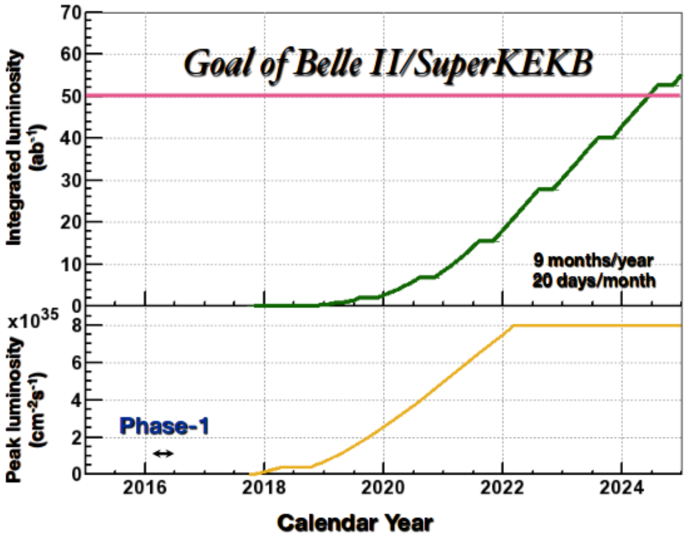


Figure 2.7: The expected integrated and peak luminosity for the upgrade Belle II/SuperKEKB as function of time. Picture taken from Belle II official website [14].

which can be shown to be the lower bound for the β_y^*. The formula for d is the following:

$$d \simeq \frac{\sigma_x^*}{\phi} \tag{2.6}$$

The numerator σ_x^* represents the horizontal beam size at the IP and ϕ is the half-crossing angle ($\sim$ 41.5 mrad). The crossing angle is doubled from KEKB value of 22 mrad to 83 mrad. If the assumption of flat beams with the same horizontal and vertical size is verified, the luminosity formula at the IP can be written as:

$$\mathcal{L} = \frac{\gamma_\pm}{2er_e}\left(1+\frac{\sigma_y^*}{\sigma_x^*}\right)\left(\frac{I_\pm \xi_{y\pm}}{\beta_{y\pm}^*}\right)\frac{R_L}{R_{\xi_y}} \tag{2.7}$$

where γ, e, r_e are respectively the Lorentz factor, the elementary electric charge and the electron classical radius, while R_L, R_{ξ_y} are the luminosity and vertical beam-beam reduction factors, whose ratio can be approximately considered 1. Since the ratio σ_y^*/σ_x^* is order 10^{-3}, the term into brakets can be approximated to one as well. Indeces $(+,-)$ refer to electron and positron beam respectively.

The fundamental parameters for manipulating the luminosity are three: the beam current (I), the vertical beam-beam parameter (ξ_y) and the beta function (β_y^*) at the IP. A comparison of the major upgraded parameters between KEKB and SuperKEKB is given in Table 2.2. In Figure 2.7 the plot for the expected luminosity foreseen during Belle II/SuperKEKB operational stages is shown. The upgraded machine will be able to provide an istantaneous luminosity of $8 \times 10^{35}\mathrm{cm}^{-2}\mathrm{s}^{-1}$, achieving at end of its running the total designed integrated luminosity of 50 ab^{-1}, more than a factor 40 over what was achieved at KEKB. It should be noted that beam currents are increased by more than a factor 2, which will result in significantly higher backgrounds in the detector. In addition, mainly to limit the RF power needed for the HER, the energy of the HER has been reduced from 8 to 7 GeV, resulting in a smaller boost $\beta\gamma = 0.28$.

2.3 The *BABAR* detector

A detailed description of the *BABAR* detector can be found elsewhere [15]. The purpose of this section is just to provide a general overview, explaining how the detector features were determined by the physics goals of the experiment. A deeper insight into the tracking devices and the subdetectors involved in charged particle reconstruction and identification is also given.

2.3.1 The B physics requirements

Precision measurements need a very high statistics, which is achieved thanks to the combined action of a high luminosity machine and a very efficient detector. Moreover, CP-asymmetry measurements exploiting the time-dependent technique depend on three fundamental features:

- the full reconstruction of exclusive final states, that means excellent performance in track reconstruction[2] and particle identification;
- B flavour tagging, in order to distinguish one B meson from its charge conjugate;
- measurement of proper time of B decays.

The *BABAR* detector was designed to satisfy these specific requirements:

- the acceptance optimization; asymmetric energies imply also an asymmetric detector in order to provide a good coverage of the forward region where events are boosted.
- the optimal vertex resolution; the decay time difference in CP asymmetry measurements is derived from the difference between the z-coordinates of B meson primary and secondary decays. Excellent resolution requires to minimize multiple scattering, that means a minimal material budget for the inner part of the detector, closest to the interaction point;
- the low p_T partcles tracking, in the transverse momentum range of 60 MeV/c $< p_T <$ 4 GeV/c;

---

[2] The process of a charged particle track reconstruction from subdetector information (hits, energy losses) will be often refered to as *tracking*.

- the efficient discrimination for e, μ, π, K, p identification over the whole kinematics range available at a B Factory, the neutral final state detecting capability and energy resolution.

2.3.2 The *BABAR* detector overview

These goals were reached combining the performances of several subdetector systems:

1. the Silicon Vertex Tracker (SVT), which provides position hits produced by charged particles traversing its layers. It is the main tracking device for low-energy particles ($p_T < 180$ MeV/c) and provides a high precision measurement of the particle production vertex;

2. the Drift Chamber (DCH), filled with helium-based gas in order to minimize multiple scattering; it is the main tracking device for charged particles with $p_T > 180$ MeV/c, allowing a precise momentum determination, and it gives also contribution to particle identification (PID) with the energy loss measurements;

3. the Detector of Internally Reflection Cherenkov light (DIRC), it is the principle subsystem for hadron particle identification;

4. the Caesium Iodide Electromagnetic Calorimeter (EMC) provides an excellent energy measurement, as well as good electron identification down to 0.5 GeV and neutral hadron identification. Its asymmetric structure, with a forward endcap, but no backward coverage, reflects the asymmetry of event angular distribution in the laboratory frame due to the Lorentz boost $\beta\gamma \sim 0.56$.

5. the superconducting coil which produces the 1.5 T solenoidal magnetic field;

6. the Instrumented Flux Return (IFR), for muon identification down to 0.6 GeV and neutral hadron detection, especially K_L^0 used in exclusive reconstruction of cross-checking channels exploited in CP-asymmetries measurements.

A view of both longitudinal and front section of the *BABAR* detector is given in Figure 2.8

2.3.3 The *Silicon Vertex Tracker*

The Silicon Vertex Tracker (SVT) is composed of five layers: three are located close to the beryllium beam pipe and they perform impact parameter measurements and vertex measurements. The outermost two layers have larger radii fundamental for tracking purpose, which is performed thanks to the combined action of the SVT and the Drift Chamber. It also allows the SVT to perform stand-alone tracking measurements for low-p_T particles, unable to reach the Drift Chamber.

Each layer consists of several ladders (6 for the innermost ones, 16 and 18 respectively for layer 4 and 5) which are based on Double-sided Silicon Strip Detector (DSSD) modules. These modules are arranged into straight ladders for layers 1-3, while the ladders of the outermost

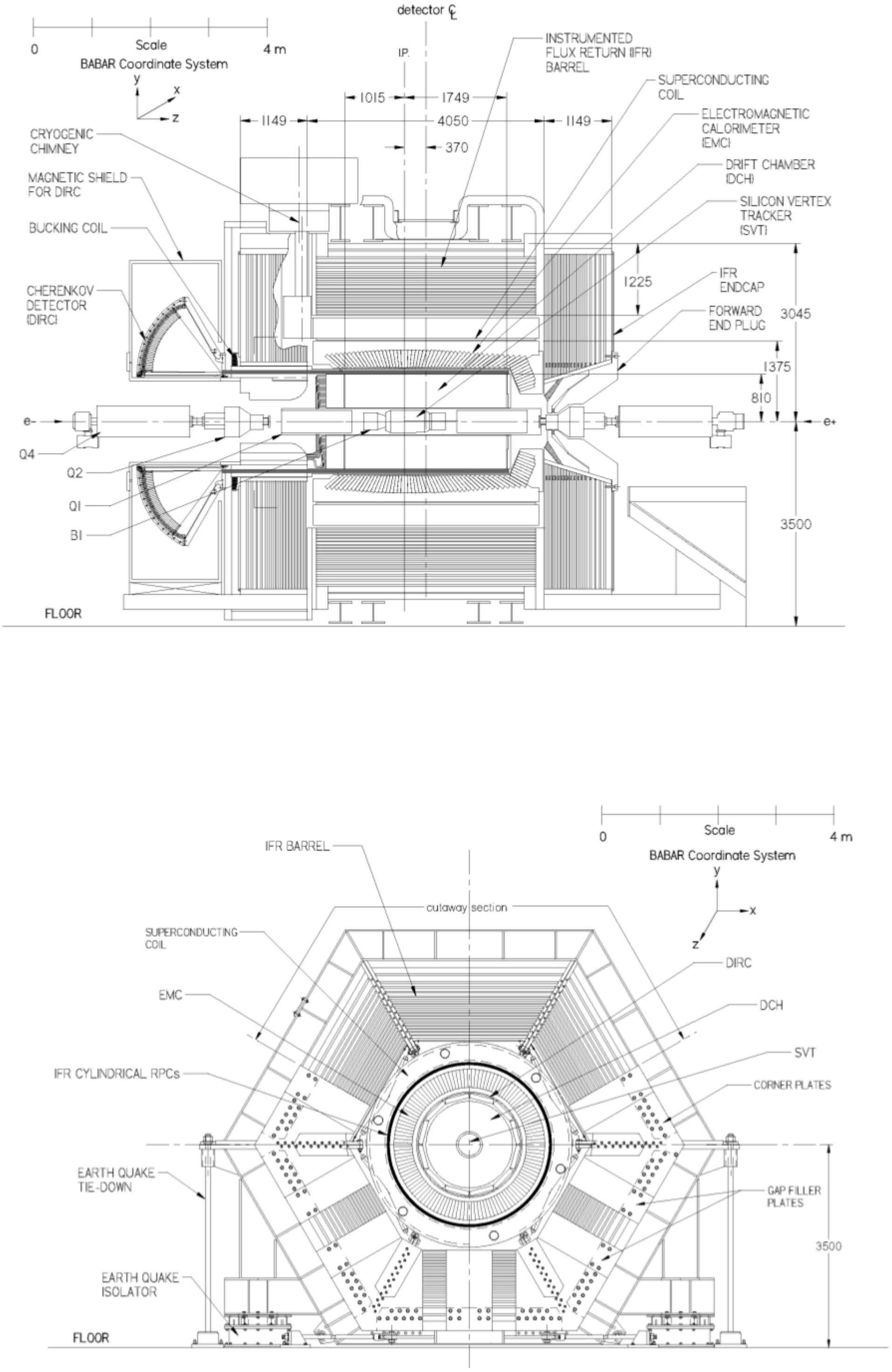


Figure 2.8: The longitudinal (top) and front (bottom) views of *BaBar* detector are shown. Picture taken from *The physics of the B Factories* [4].

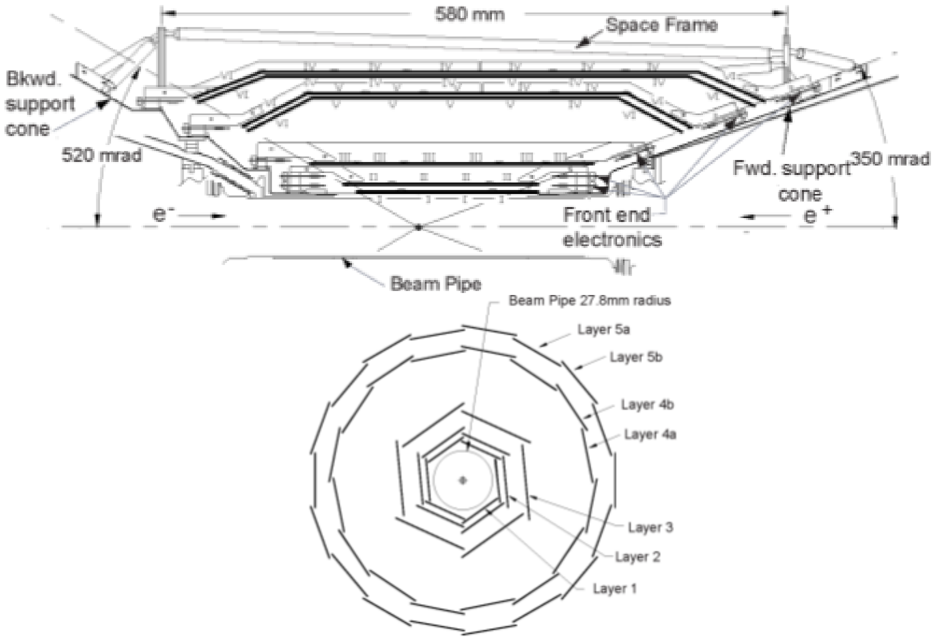


Figure 2.9: Schematic view of the longitudinal (top) and front (bottom) sections of the *BaBar* Silicon Vertex Tracker, taken from *The Phisics of the B Factories.*

layers are arch-shaped (forward and backward modules tilted). This configuration is aimed at minimizing the particle path lenght through silicon, limiting interaction probability and multiple scattering, as shown in the longitudinal view of the five layers in Figure 2.9. It also ensures an angular coverage in the polar angle θ from 20° to 150° in the laboratory frame, corresponding to a 90% solid angle coverage in the CM frame.

The SVT modules consists of double-sided silicon sensors electrically separated into two half-modules and readout at the ends. The ϕ side of the sensors has strips parallel to the beam axis in order to measure the azimuthal coordinate of the hit ($r\phi$), while the z-side has strips orthogonal to the beam direction and measures the coordinate along the beam axis (z). All 340 sensors are aligned *in situ* exploiting dimuon and cosmic ray events.

The average effciency of the SVT, calculated for every half-module as the ratio of the hits associated to each track with respect to the total number of tracks crossing that module, was around 95%. Energy-loss measurements allowed to reach a 2σ separation between kaons and pions up to 500 MeV/c momenta. Finally, the resolution on the measured coordinates depends on the measured quantity and on the polar angle of the crossing track, being maximum for 90° tracks and degrading in forward and backward regions: it corresponds to 15 μm for the z coordinate and 40 μm for azimuthal coordinate.

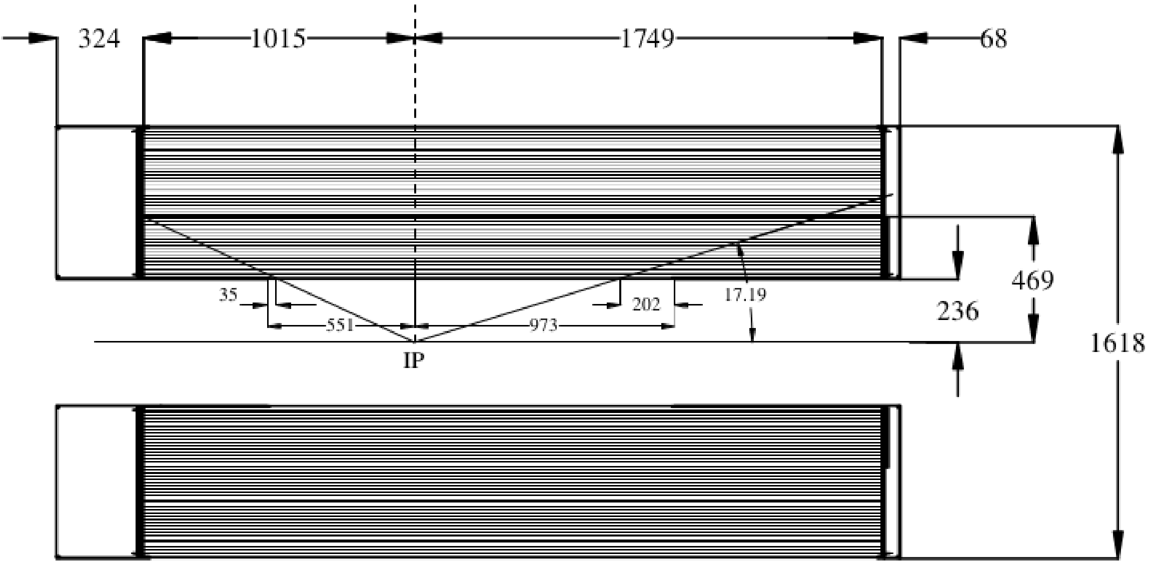


Figure 2.10: Schematic view of the longitudinal section of the *BABAR* Drift Chamber, with dimensions given in mm. Picture taken from *The BABAR physics book* [4].

2.3.4 The *Drift Chamber*

A longitudinal section of the Drift Chamber (DCH) in *BABAR* is shown in Figure 2.10. This subsystem was mainly involved in tracking and thanks to the ionization energy loss measurements (dE/dx) it was also used for charged particle identification. It is of fundamental importance for low-momentum tracks, not able to reach the DIRC (see Section 2.3.5), which also covered only the barrel of *BABAR* detector: the DCH was the only device providing particle identification in the forward region. It also had an important role in the L1-trigger. A 280 cm-long cylinder composed of two layers, the inner at a radius of 23.6 cm and the outer at 80.9 cm, contains aluminium field wires and a gas mixture of 80:20 Helium:Isobutane for multiple scattering minimization. It is closed by aluminium end plates, on which the readout electronics is mounted (on the backward side only).

The DCH is optimized to minimize the material budget and the amount of material inside the chamber accounts for only 0.2% radiation lenghts. It is composed of 40 layers of small hexagonal cells gathered into 10 superlayers, where the layers have all the same orientation. The disposition of superlayers follows the ordering AUV, which stands for *A=Axial, U=positive stereo angle, V=negative stereo angle.* The stereo angle (50-70 mrad) refers to the intersection of the direction of wires in subsequent layers, introduced to measure also the z coordinate. The spatial resolution is better than 140 μm and the DCH performed as expected during all the operational stages, with a resolution on dE/dx measurements of 8%. The resolution on the transverse momentum measurement achieved by *BABAR* is approximately 0.3% for 1 GeV/c pions at 90°.

2.3.5 The *Detector of Internally Reflected Cherenkov light*

The Detector of Internally Reflected Cherenkov light (DIRC) provides the main contribution to particle identification (PID) of charged particle, expecially for π/K separation up to high momenta (4 GeV/c), by measuring the emission angle θ_C of the light produced by the Cherenkov effect.

This effect consists in a charged particle passing through a medium at a velocity v_P larger than the speed of light in the medium (c/n), therefore generating a shock front of conical radiation propagating at an angle with respect to the direction of flight of the particle defined by $cos\theta_C = \frac{1}{\beta n}$. The medium exploited by the DIRC consists of 144 synthetic quartz bars of 4.9 m lenght $\times$ 6 cm^2 transverse area, with a refractive index $n = 1.474$, organized in a 12-sided polygonal barrel. In Figure 2.11 a schematic view of the longitudinal section of the DIRC is shown.

Only the backward bar end is instrumented, while in the forward region the bar end has a mirror which reflects the emitted photons towards the backward side. Once they reach the quartz edge, they are reflected at larger angles and they exit the quartz, entering the so-called Standoff Box (SOB) which is a region filled with purified water needed for the Cherenkov cone expansion. The SOB drives the photons beyond the end of the magnet, at a distance of 1.2 m far from the bar end, towards the wall equipped with almost 11000 Photomultiplier Tubes (PMTs). The PMTs detect light in the visible and near-UV spectrum and they need a dedicated shielding to prevent damages from magnetic fringe field.

Both the positions of the detected photons and their arrival times are used to reconstruct the Cherenkov angle θ_C. For every detected photon, there are several ambiguities coming from the ignorance in the number and of the nature of the reflections undergone by the radiation propagating in the quartz. They are used as input in reconstruction algorithms based on unbinned maximum likelihood fits, which produce as results likelihood values for each of the five stable charged particle hypothesis (e, μ, K, p), the Cherenkov angle estimate and the number of signal and background photons.

The typical resolutions achieved on the measured angle are 10 mrad per photon and 2.5 mrad per track, allowing to separate at 4σ level pions from kaons up to 4 GeV/c momenta.

2.3.6 The *Electromagnetic Calorimeter* and the muon detector

The *BABAR* Electromagnetic Calorimeter (EMC) consists of a barrel of 48 rings of 120 CsI(Tl) crystals each and a foward end cap composed of 820 crystals organized in 8 rings, as it is schematically shown in Figure 2.12. These crystals respond to particle excitation by emitting photons according to the mechanism of the scintillation. The incoming particle starts producing the electromagnetic shower into the crystals and it causes the CsI to scintillate. The number of photons is related to the energy loss of all the particles in the shower and consequently to the total ionization energy deposited in the calorimeter.

The precision improves as energy increases and the energy resolution σ_E/E goes from 5% (at 6.13 MeV) to 2% (at 7.5 GeV), while the angular resolution varies from 12 to 3 mrad. The

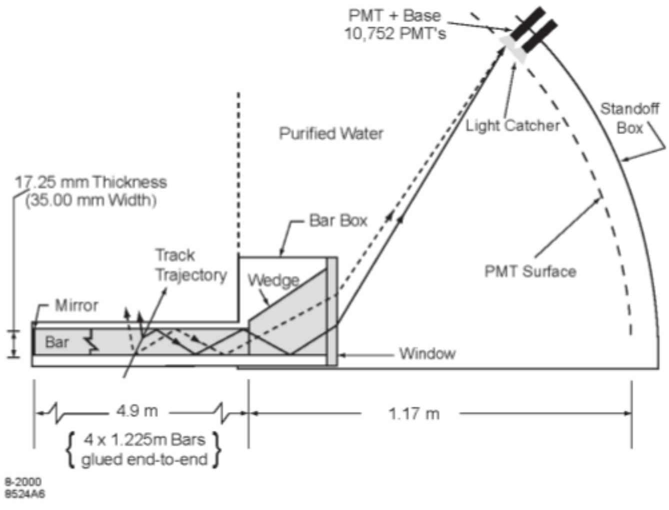


Figure 2.11: Longitudinal view of *BABAR* DIRC, from the forward (left) to the backward (right) region. Figure taken from *The Physics of the B Factories* [4].

EMC is the device which provides the most reliable method for electron identification, using the ratio E/p, where E comes from shower measurements and p from tracking.

The most external sub-system, devoted to muon detection and helping in neutral hadron identification, is the Instrumented Flux Return (IFR). It consists of a segmented iron structure instrumented with Resistive Plate Counters (RPCs). They are more than 800 organized in 19 layers in the sextants of the barrel region, and in 18 layers in the end caps. The barrel structure, the forward and the backward doors are schematically reported in Figure 2.12. The RPCs of the forward door have been completely replaced since they quickly showed a reduction of the muon detection efficiency, with a drop of about 15% and the sextants were also upgraded during *BABAR* operational stages. The average efficiency of the barrel section at the end of the data-taking period was around 88%, just slightly below the 92% geometrical acceptance.

2.3.7 Trigger

The trigger system in *BABAR* consists of two levels: the trigger L1, implemented in hardware, and the trigger L3, which has a software implementation. The combined efficiency of the two stages fully statistfies the trigger requirements providing a 99% efficiency for $B\bar{B}$ events, 95% for continuum decays ($uds, c\bar{c}$) and about 92% efficiency for tau pairs events.

The L1 trigger receives inputs from the DCH charged tracks, the EMC showers and the IFR. The drift chamber trigger (DCT) and the calorimeter trigger (CMT) are globally processed in the global trigger (GLT). They independently provide redundant information and satisfy all the trigger requirements. Originally the DCT provided only $r - \phi$ information: after the improvement for a 3D-tracking, the added z-information allows to reject the background coming from beam-gas interactions, where tracks are produced more than 10 cm away from the IP, resulting in the capacity of following the increasing luminosity with no significant impact on the dead time. The goal of the GLT, receiving primitive information, is to send out a trigger

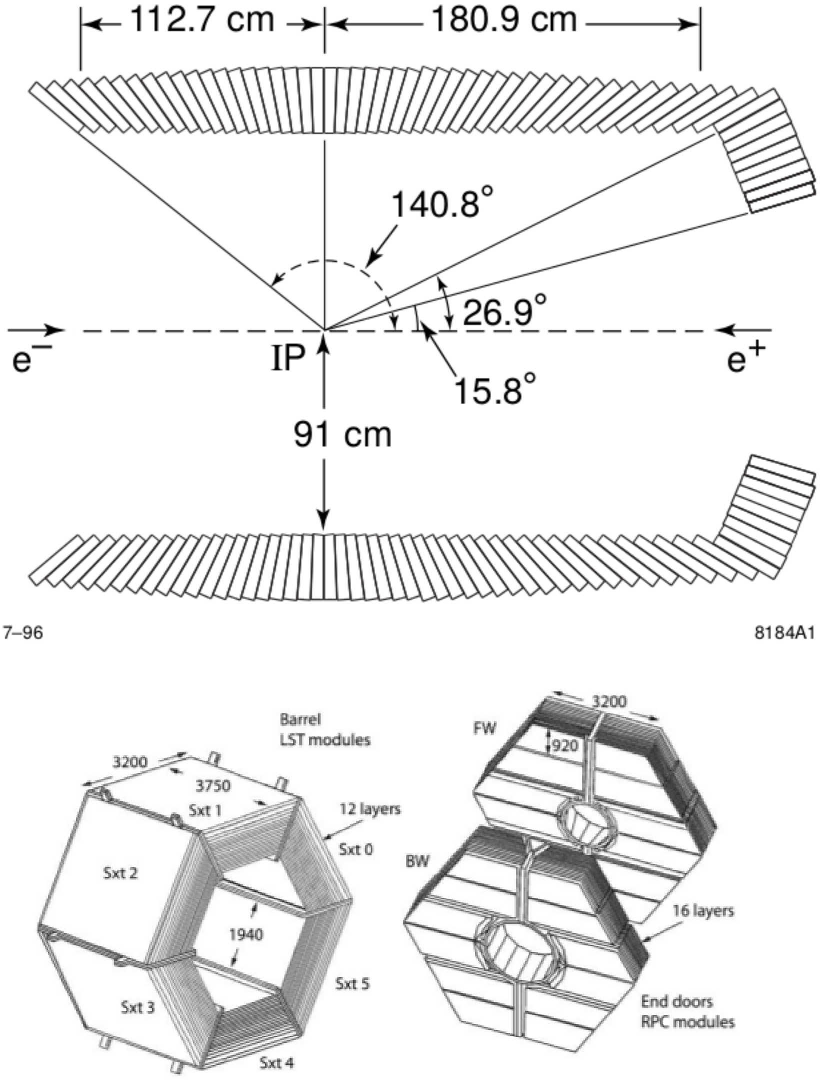


Figure 2.12: Top: the Electromagnetic Calorimeter longitudinal view is shown, with the central barrel and the forward end cap. Bottom: the final version of the Instrumented Flux Return is illustrated. Figures taken from *The physics of the B Factories* [4].

outcome to the *BABAR* DAQ system which accepts the the signal and provides the readout of all the event.

The L3 software trigger refines and corroborates the L1 selection. Dedicated software algorithms select events of interest which are then stored for processing. The input trigger signals from L1 are further reduced and the main physics sources are identified and flagged for the specialized categories of events needed for luminosity evaluation, diagnostic and calibration purposes. Thanks to its spare capacity, the L3 trigger has allowed to accept data at a rate of 800 Hz, to be compared to the intial design value of 120 Hz.

2.3.8 Simulation and reconstruction at *BABAR*

The comparison of data and simulations plays an important role in most analyses and therefore two kinds of events collections are available at the *BABAR* Event Store: the real data collections from Prompt Reconstruction lists (PR) and the simulated events from Monte Carlo official production.

The real data coming from PR lists are the output of a reconstruction process which starts with raw detector signals in *BABAR*. They are digitized and if they passed the trigger, they are reconstructed forming tracks and calorimetric clusters.

The tracks reconstruction relies on the information collected by the DCH and the SVT subdetector, from which track parameters and their errors are extrapolated by the dedicated fit implemented in *BABAR* software, based on the Kalman filter fit technique [16]. The Kalman filter fit is a tracks finding procedure which correctly takes into account effects of multiple scattering and the energy loss in the material. Once the tracks are reconstructed, their momenta are measured from the track curvature and calorimetric clusters are reconstructed from adjacent over-threshold crystals and can be associated to charged tracks or identified as neutral object. Particle identification algorithms (see Section 2.3.9) are then applied to establish particle identities.

At this point, the event reconstruction continues with the combination of stable particles to form a common vertex which indicates composite particle decay. The whole decay chain is reconstructed and the event is stored.

Monte Carlo simulation uses event generators, consisting of several packages in *BABAR* framework that, according to theoretical models of particle interactions, produce sets of final state four-vectors describing e^+e^- collisions. A particle identity is associated by the generators to each four-vector. A combination of both background and signal generators is required to properly mimic the event stream of real data. The production of hadronic event from e^+e^- collisions is handled by two generators:

- the `EvtGen` package[17], which is used for Υ resonances simulation and for B decays to exclusive final state, as the one analysed in this thesis;
- the `JetSet` generator [18], also known as `Pythia`, which simulates the production of quark-antiquark pairs from e^+e^- collision; it is used for generic continuum and inclusive B decay simulation.

In this thesis, samples simulating lepton pair events have been also analysed and they are simulated with KK2F [19].

Background event generators also include Bhabha scattering, Bremsstrahlung, QED and machine background, intial state radiation and cosmic rays.

After the generation, the simulation continue with particle propagation and transport processes which end up with the interaction of the particle with the *BABAR* detector, supplied by the GEANT4 simulation toolkit[20]. Electronic signals which simulate the detector response are computed and they are finally reconstructed in the same way as signals from real data.

The *BABAR* software used for performing analyses is called `Beta`, a framework written in C and C^{++} which provides the basic interface to the complete reconstruction of *BABAR* data. The software release exploited in this thesis is version 24.5.6. The `Beta` framework consists of several packages, *e.g.*, providing the result of the fit to the decay vertex, information about the quality of the reconstructed tracks and the particle identification.

2.3.9 Particle Identification at *BABAR*

Particle Identification, usually referred to as PID, is the process which analyses information from the various subdetectors in order to identify the nature of the charged particle passing thourgh the detector at a given time. At the energy of the *BABAR* experiment, only 5 types of stable or long-lived charged particles can interact with the detector: electrons, muons, pions, kaons and protons. For each particle, the momentum $p = mc\beta\gamma$ is given by the track curvature and if its velocity $v = \beta c$ is measured, its mass and, therefore, the particle identity can be determined.

Depending on the particle momentum, the velocity can be measured through the reconstruction of the Cherenkov emission angle in the DIRC or through the measurement of the specific energy loss dE/dx in the SVD and in the DCH. This method is applied mainly for $\pi/K/p$ identification. Electron identification relies on matching the measured momentum to the total energy released in the EMC, while muon identification is based on the penetration depth in the IFR detector. In Table 2.4 the list of all the PID information provided by the various components of the *BABAR* detector is given. As shown in the table, for proton identification the needed information is provided by the tracking system (SVT, DCH) and by the DIRC.

Finally, dedicated algorithms, the PID selectors, receive as input all this information and provide as output the particle identity of the reconstructed track. They are available in different levels of tightness and they are characterized by two numbers: their efficiency for identifying a particle type α and the mis-identification probabilities for incorrectly assigning particles of type $\beta, \gamma, ..$ to the paticle type α. These numbers vary as a function of the momentum and the polar angle of the track.

Different PID selectors have been implemented in the history of *BABAR* analysis software. They can be divided into two main categories: the older generation Likelihood selectors, more stable during the different data-taking periods, and the PID selectors relying on statistical tools of multivariate analysis (Bagged Decision Trees, Neural Networks), more performing but also more sensitive to the detector operational status. It is possible to use dedicated *lists*, available

Detector subsystem	PID information	Particle Type
SVT	$dE/dx, \vec{p}$	K, π, p
DCH	dE/dx	K, π, p
DIRC	$cos\theta_C$	K, π, p
EMC	E/p	e
IFR	IFR count	μ

Table 2.4: The list of the PID information provided by the different subdetectors, according to the type of particle to detect, is reported in the above table.

for each particle type, formed by a combination of both Likelihood and Decision Tree selectors, sharing the same tightness level. This enable to choose afterwards, once the reconstruction is performed, the selector that gives the best sensitivity for the specific analysis. The PID list I used in the reconstruction phase is a combination of *very loose* selectors for protons. Further optimization on PID choice is achieved thanks to the Decision Tree based selector *pKM*, as reported in Section 5.1.

Decision Trees in particle identification

An important category of PID selectors implemented in *BABAR* framework is based on multivariate analysis methods. The KM selectors rely on the Decision Tree technique and they are used for π, K, p and e identification. They can be applied at different levels of particle efficiency, from *super loose* to *very tight*.

To identify particle types, a decision tree is trained on a set of discriminating variables, based on the information collected from various subdetectors. For the training, input data are splitted into rectangular nodes, considering all the possible splits in each dimension (discriminating variable). The goal is finding the split which provides the highest signal to background ratio.

The discriminating variables are chosen according to the type of particle to identify. Typical information used is the particle momentum, the polar angle and charge, the number of hits in the DCH, the number of DIRC photons and, for electrons, also the EMC shower energy and depth.

The decision tree based proton selectors are indicated as *pKM* selectors among *BABAR* analysis tools. Performance studies have shown that they provide the optimal PID in analysis dealing with protons.

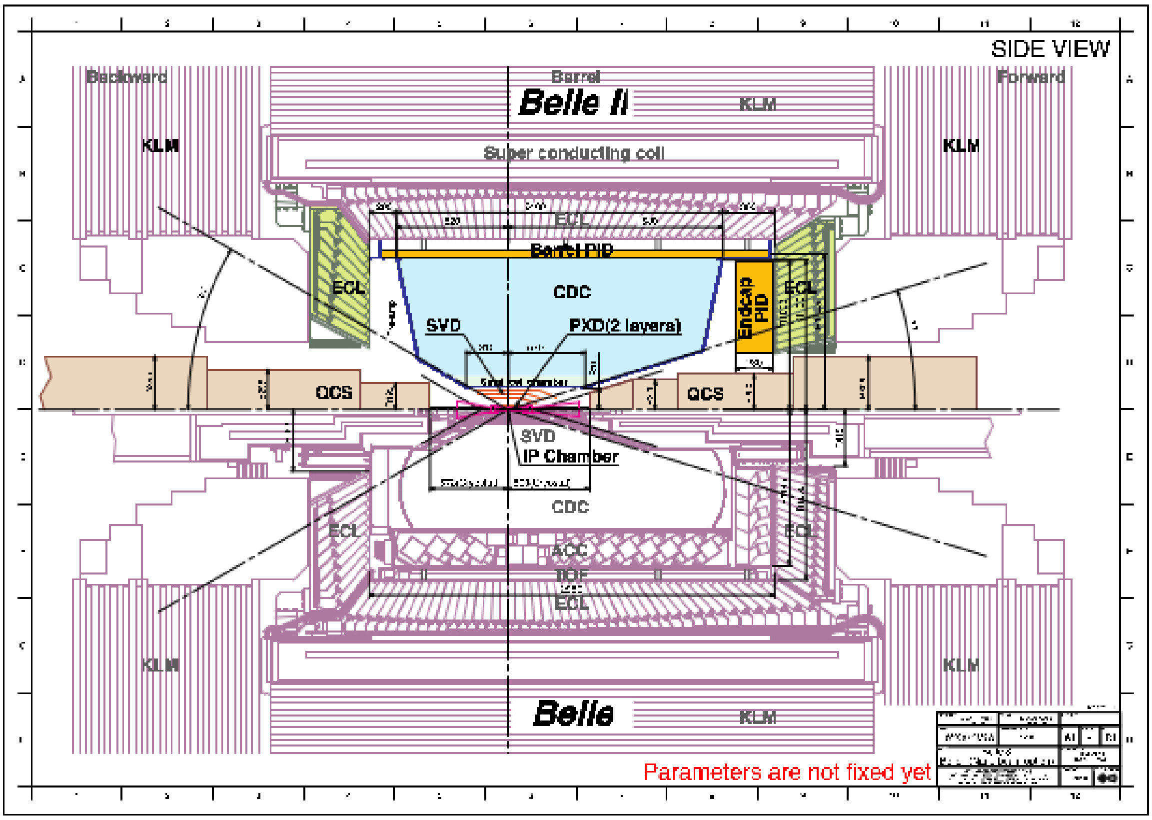


Figure 2.13: The top half shows the upgraded Belle II spectrometer as designed in the TDR[13]; in the bottom half the longitudinal view of the old Belle detector is shown.

2.4 The Belle II detector

The Belle II detector is described in detail in the Technical Design Report [13]. As done for the *BABAR* detector, a general overview is given. The main changes and improvements with respect to the Belle detector will be underlined in the following Section.

2.4.1 Belle II upgrade overview

The aim of the Belle II upgrade is to mantain the same good performance of Belle detector in the new experimental environment with much higher background levels. A schematic comparison of the longitudinal view of the new Belle II spectrometer with respect to the old Belle detector is shown in Figure 2.13. Since the SuperKEKB boost is reduced with respect to KEKB (0.28 instead of 0.43) achieving the same proper time resolution requires better vertex resolution, that can be achieved with the reduction in radius of the first active detector layer. Studies estimate a conservative 20 times increase of the background hit rate, to be compared to a 50 times increase of the event rate. The modifications studied to ensure a comparable or even better performance for the Belle II detector are the following:

- the inner part of the vertex silicon strips detector, next to the beam pipe, is replaced by a two-layer silicon pixel detector based on DEPFET technolgy, placed at smaller radius and with a very light material budget;

- the silicon strips detector starts from just outisde the pixel detector and reaches a larger radius than in Belle;
- the silicon strips readout is based on APV25 chips which have a shorter shaping time with respect to Belle VA1TA chips to cope with higher background;
- the central drift chamber, the main tracking device, extends to a larger radius and it has smaller drift cells;
- the new particle identification devices, both in the barrel and in the end caps, are more compact and highly performant, exploiting the Cherenkov image detection method and a very fast readout electronics;
- the electronics of the electromagnetic calorimeter is replaced with a fast wave-form-sampling circuit;
- the barrel part of the muon and K_L detector is modified with additional inner RPCs maybe some of them operating in the proportional mode, while the end caps are replaced by scintillators and PMTs.
- a new data aquisition system for handling the higher event rate is provided.

The improvements foreseen for Belle II performance are various, first of all the vertex resolution which benefits from the two innermost pixel layers; second, a good improvement of the reconstruction efficiency for the K_S decaying to charged pions, which appears in the *golden mode* exploited in CP asymmetries measurements. In addition, the new particle identification devices ensure a better pion/kaon separation in the whole kinematics range of the experiment. Finally, a better performing electronics for the calorimeter readout reduces the pile up, improving missing energy measurements.

2.4.2 The *Pixel Detector*

The expected background rate at small radius (~ 20 mm) makes the use of silicon strip sensors not feasible because of the high occupancy. Pixel detectors have a much finer segmentation and therefore smaller occupancy (3%). Strips detectors at SuperKEKB can be safely located at larger radii.

The pixel detector (PXD) concept implemented at Belle II is based on the DEPFET (DEPleted Field Effect Transistor) technology. It allows to have very thin pixel (50 μm) with readout electronics moved outside the acceptance region. This implies that the active cooling needed for readout operation is not contributing to the multiple-scattering material budget. For the sensors, air cooling is sufficient since they are not consuming too much power. Also radiation tolerance can be achieved engineering the sensors in a dedicated radiation hard-technology.

DEPFET is a semiconductor detector concept able to combine detection and amplification in one device. A DEPFET pixel consists in a p-channel MOSFET structure integrated onto a fully depleted silicon substrate. The MOSFET has an internal gate where the electrons generated

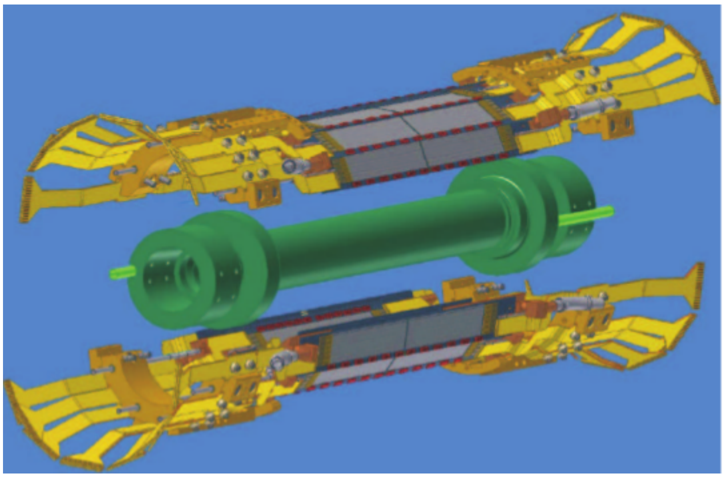


Figure 2.14: The two-layer pixel detector is shown, with the supporting structure and in place on the beampipe. Picture taken from the Belle II TDR [13].

when a charged particle passes through the device are collected. The internal gate modulates the current at readout time allowing the detection of the collected charge.

The DEPFET sensors are organized into planar modules, the ladders, which compose the two layers of the PXD, with radii respectively of 14 and 22 mm. The sensitive lenghts of each layer are determined by the acceptance requirements of the detector: a polar angular coverage from 17° to 155° is ensured. The total amount of pixels composing the PXD is around 8 milion organized into arrays. Each DEPFET pixel represents a cell, a monolithic structure with internal amplification, therefore much thinner than other devices which need additional amplification. Typical pixel sizes have to meet the vertex resolution requirement, which is 20 μm, and also the thinning feasibility, resulting in a $50{\times}50$ μm^2 and $50{\times}75$ μm^2 cells respectively for the inner and outer layers.

Another challenging requirement is the short frame readout time, that is achieved by splitting the readout on both sides of the pixel matrix and introducing high level of parallelization (a four rows simultaneous readout): the processing time for each pixel row is then manageable, of order of 100 ns, and all the 1600 rows can be read in a total frame time of 20 μs.

The sensors are finally mounted on the supporting structure which is able to slide on the beam pipe to accomodate for thermal expansion. A schematic view of the total structure of PXD is reported in Figure 2.14

2.4.3 The *Silicon Vertex Detector*

The physics of the B Factories deals with low-momentum particles and the main issue for Belle II tracking is the multiple scattering. The fundamental requirement for the Silicon Vertex Detector (SVD) is keeping the material budget low. A short shaping time is also necessary to limit the occupancy under few percent. These features determine the geometrical shape of the sensors and also the readout electronics adopted. The former issue is solved by adopting the largest possible silicon wafers and equipping the forward region with trapezoidal slanted sensors. For

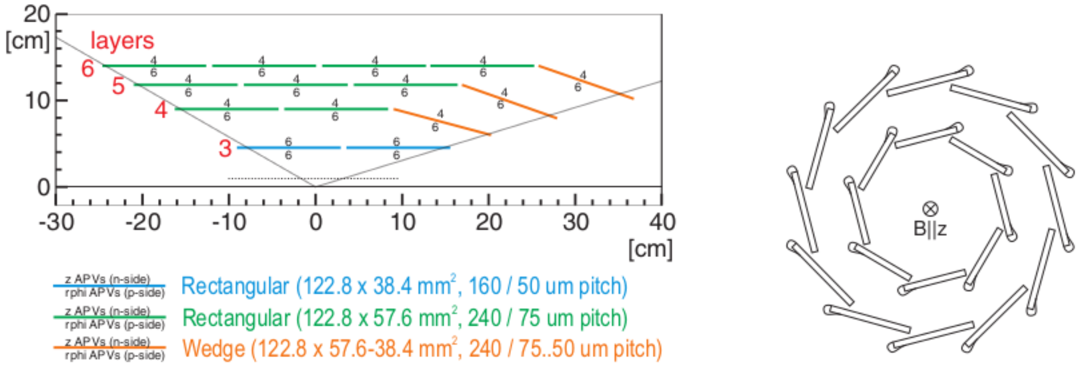


Figure 2.15: Left: Schematic view of the SVD longitudinal section, showing the sensors geometry for each layer and the number of APV25 chips on each side of the sensors. Right: the windmill structure of the inner sensors, with the Origami concept applied on the upper edge of silicon modules, is shown. The solenoidal $\vec{B}$ field parallel to the z-axis is also indicated in the picture.

the latter, the solution came from inheriting the APV25 chips first used in the tracker of the CMS experiment (LHC, Cern). The APV25 chips have 128 channels each and a shaping time of 50 ns. It has a sampling mode of operation in which, performing a fit on the sampled signal, the peak time can be measured with a resolution 3 ns. This results in a innermost layer occupancy smaller than 1%.

The SVD is composed of four layers of Double-sided Silicon Strip Detectors (DSSDs), located respectively at radii of 38 mm (*layer 3*), 80 mm (*layer 4*), 104 mm (*layer 5*) and 135 mm (*layer 6*). The azimuthal angular coverage goes from 17° (the *forward* region) to 150° (*backward* region). The asymmetry is due to the Lorentz boost in the laboratory frame, which favors events boosted in the forward direction.

There are three different types of sensors: the rectangular ones are produced by Hamamatsu Photonics (HPK), while the Micron Semiconductur (UK) produces the trapezoidal ones. Sensors are produced starting from six-inches 320 μm-thick silicon wafers. The rectangular sensors are developed in two sizes: for layer 3, which has only the barrel section, smaller sensors are adopted, while for layers 4 to 6 larger sensors are used for the barrel region. The forward region of layers 4 to 6 is convered by trapezoidal slanted sensors which minimize the material budget, simultaneously optimizing the acceptance. For the rectangular sensors, the long strips measuring the $r\phi$-coordinate on th p-side are in total 768, readout by 6 APV25 chips per sensor. They are parallel to and facing the beam axis. The short strips on the n-side, which measure the z-coordinate, are 512, therefore they need only 4 APV25 chips for readout. They are located on the sensor side towards the outside. For the slanted sensor of the forward region the geometry is the same, except for the p-side strips which are not parallel to z-axis anymore. Dimensions are given in mm in Figure 3.1.

The *Origami chip-on-sensor* concept is a readout modification introduced in Belle II for the inner sensors, while the forward and backward sensors are readout by conventional APV25 chips located outside the acceptance. The Origami concept refers to bent *fanout* pieces that can only be applied to the outer edge of the sensors in the windmill structure, as shown in Figure 3.1.

The strips on the outer side are read by a planar *fanout*, while for the strips on the inner side, a flexible *fanout* bent around the upper edge of the sensor is needed. This technique aligns all the chips on the outer side, enabling to use only a single cooling system for all the ladder. The cooling is provided by a dual phase CO_2 system, whose design aims at minimizing the material budget, adopting 1.4 mm-large tubes with 0.05 mm-thick walls.

2.4.4 The *Central Drift Chamber*

The Central Drift Chamber (CDC) fulfils three main tasks. It contributes with the SVD and PXD to charged track reconstruction and it measures the momentum. Second, it helps in particle identification thanks to the measurements of the energy loss due to charge particle ionization in its gas volume. Finally, it provides reliable trigger signals for charged particle events.

The CDC follows the successful design of its predecessor in Belle: the same structure is implemented and the same material and gas mixture are used. The main modifications regard:

- the readout electronics, which needs to be much faster in order to handle higher trigger rates;

- the cylinder radii, where the innermost ones has to be larger to account for a bigger SVD, while the outermost radius has to deal with the more compact barrel detector for particle identification, resulting in a 160 mm and 1130 mm radii respectively. The resulting acceptance in azimuthal angle goes from 17° to 150°. Also the wire configuration has to be adapted.

- the three-dimensional trigger information, providing a z trigger (see Section 2.4.8) which efficiently helps in reducing beam backgrounds without sacrifying physics events.

The CDC consists in 56 *layers* organized in 9 *super-layers*. The length of the squared cells varies from 10 mm for the innermost super-layer to 18.2 mm for the outermost ones. The 50% helium-50% ethane gas mixture is filled with a total of 14336 tungsten sense wires of 30 μm diameter. For the electric field generation, aluminium wires of 126 μm-diameter are used. The structure is supported by two carbon-fiber reinforced cylinders, terminating with aluminum endplates.

The measured spatial resolution of the CDC is about 100 μm. The resolution on the dE/dx measurements for particles with an incident angle of 90 ° is around 12%.

2.4.5 Particle Identification devices

The particle identification in Belle II is performed by two different types of detector: the *Time-Of-Propagation* counter in the barrel region and the *Aerogel Ring-Imaging Cherenkov counter* for the end-cap. Both the sub-detectors exploit the Cherenkov effect to identify the charged particle traversing the spectrometer.

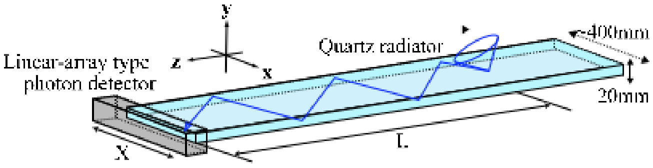


Figure 2.16: Top: schematic view of the Time-Of-Propagation counter. Both the arriving time and position of the internally reflected Cherenkov photons are measured.

The *Time-Of-Propagation* counter

The Time-Of-Propagation (TOP) counter, shown in Figure 2.16, measures the time of propagation of the Cherenkov photons internally reflected in the quartz bar. Together with the detected arriving position (x, y coordinates in the picture), this information allows to reconstruct the 3-D image of the Cherenkov cone and its opening angle θ_C. From the Cherenkov angle the particle velocity v can be inferred and the likelihood for different mass assignment hypothesis is calculated.

The TOP consists in 16 modules surrounding the outer wall of the CDC. Each module is a quartz radiator composed of two quartz bars glued together for a total lenght of 2500 mm and a transverse area of 44×20 mm^2. Photons emitted in the forward direction are reflected backward by a mirror placed at the forward end of the quartz bar. Photons propagating in the backward direction are collected by different PMT channels, depending on their reflection angles. The critical issue in the TOP performance for PID is the broadening of time resolution due to the photons chromaticity, which is cured by a focusing system consisting in a slightly concave-shaped mirror, aimed at compensating the chromaticity dispersion of Cherenkov photons: the parallel rays are focused into the same pixel of a photo-sensor, while the chromatically-dispersed one are detected by separate channels.

The required time resolution for a single-photon detection is around 50 ps. The goal is achieved by employing as photon detectors the Multi-Channel Plate Photo Multipliers (MCP PMTs), two-stage amplification devices, with a gain of 10^6 and a very fast reponse.

The *Aerogel Ring-Imaging Cherenkov* counter

In the forward end-cap, the particle identification is provided by the proximity focusing Aereogel Ring-Imaging Cherenkov counter (ARICH), whose schematic view is shown in Figure 2.17. It consists of an aerogel radiator producing Cherenkov photons when a charged particle passes thorugh its volume; an expansion volume of 20 cm divides the aerogel tile from the photon detector surface and allows the Cherenkov photons to enlarge into rings. The array of position sensitive photon detectors provides single photon detection in a high magnetic field (1.5 T) with optimal resolution in two dimensions.

The main parameters describing the detector are the number of detected photons $N_{p.e.}$ and the resolution on the Cherenkov angle σ_θ. Increasing the radiator thickness, the number of detected photons increases and the resolution per charged track improves as $\sigma_\theta/\sqrt{N}$, but the

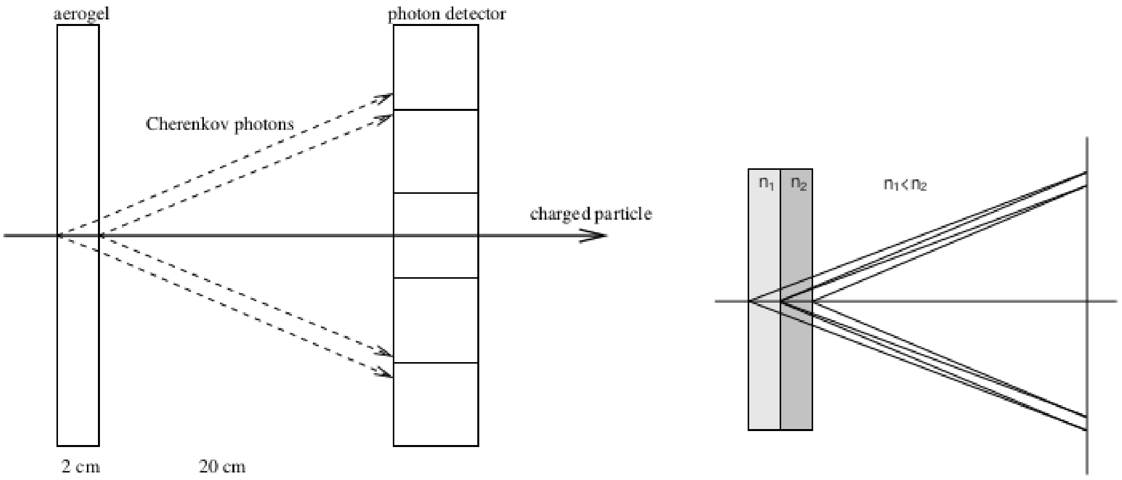


Figure 2.17: Left: a schematic drawing of the ARICH is shown. Dimensions are given in cm and from left to right the arrangement of the aerogel radiator, the expansion volume and the position sensitive phton detector surface is illustrated. Photons trajectory are reported in dashed lines. Right: the proximity focusing system with non-homogenous dual-layer aerogel is shown. The two rings coming from the different aerogel layers, with an appropriate choice of the refractive indeces, totally overlap on the photon detector surface.

angular resolution on the single photon (σ_θ) degrades due to the uncertainty of the emission point. Combining two layers with appropriately tuned refractive indices in order to have two overlapping rings will reduce the spread due to the emission point uncertainty, achieving a resolution on the Cherenkov angle $\sigma_\theta \simeq 13$ mrad. The performance is optimized for charged tracks with momentum larger than 3.5 GeV/c, but no significative degradation is shown also for lower momentum tracks.

A non-homogenous aerogel radiator (see Figure 2.17) is provided and the proximity focusing technique is applied by choosing as upstream layer the aerogel with lower refractive index ($n_1 = 1.046$) and increasing the refractive index of the downstream layer ($n_2 = 1.056$). Studies have shown that a twice as thick layer of homogenous aerogel radiator would provide the same number of photons, but with a worse resolution $\sigma_\theta \simeq 20$ mrad.

The development of reliable photo sensors with good efficiency for single-photon detection in the high magnetic field is achieved with the *Hybrid Avalanche Photo-Detectors* (HAPD) by Hamamatsu Photonics.

The ARICH global acceptance covers the azimuthal angular region from 15° to 30°. Considering the studies on Cherenkov angle resolution, for a total measurement of 20 Cherenkov photons (10 per ring) the precision achieved is:

$$\sigma_{track} = \sigma_\theta/\sqrt{N} \sim 3 \text{ mrad.}$$

2.4.6 The *Electromagnetic Calorimeter*

The electromagnetic calorimeter (ECL) in Belle II covers many important tasks: it has to provide a high energy resolution, since more than a third of the B meson decay products consist of π^0 and neutral particles. The calorimeter needs to detect with high efficiency photons over a wide energy range, from 20 MeV to 4 GeV, and it has to precisely determine their energies and angular coordinates. It must provide electron identification, luminosity measurements and proper trigger signal generation; finally, the calorimeter and KLM combined performance ensures kaon detection.

The choice for the Belle II detector has been to exploit the Belle calorimeter, replacing the readout electronics in order to cope with the increased background rates. The calorimeter is composed of a barrel section of 3 m length and 1.25 m inner radius, with an azimuthal angular coverage of $12.4° < \theta < 155.1°$. The total structure contains 8736 CsI(T) crystals shaped as truncated pyramid with a 6×6 cm^2 cross section and of 30 cm length, corresponding to 16.1 radiation lengths (X_0). The size of the crystals enables to collect the 80% of photon energy deposit, limiting the resolution degradation at high energy rate caused by the shower energy loss fluctuations. The scintillation light is detected by two $10{\times}20$ mm^2 Hamamatsu Photonics S2744-08 photodiodes glued onto the crystal with a 1-mm plexiglass plate.

The structure of the barrel section consists of 8 cells separated by 0.5 mm thick aluminium septum walls, closed by two cylinders. A water cooling system handles the heat generated by the two preamplifiers. The structure for the end caps is quite similar to that of the barrel section.

From calibration with cosmic muons it is known that the average output signal of the current crystals gives 5000 photoelectrons per 1 MeV of deposit energy, with an average noise level of 200 keV. The higher background rate is handled with a new readout electronic based on the signal sampling method.

The instrinsic energy resolution of the calorimeter from a prototype test is given by the approximated formula:

$$\frac{\sigma_E}{E} = \sqrt{\left(\frac{0.066\%}{E}\right)^2 + \left(\frac{0.81\%}{\sqrt[4]{E}}\right)^2 + 1.34\%^2} \tag{2.8}$$

E is the energy in GeV and the first term refers to the electronic noise contribution.

2.4.7 The neutral kaon and muon detector

The K_L and muon detector (KLM) is located outside the superconducting solenoid which provides the 1.5 T magnetic field used for momentum measurements and it consists of alternating 4.7-cm thick iron plates and active material detectors. A schematic overview is shown in Figure 2.18. The iron is needed as magnetic flux return for the solenoid and it also provides the material budget for 3.9 interaction lengths, allowing the K_L to hadronically shower in its volume. The octagonal barrel section covers the azimuthal region $45° < \theta < 125°$. Adding the endcaps coverage, the KLM reaches the angular acceptance $20° < \theta < 155°$. The barrel is equipped with glass electrode Resistive Plate Chambers (RPCs), while in the end caps scintillators readout by SiPM are installed. This replacement is needed to avoid the RCPs efficiency drop in a high rate environment.

The use of scintillator strips equipped with two SiPMs, one at each end, ensures an optimal

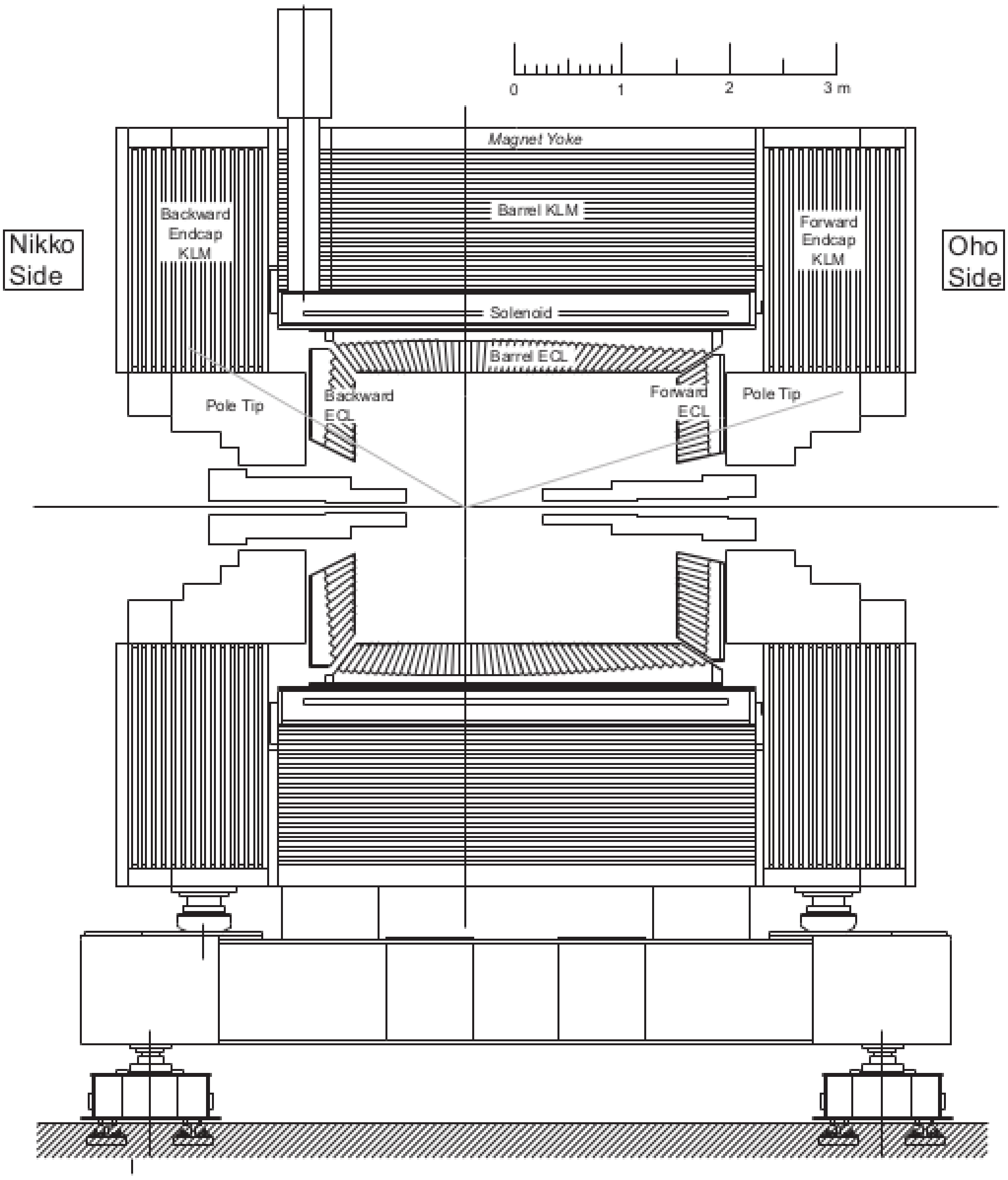


Figure 2.18: A schematic overview of the neutral kaon and muon detector at Belle II is given in the picture.

time resolution, which is measured with a trigger provided by cosmic ray events; the time dispersion difference gives $\sigma_t \sim 0.7$ ns. The good timing performance enables to exploit the KLM endcaps as time of flight detector for K_L. The barrel section of the KLM provides a μ-reconstruction efficiency of 89% for 1 GeV momentum tracks and a plateau for K_L-reconstruction efficiency around 80% for particles with momentum higher than 3 GeV.

2.4.8 Trigger

The high luminosity upgrade will increase the total event rate to almost 20 kHz, requiring a robust and at the same time flexible trigger for discriminating the signal events from the background rate, which are expected 20-30 times larger than at KEKB.

Considering the high charged track multiplicity in $B\bar{B}$ event final states, an obvious trigger requirement is on the number of tracks in the event. However, there are physics processes, *e.g.*, τ leptonic decays and dark matter search, that have zero or two tracks only in their final state, which causes much more difficulties in signal and background discrimination.

In addition, some low multiplicity processes and Bhabha scattering, which has a huge cross section at $\Upsilon(4S)$ CM energy, show the same topology and when vetoing Bhabha events for reducing the total event rate, large part of the trigger efficiency on low multiplicity events is lost. In order to preserve high efficiencies on the processes of interest, an effective trigger system is essential.

Its implementation will follow the same principle successfully adopted for Belle, but with totally new technologies able to support the increased event rates. The trigger system is composed of the Level 1 trigger (L1), implemented in hardware, and the software implementation of the High Level Trigger (HLT).

The L1 Trigger

The L1 trigger system gathers many sub-trigger systems which provide information from various subdetectors. A first refinement is provided by the global reconstruction logic (GRL), whose output is sent to the global decision logic (GDL), responsible for the issuing of the trigger signal. In Belle II, all the trigger system components have been replaced by new technologies, based on the Field Programmable Gate Array (FPGA), which makes the logic configurable rather than hard-wired. The first combination of detector information is performed by the GRL and it is mainly based on the drift chamber trigger and on the ECL trigger.

The CDC trigger finds and characterizes the tracks detected in the drift chamber. The 2D tracking basically follows the same implementation as Belle. The Belle II CDC detection improvement with the 3D tracking, which adds the information on the z-coordinate, allows to reject the most part of the background coming from Touschek intra-bunch scattering. In fact, these scattering events show a large displacement in their z-position with respect to the nominal interaction region. The maximum 5 μs latency of the trigger system requires a very fast z-vertex recosntruction, to be performed in no more than 1 μs, which is achieved thanks to neural networks techniques.

The calorimeter trigger system adopts two complementary trigger schemes: a total energy trigger and an isolated cluster counting technique, the former being senstive to events with high electromagnetic energy deposition, the latter to multi-hadronic physics events with low energy clusters and/or minimum ionizing particles. Moreover, the ECL trigger is able to recognize the back-to-back topology of BhaBha scattering with a high purity, which is fundamental for ensuring a high trigger efficiency of low multiplicity processes.

The High Level Trigger

The HLT trigger is developed to improve the backgrounds suppression. It is composed of two parts: the Level 3 Trigger (L3) and the physics trigger. The main task of the former is to reduce the load of the physics trigger. The event rate before and after the L3 trigger passed from 30 to 15 kHz, using only the information from CDC and ECL triggers.

Events that passed the L3 trigger are then sent to the physics trigger. They undergo the full offline reconstruction, involving all output from the subdetectors except the PXD. The physics trigger is then able to classify the event category ($B\bar{B}$, $e^+e^-\to\tau^+\tau^-$, etc.), providing a further reduction of the event rate to 10 kHz is provided. Only at this point, the HLT processing results containing the event tag and the track parameter estimate are combined with the information from the PXD, which enables to calculate the Region of Interest (ROI). Only the pixels belonging to that region are readout and the selected PXD hits are used to complete the event reconstruction.

Chapter 3

Assembly of the SVD modules

The SVD group in Pisa is involved in various aspects of the Belle II detector construction and in particular in the assembling and electrical testing of the *forward* and *backward* modules of the *layer 4*, *layer 5* and *layer 6* of the Silicon Vertex Detector. We perform our activity in the clean room of the High Technology Laboratories of the INFN, Sezione di Pisa, in the dedicated Clean Room.

The goal is to produce and test a total of nearly 100 modules which, once finished, are shipped to the ladder production sites: *layer 4* at TIFR [1], *layer 5* at HEPHY [2] and *layer 6* at IPMU [3]. Layer 3 which does not involve forward/backward modules is assembled in Melbourne. The final ladder mounting will be performed at KEK, where all the produced ladders are shipped for being assembled into the 4 layers that will be placed onto the supporting structure within the Belle II detector. A tridimensional view of the SVD in place on the beam pipe is given in Figure 3.1.

Besides the analysis on *BABAR* dataset for the measurement of the branching fraction of the rare baryonic mode $B^0(\bar{B^0})\rightarrow pp\bar{p}\bar{p}$, my thesis work also concerns the assembling and testing activity on forward and backward modules with the SVD Pisa group. In Section 3.1 the operating principle of Double-sided Silicon Strip Detectors (DSSD) is explained and an overview of the SVD structure and of forward and backward modules is given (see Section 3.2 and 3.3). The assembling and testing procedure are then discussed in Section 3.4 and 3.5. The quality of the produced sub-assemblies is analysed in Section 3.6.

3.1 Operating principle of DSSD modules

The Vertex Detector in Belle II has to measure precisely track parameters to determine the positions for both primary and secondary vertices in the event. For particles with low transverse momentum, the Silicon Vertex Detector is the only tracking device and the requirement of a low material budget is fundamental to precisely measure the low energy tracks. The Double-Sided Silicon Strip Detector technology provides the optimal sensors to fulfil all these requirements.

---

[1] Tata Institute of Fundamental Research, Mumbai, India

[2] Institute of High Energy Physics, Austrian Academy of Sciences, Vienna

[3] Kavli Institute for the Physics and Mathematics of the Universe, University of Tokyo, Kashiwa, Japan

Figure 3.1: A rendering of the Belle II SVD in place on the beam pipe is shown in the picture.

They allow a high precision measurement of bi-dimensional positions. The operating principle of such sensors is the reversely polarized p-n junction.

The physical phenomenon exploited for particle detection is the ionization energy loss in a medium at the passage of a charged particle. In this case, the medium is the silicon and the energy released by a charged particle at its minimum point of ionization (MIP) is usually around 2 MeV cm^2/g. The mean energy loss per unit lenght in Silicon is then $\left(\frac{dE}{dx}\right)^{mean} = 388\ \mathrm{eV}/\mu\mathrm{m}$. In a semiconductor this energy is converted into electron-hole (e-h) pairs. The number of pairs (N_{e-h}) created by a charged partcle in the medium depends on the silicon thickness (d_{Si}) and on the mean energy for producing a pair e-h in silicon (E_{e-h}). Assuming as typical size of a DSSD sensor $d_{Si} = 300\ \mu\mathrm{m}$ and given that $E_{e-h} = 3.63$ eV, the calculation for the numbers of e-h pairs gives:

$$N_{e-h} = \frac{\left(\frac{dE}{dx}\right)^{mean} \cdot d_{Si}}{E_{e-h}} = 3.2 \times 10^4 \tag{3.1}$$

This number has to be compared to the number of intrinsic e-h pairs in the silicon at temperature T=300 K. Assuming an intrinsic carrier density $n_i \sim 1.45 \times 10^{10}$ and a typical surface area $A_{Si} = 100\ \mu\mathrm{m} \times 10\ \mathrm{cm} = 0.1\ \mathrm{cm}^2$, the number of intrisic e-h pairs is:

$$N^i_{e-h} = n_i \cdot A_{Si} \cdot d_{Si} = 1.45 \cdot 10^{10} \cdot 0.1 \cdot 2.5 \cdot 10^{-2} = 3.63 \times 10^7 \tag{3.2}$$

The number of intrinsic e-h pairs is about three orders of magnitude larger than the number of e-h pairs produced by the passage of a MIP in silicon. The technique developed to reduce the number of intrinsic carriers is the reverse polarization of a p-n junction.

A junction is the contact region between differently doped silicon substrates. Its fundamental feature is the space charge region, also called the depletion region: the e-h pairs recombine and the induced electric field prevent instrinsic carriers from drifting in the depletion area, which is then free of charges. In a DSSD sensor an asymmetric p-n junction is created by depositing a surface layer of highly p-doped silicon on a n-substrate. When a negative voltage V_{bias} is applied to the p contact, the junction becomes reversely polarized and the space charge region increases.

At sufficiently high voltage, the n-substrate is totally depleted of intrinsic charge carriers. The reverse polarization induces an electric field which is responsible for the drift of the e-h pairs towards the electrodes (typically in aluminium) on the silicon surface.

The response of the sensor to the passage of a MIP is a current signal induced on the electrodes by the drift of the e-h pairs. In the case the electrodes on the silicon surfaces are segmented into strips, most part of the induced signal is collected by the strip closest to the traversing point of the charged particle. The sensitive strips in the electrodes are electrically isolated thanks to a silicon oxide layer which creates a decoupling capacity and avoids the passage of direct current.

The electrode segmentation on one side of the silicon surface gives a silicon strips detector able to measure the one-dimensional position of a charged particle traversing the sensor. When also the electrode on the opposite surface is segmented in a orthogonal direction, the sensor is called *double-sided silicon strip detector* (DSSD). The two-dimensional postition is provided using the same amount of material budget of an ordinary strip sensor. The electrode segmentation on the n-side of the silicon substrate results in technical problems due to the electron conductive channel created between the silicon oxide layer and the substrate surface, which might short the strips. This is solved by introducing *p-stop* specific doping, which insulate each strip from the other.

The high bias voltage is supplied to each strip through a high resistivity integrated on the sensor. The typical size of a strip is around tens of μm-width and the typical distance between two consecutive strips, called *pitch*, varies from 25 to 100 μm. A schematic view of the operating principle of DSSD sensors is shown in Figure 3.2.

The position of the traversing particle is given by the position of the strip on which the signal from the e-h pairs drift is induced. The spatial resolution is then given by the pitch size, according to the formula $\sigma_{dig} = d_{pitch}/\sqrt{12}$, which is known as digital resolution. When more than one strip collects the induced signal, the position is given by the centre of gravity $\bar{x}$ of the hit strips, calculated as follows:

$$\bar{x} = \frac{\sum_i x_i \cdot S_i}{\sum_i S_i}, \tag{3.3}$$

where x_i stands for the position of the strip on which a signal of intensity S_i is induced. In this case, the spatial resolution is determined by the signal to noise ratio, which might be better than the digital resolution determined by the size of the pitch.

3.2 The SVD ladder structure

The general overview of the SVD has already been presented in Section 2.4.3. It has a coverage in the azimuthal angle from 17° to 150° and it is composed of 4 layers equipped with DSSD modules. Apart from the innermost one (layer 3), layers 4-6 has a slanted geometry in the forward region, in order to optimize the acceptance and to minimize the traversing path length of the incident particle. Each layer is composed of ladders, which consists of a varying number of DSSD modules, depending on the layer, glued onto sopporting carbon fiber ribs.

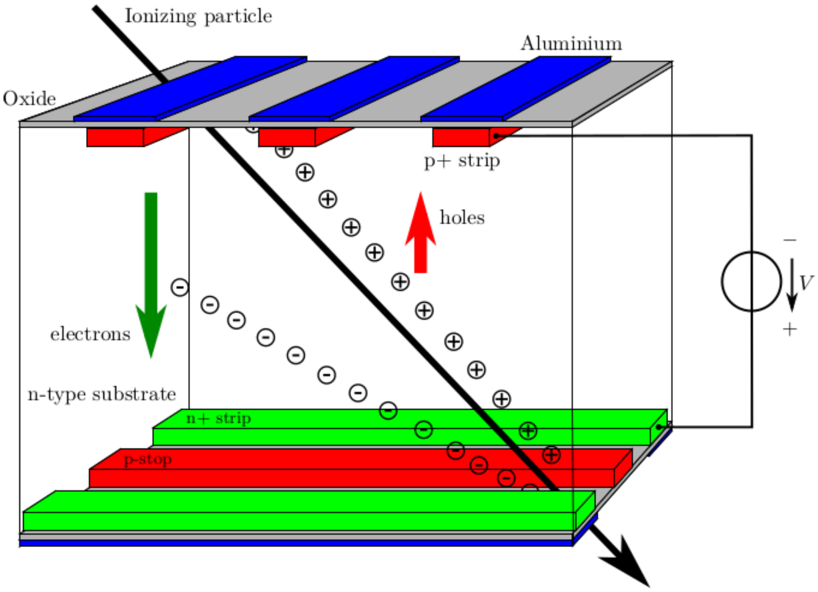


Figure 3.2: The working principle of a DSSD is illustrated. The passage of a charged particle produces e-h pairs in the silicon substrate. A reverse bias voltage is supply between p-strips and n-strips. The e-h pairs drift according to the electric field induced by the bias voltage and segmented electrodes collect the current signal induced by the drifting.

In a ladder, the forward and backward sensors are equipped with flexible circuits which provide the connection to the readout electronic, consisting in APV25 chips mounted on the *Hybrid Sandwich* board. The connection to the readout electronics for the central sensors is provided by flexible circuits bent onto the outer edge of the sensor, according to the Origami chip-on-sensor method (see Section 2.4.3). This technique allow the alignment of the APV25 chips on one single edge of the ladder, favoring the cooling operation which is provided by a single pipe based on a CO_2 cooling system. This method, combined with the chip thinning down to 100 μm, help reduce the material budget in the SVD construction: the total amount of material corresponds to $0.3\% X_0$[4]. A 3D-view of the typical ladder structure (layer 5) is reported in Figure 3.3.

Since the structure of forward and backward DSSD modules is the same for layer 4-6, they are all produced at the INFN Laboratories in Pisa and then they are shipped first to the dedicated sites for the ladder assembling and then to KEK for the ladder mounting.

The procedures developed by the SVD group in Pisa for the assembling and the electrical testing of these modules are described in the following section.

3.3 The forward and backward sub-assemblies

The backward and forward modules[5] of the SVD consist of a DSSD, of the *Hybrid Sandwich* board on which the readout electronics is mounted and of the flexible circuit called *Pitch Adapter*

---

[4] Radiation length.

[5] Referred to also as sub-assemblies.

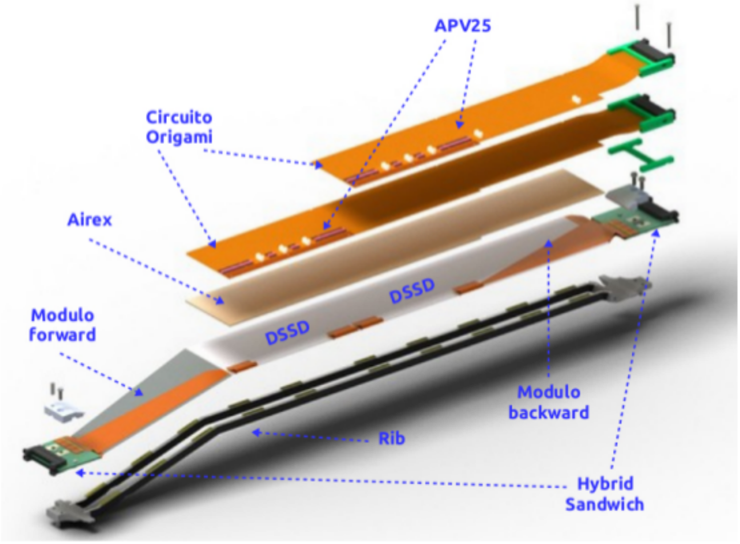


Figure 3.3: The rendering of a ladder of the layer 5. The main components are indicated in the figure.

which connect the sensor to the APV25 chips on the *Hybrid Sandwich*.

Despite the different geometry and sensors, the forward and backward modules are essentially identical in their structure, as shown in Figure 3.4. The most relevant technical details for both trapezoidal sensors used for the forward modules and the rectangular ones used fo the backward modules, are given in Table 3.1. The modules will be placed onto the beam pipe with the p-side facing the IP and measuring the $r - \phi$ coordinates, as already explained in the general SVD overview in Section 2.4.3. The n-side hits provide the z-coordinate measurement.

The *Hybrid Sandwich* (HS) is composed of two single hybrid boards, one for each side of the sensor: on the p-side single hybrid, 6 APV25 chips are implanted, while 4 chips are mounted on the n-side single hybrid. The two single hybrid boards are then glued together to form the HS, which consists of 10 chips. Each chip collects the signal from 128 strips, with a shaping time of 50 ns, which represents a factor 10 improvement with respect to the previous Belle readout electronic performance.

The *Pitch Adapter* (PA) purpose is to connect the silicon strips to the APV25 chips through interconnection copper lines, among several layers of polyamide, matching the sensor and the chips. All the connections between PA, sensor and chips are made by Al/Si 25 μm-thick wire bonds on Ni/Au bonding pads. Four different types of PA, varying in PA and pads dimensions, are needed, one for each module side (forward p/n side, backward p/n side). The p-side PA has typical dimension of 6×2 cm^2, while on the n-side a larger PA is used, 6×16 cm^2. The bonding pads have different pitch on the sensor side with respect to the HS side: the former vary from 45 to 240 μm, while the latter is 45 μm. The PA has a 80 μm thickness and two small holes on its surface, provided for the alignment during the assembling procedure.

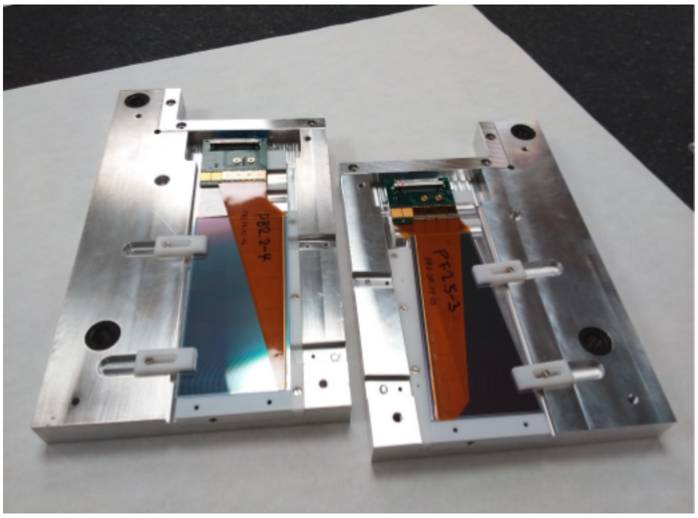


Figure 3.4: A backward (left) and a forward (right) module produced in Pisa, placed on the dedicated support (multi-purpose chuck). The *Pitch Adapter* corresponds to the orange part seen on both sensors. The modules are n-side up.

module	type of sensor	strips [p/n side]	strip pitch [p/n side] μm	size mm^2	active area mm^2
forward	trapezoidal	768/512	50-75/240	$126 \times (61 - 41)$	5893
backward	rectangular	768/512	75/240	125×60	7030

Table 3.1: Technical specifications for the forward and backward module sensors employed in the Belle II SVD construction.

3.4 The assembling procedure

The assembling procedure, unless for the specific mechanical tolerances, are the same for both forward and backward modules, whose type will not be specified anymore in the following section, and also for the p/n-side. The first step consists in the visual inspection of each component of the module, aimed at verifying its status and finding eventual damages that might prevent from using that component in the sub-assembly.

The second step is the HS assembling. The chip positions on the two single hybrids are measured and their shift with respect to the nominal position has a maximum tolerance of 200 μm. They are electrically tested and if all the measurements show a correct behaviour of the hybrids, they are glued together to form the HS. The relative alignment of each single hybrid is measured and a second visual inspection and electrical test follow the gluing. This further check is aimed at verify that the assembling phase did not introduce any new defects on the HS. Finally the HS is ready to be used in the module assembling.

The assembling is performed on mechanical positioning systems (jig) specific for each geometry and side of the sensor (backward, forward, p/n side), placed on the granite table of a Coordinate Measurement Machine (CMM), which provides a resolution on the relative placement of every component of 5 μm. The jig surfaces are covered with teflon and they can move in the vertical direction thanks to micrometer screws. The sensor and the HS are placed on the

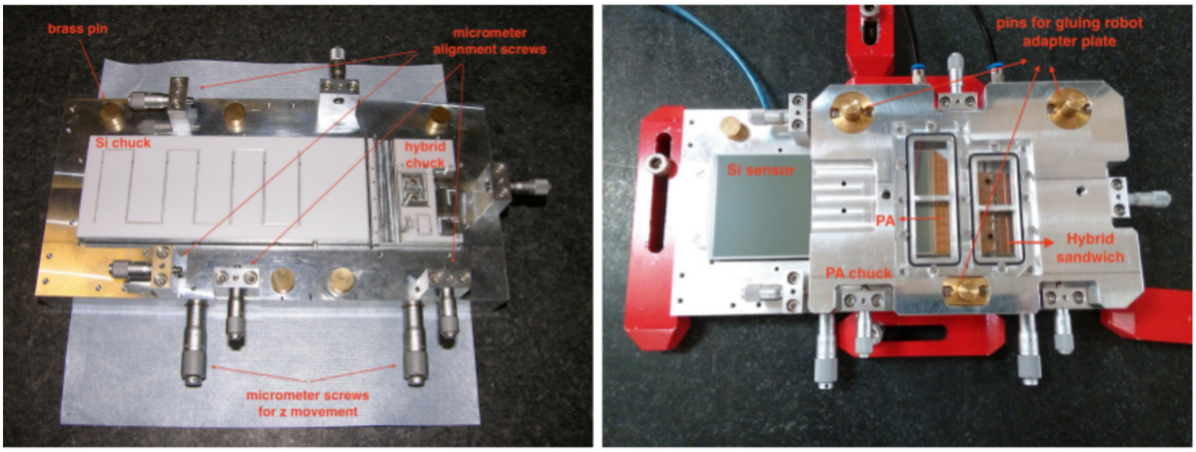


Figure 3.5: Left: the jig for the mechanical assembling of the module; sensors and HS supports are labeled in red, as well as the micrometer screws and the brass pins for the PA chuck positiong. Right: the PA chuck in place onto the jig.

jig and they are fixed to their nominal position by a vacuum system for the silicon sensor and by screws for the HS. The CMM provides coordinate axes in the reference system of the sensor for measuring the position of every component at the nominal reference crosses and calculating the right alignment of each piece, with a tolerance of 100 μm beteween sensor and HS. The final coordinates are registered and the jig is moved down in the z direction to prevent any sensor or HS contact with the PA.

The PA is positioned onto an appropriate chuck, on which it is kept in place by a vacuum system. The chuck is placed onto the jig on which the sensor and the HS have been already aligned. Micrometer screws allow the correct alignment of the PA chuck with respect to its nominal position on the jig (see Figure 3.5). The possible shrink of the PA with respect to its nominal dimension is kept into consideration to achieve a final placement within 150 μm maximum tolerance. The planarity of the region to be glued, with a maximum tolerance of 50 μm, is measured both for PA-sensor and for PA-HS side.

When the desired alignement is achieved, the glue is deposited onto the PA by a robot which employs a programmable coordinate system according to the type of PA to be glued. The glue preparation phase is extremely important to provide a controlled homogeneous deposition which avoids any bubble formation. The controlled environment of the clean room ensures also stability in umidity, pressure and temperature, which are fundamental factor for obtaining a correct and constant viscosity during consecutive gluing oprations. The PA chuck placed on a metallic support on the robot is positioned at the nominal deposition coordinates through micrometer screws: a dedicated alignement software implemented on the gluing robot tests the correct positioning of the PA chuck and finally the glue is deposited from a syringe moved by the gluing robot. The typical needle height on the PA surface is 200 μm. The PA chuck positiong on the gluing robot and the glue deposition are illustrated in Figure 3.6

After the glue deposition, the PA chuck is moved back to its orignal position onto the jig and then the height of the HS and of the sensor is gradually increased through micrometer screws until they reach the PA surface. The HS is first glued and then the sensor, as indicated by the

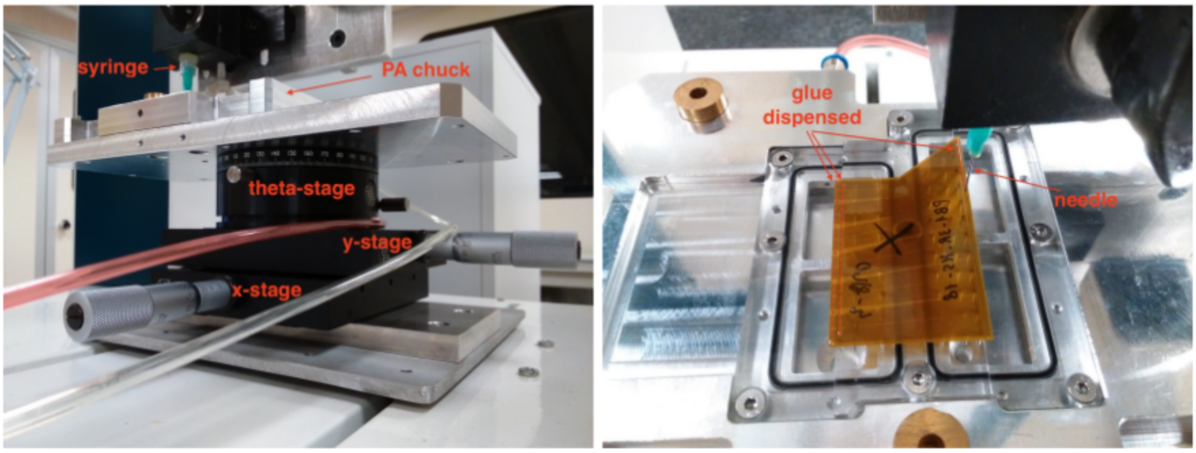


Figure 3.6: The PA chuck alignement with respect to the gluing robot coordinate system through the micrometer screws (left) and the glue deposition (right) are illustated in the picture.

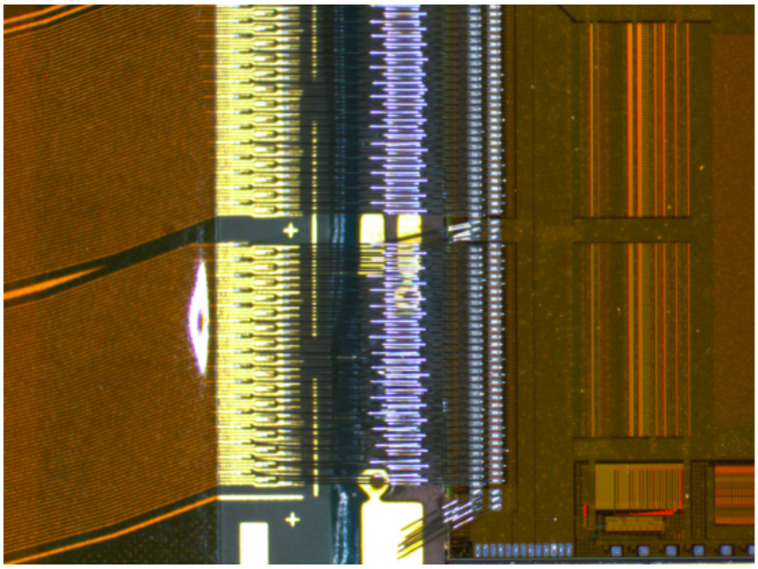


Figure 3.7: Microbonding between PA and chips; under the wire bonds the deposited glue is visible (blue halo).

approved gluing procedure. The glue diffusion is constantly monitored by the CMM: it has to uniformly spread out without covering the active area of the sensor or of the HS.

Once the glue has spread throughout all the gluing area, the micrometer screws are locked and after 15 hours curing time the chuck is removed and the gluing is complete. After the gluing, the reference crosses are measured again by the CMM and the visual inspection checks that the assembling has been successfully performed. After the module is safely turned up-side-down thanks to the appropriate fixture, the same gluing procedure is repeated on the opposite side (n) of the sensor and a second PA is glued on the sensor and HS.

At the end of the n-side gluing, after the cross-check measurements and the visual inspection, the assembling is concluded and both sensor and HS sides of the PA are wire bonded by an expert technician (see Figure 3.8). A last visual inspection is made to check the status of the sub-assembly before strating with the electrical testing.

3.5 The electrical testing procedure

The goal of the electrical tests performed in Pisa is to find and classify the defects of the modules, after the assembling. The sensors have been already tested in the sites of their production and the list of defects detected before the assembling is compared to that produced after the electrical testing in Pisa. The comparison underlines if the assembling procedure has introduced some new defects.

For the testing procedure, the HS is connected to a dedicated measurement station called *test stand* which have been realised in the INFN High Technology Laboratories in Pisa. The test stand setup is composed of:

- the APVDAQ acquisition system and its software, developed at HEPHY;
- low voltage power supply for APV25 chips;
- high voltage power supply for sensor reverse polarization (V_{bias});
- infra-red pulse laser for simulating strip response to incident radiation;
- motor for laser scan;
- the testing box which provides electromagnetic and light shielding for the module to test.

The electrical testing is performed through all its stages by the acquisition software. A LabView interface has been programmed to interact with both APVDAQ and the test stand setup (especially with the power supply and the laser scan).

Electrical tests are performed on single hybrids, HS and on the complete assembled module and they consist of:

- *passive measurements*: all the passive components are measured and compared to the nominal values;
- *I-V curve* (module only): at varying the reverse polarizing voltage V_{bias} in the range $0 - 100$ V (forward modules) or $0 - 200$ V (backward modules), the total leakage current is measured;
- *pedestal run*: when $V_{bias} = 100$ is applied, the mean signal level collected by each strip and its variance (noise) are measured; strips (or hybrids) are not injected at this stage and an internal random trigger select 600 events;
- *internal calibration run*: a pulse is injected into each chip channel and at varying of its timing Δt, the output signal shape is measured. A fit to the amplitude of the response signal, as a function of the pulse timing, allows to extract the maximum peak gain of output signal and the peak time;
- *V_{sep} scan* (module only):a fixed Δt pulse is injected into each chip channel and the output gain amplitude is measured; at fixed V_{bias}, the measurement is repeated varying the

voltages applied between the chip power supply and the sensor reference voltage, for each sensor side (the varied voltage is called V_{sep}), in the range $-5 < V_{sep} < 5$;

- *laser scan* (module only): sensors strips are illuminated by an infra-red pulse laser and the hits (defined by a signal to noise ratio higher than 6) registered by each strip are counted; the test is repeated twice with the laser moving in the orthogonal direction with respect to the sensor strips.

The acquired data are then analysed and according to the given definitions for each defect, the full list of the found defects is produced:

- *open*: strip, chip channel or bond broken; the defect is identified by high noise in the measured channel, unusual peak time and peak gain, low number of hits;
- *short*: contact between two consecutive strips or bonds; the defect is identified as low gain and low hit consecutive channels;
- *pinhole*: breaking of the silicon oxide layer between the aluminium sensor strip and the n-substrate; this defect corresponds to strips with lower gain at $V_{sep} = 5$ V or $V_{sep} = -5$ V, while the usual gain (between 70-100 ADC units) is recovered at $V_{sep} = 0.75$ V;
- *laser response*: lower hit number as response of the laser incident radiation.

Sometimes noisy channels might depend on the test stand setup: it is always fundamental to qualify the test stand before any testing session.

3.6 The sub-assembly quality overview

The mean quality of the assembled modules in Pisa is optimal. The 47 backward sub-assemblies, that were planned to be produced in Pisa, have been completed at the end of July, while the forward module production will terminate before the end of October, as scheduled. All these sub-assemblies are going to be used in the ladder assembling.

As concerns the mechanical quality, no defects (apart from three initial failures in forward modules gluing) have been detected in the assembled modules and all the reference points have been measured to be at nominal position, within the established tolerances.

The jigs show no evidence of degradation during time and the module quality stays constant through all the production period.

The electrical quality is mainly evaluated from the AC defect rate (opens, shorts and pinholes) found in the analysed modules, which over a sample of 33 backward and 25 forward sub-assemblies gives less than 1% defective strips per side. Including in the overall quality also the noisy strips, the percentage of defective strips remains lower than 1, while adding also the laser response result, the forward p-side only shows more than 1% strips as defective. However, the laser response is strictly dependent on the laser scan setup and it was found that in 4 modules an uncorrect setup was the real cause of the apparently bad response of the strips to the laser radiation.

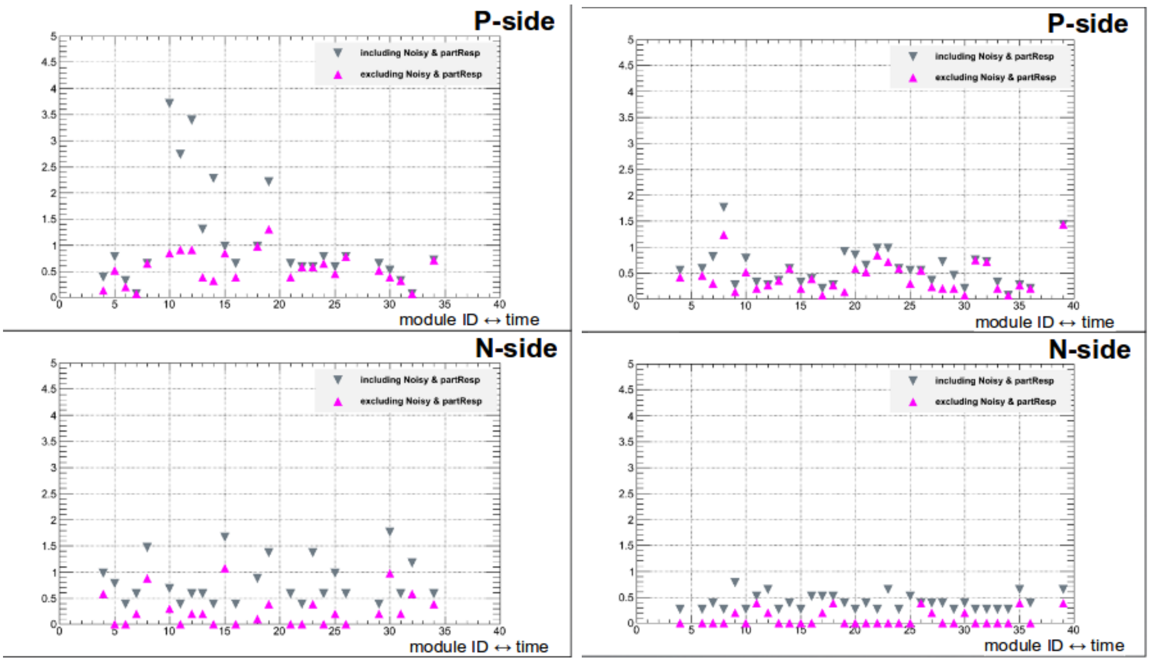


Figure 3.8: The total AC defects rate distribution versus the module identity number (alias versus time), for forward (left) and backward (right) modules. The percentage of AC defective strips (purple triangles) is constant and lower than 1. A slightly increased trend is shown if noise and laser response are included in the defect list.

As conclusion, the overall quality of the produced modules is excellent and the sum of all the DSSD AC defects and the new ones introduced during the assembling procedure in Pisa results in less than 1% defective strips per assembled module. The rate of the AC defects (summed between the DSSD AC defects that were already found on sensors and the AC defects introduced in Pisa) is shown as a function of the module identity number (progressive in time) in Figure 3.6 for both forward and backward modules.

In the course of my thesis period I actively participated in all phases of module assembling, gaining experience with the complex and delicate procedures. Not all modules were working perfectly from the beginning, and I contributed to understanding the effects of various kinds of defects on the sensors and on the modules.

A look forward: SVD performance in the analysis

The SVD performance, combined to the CDC measurements, is the main responsible for the track reconstruction. Its improvement is fundamental to handle the high luminosity challenge, which forsees a 20-30 times larger background rate with respect to the one measured at Belle. The global upgrade project of the VXD in Belle II will confirm the same and even better performances achieved by its predecessor.

Studies have shown that the track parameter resolution will be improved by a factor 2 with respect to the *BABAR* resolution. The output of the MC preliminary studies is shown in Figure 3.9. A better track reconstruction means better performance in the kinematic fit to find the common vertex of the decay products. This is a basic requirement exploited in the event

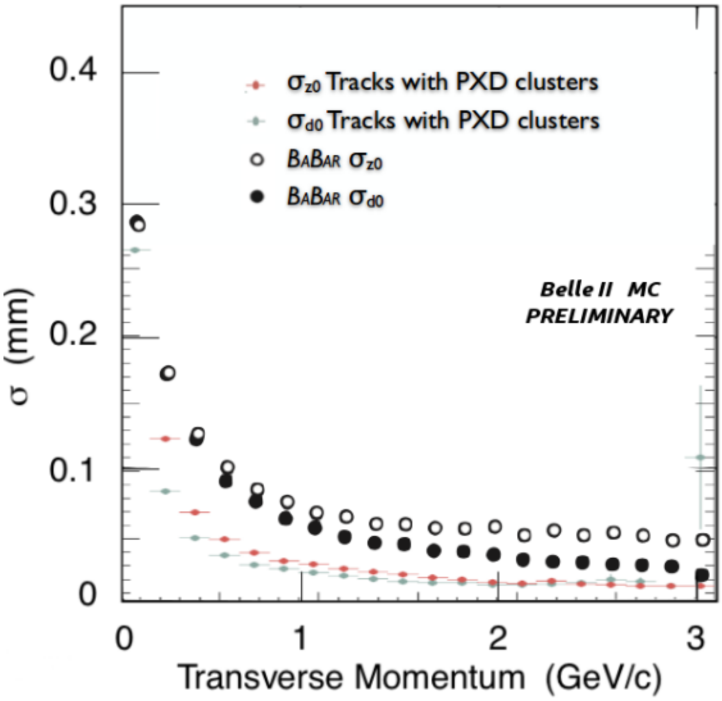


Figure 3.9: The Belle II resolution on the impact paramters d_0, z_0 as a function of the transverse momentum, compared to *BABAR* resolution.

reconstruction implemented in the analysis. The analysis method and the importance of the impact parameter in the selection optimization will be described throughout the next chapters.

Chapter 4

Analysis Method Overview

The goal of this thesis is to measure the branching fraction for the $B^0(\bar{B^0}) \rightarrow p p \bar{p} \bar{p}$ decay mode. The first step is to reconstruct B candidates decaying to the final state $p p \bar{p} \bar{p}$, which defines the signal signature for this analysis.

B meson candidates are reconstructed summing the four momenta of final state particles and then identified using kinematic variables sensitive to B mesons properties, such as the beam-energy substituted mass m_{ES} and the energy difference ΔE (see Section 2.1.1). In this chapter a general overview of the BF($B^0(\bar{B^0}) \rightarrow p p \bar{p} \bar{p}$) measurement method is given. The concept of a *blind* analysis and the strategy for the signal yield extraction are introduced (Section 4.1).

In Section 4.2, I will focus on the first analysis step, which consists in the B meson reconstruction from the final state $p p \bar{p} \bar{p}$. The reconstruction efficiency is evaluated on the signal Monte Carlo sample.

The data sample used for this analysis is the full dataset collected with the BABAR detector during Run$(1-6)$, for a total on-peak integrated luminosity of 424 fb^{-1}. The experimental data sample, together with the Monte Carlo samples on which the optimization studies are performed, are described in Section 4.3, which also includes an estimate for the expected number of signal and background events, before any further selection for background rejection is applied, assuming a signal branching fraction of 10^{-7}.

4.1 The analysis method

The strategy that I have implemented for this analysis can be summarized in 4 main steps:

1. The *reconstruction* of the event;
2. The *event selection*;
3. The *fitting procedure*, which can be divided into two parts:
    - The modeling of signal and background shape in the m_{ES} distribution;
    - The signal yield extraction.
4. the branching fraction calculation.

In this thesis, the variable studied to extract the signal yield is the beam-energy substituted mass m_{ES}. The strategy of extracting the signal yield consists in performing an unbinned extended maximum likelihood fit to the m_{ES} distribution, in the whole reconstruction range $5.2 < m_{ES} < 5.3$ GeV/c^2.

In the *reconstruction* step, four tracks identified as protons [1] are combined and the m_{ES} and ΔE variables are calculated. Only the candidates in a loose range of m_{ES} and ΔE and whose decay vertex has been successfully reconstructed are accepted. At that point, the *event selection* is applied, in order to discriminate signal from background events among all the reconstructed candidates. This selection consists of two steps: a first *preselection* aims at rejecting the bad-quality reconstructed events. It is based on track quality requirements, on the probability of the fit performed to the common B-vertex and on the particle identification response. The second step relies on the implementation of a Multivariate Analysis selector trained on different event variables, whose distribution helps separate the signal events from the hadronic events coming from *continuum* backgrounds (see Section 4.2 for appropriate definition). The Monte Carlo samples are split in two equal subsets, one used for the training of the Multivariate Analysis methods and the other for the testing and for application. The full selection is explained in detail in Chapter 5.

The m_{ES} distribution for all the events that passed the selection is then fitted with an unbinned extended maximum likelihood fit. The *fitting procedure* consists of two parts: first the signal and background shape in the m_{ES} distribution is studied on Monte Carlo samples and appropriate probability density functions are defined; second, the signal yield extraction procedure is implemented, once the shape parameters have been fixed to the fit results from the modeling studies.

A blind analysis strategy is employed to avoid any experimenter's bias. For the signal shape modeling, the *blinding* procedure is applied by fitting the signal Monte Carlo sample. For the background modeling, the m_{ES} distribution shape is studied on both the background MC samples and on the side band region of on-peak data (and also off-peak data). The region of data where the signal is expected is kept hidden, until the full analysis is complete.

The side band interval corresponds to the range in m_{ES} going from the beginning of the reconstruction range, 5.2 GeV/c^2, to the beginning of the signal window. The signal window consists in the range $5.27 < m_{ES} < 5.29$ GeV/c^2 and corresponds to a 3σ region around the B mass peak at 5.279 GeV/c^2. Background modeling studies performed on on-peak data refer exclusively to the side band region, as defined above. No fit on data has been performed in the signal window of m_{ES}, since the unblinding has not been allowed yet.

Once the unblinding is allowed, the unbinned extended maximum likelihood fit to the m_{ES} distribution for on-peak data will extract both the number of signal and background events. The signal yield N_{sig}^{obs} is then exploited to calculate the branching fraction of the baryonic mode $B^0(\bar{B}^0) \to p p \bar{p} \bar{p}$ as explained in the Chapter 7.

---

[1] In the rest of the thesis, I will refer to the signal mode as 4 proton final state, without distinguishing between protons and antiprotons. When discrimination between positive and negative charged particles is needed, it will be explicitly stated.

4.2 The candidate reconstruction

The search for the fully reconstructed $B^0(\bar{B^0}) \rightarrow pp\bar{p}\bar{p}$ starts with the combination of 4 charged tracks. These tracks are the result of tracking fits to the hits and energy deposits that the final state particles leave in the *BABAR* subdetectors when propagating in the 1.5 T magnetic field from the interaction point through the Silicon Vertex Tracker and the Drift Chamber (see Section 2.3.8).

Once tracks are reconstructed and assigned their particle identities, a kinematic fit combines their four-momenta to form a common vertex, which represents the mother particle candidate. Wrong combinations are the source of what is called *background* for a given decay mode, consisting mainly of two classes:

- the *combinatorial background*, composed of random combinations of tracks which are not coming from the same decaying particle but which roughly produce the invariant mass of the searched mother candidate. An important source of combinatorial background is the *continuum background*, which arises from e^+e^- events where pairs of light quark-antiquark are produced instead of B mesons. This is the dominant background component in rare B decays, which do not proceed via the favored $b \rightarrow c$ transition.

- the *physics background*, which originates from B meson decays to final states that can be mis-identified as the signal final state, because they show a similar topology to the $B^0(\bar{B^0}) \rightarrow pp\bar{p}\bar{p}$ tracks. It is related to PID mis-identification probabilities, when at least one track is wrongly identified as a proton. This is potentially more dangerous than the combinatorial background, since the invariant mass distribution of the fake candidates usually peaks at the same value as the signal.

The candidate reconstruction consists of three steps:

1. the selection of 2 positive and 2 negative charged tracks, identified as protons and antiprotons respectively. For the particle identification a combined list of *very loose* selectors is used, as described in Section 2.3.9. The main characterization of the PID selectors exploited in the reconstruction is given by their efficiency and their mis-identification probabilities, which differ according to the type of the selector. PID performance studies show that the exploited selectors have an efficiency higher then 0.98 for protons and antiprotons of 1 GeV/c momentum and mis-identification probabilities lower than 0.05. The positron and electron fake rate for Likelihood selectors increases until 0.2 for 1 GeV/c momentum tracks, but it is cured in the Decision Tree PID method.

2. The kinematic fit on the 4 charged tracks to form a common vertex. This procedure, known as *vertexing*, is based on the `TreeFitter` algorithm implemented in *BABAR* software to reconstruct the mother particle from the detected tracks. The script relies on the Kalman Filter fit technique and more details about its implementation can be found in [21]. At reconstruction level, no cuts are imposed on the probability of the fit result and only its convergence is required. The vertexing procedure assigns a position and a four momentum

to the common vertex of the four tracks, identified as the mother particle candidate of the event.

3. The kinematical cuts on the values of ΔE and m_{ES}, calculated on the four-momentum assigned to the mother particle candidate, according to the formula described in Section 2.1.1. Loose acceptance ranges are defined for both variables. Reconstructed candidates are accepted if the following conditions are satisfied:

$$5.2 < m_{ES} < 5.3 \text{ GeV/c}^2 \;; \qquad |\Delta E| < 0.2 \text{ GeV}. \tag{4.1}$$

Only when all the above requirements are fullfilled, the event is reconstructed and all the kinematic information about the 4 proton tracks and the B meson candidate are filled in a *ntuple* [2]. The same procedure is applied both to Monte Carlo-generated and real data events.

4.3 The samples

In this analysis, three types of samples have been used: the first two types are the on- and off-peak events as described in Section 2.2.1, which are real data collected by the BABAR experiment. The third type consists of the Monte Carlo samples. The latter type is used for optimization studies, for the BDT training and testing after being split in two equal subsets, for the efficiency calculation and the signal and background modeling. The formers are exploited for the analysis and cross-check between MC and data results. The software release version used for processing all the three types of data is given in Section 2.3.8.

4.3.1 Experimental dataset

The dataset used for this analysis consists of on-peak and off-peak data collected from Run 1 to Run 6. The luminosity of each Run processed and the corresponding yield of $B\bar{B}$ pairs (for on-peak data only) are calculated with the *B counting* method [22] explained in Section 7.1.3.

To directly compare the on-peak and off-peak samples, the weight

$$w_{off} = \frac{L^{\text{onpeak}}}{L^{\text{offpeak}}} \cdot \frac{E^2_{\text{offpeak}}}{E^2_{\text{onpeak}}} = 9.58557$$

is applied to off-peak data. The scaling is given by the ratio between the two integrated luminosities times a factor $(E^2_{\text{offpeak}}/E^2_{\text{onpeak}})$ which takes into account the dependence of the cross section on the *centre-of-mass* energy.

In Table 4.1, integrated luminosity and $N_{B\bar{B}}$ are listed for each on-peak Run processed, while integrated luminosity for off-peak data is shown in Table 4.2.

---

[2] A *ntuple* is the basic data structure of the ROOT-based analysis system.

Run	$\int \mathcal{L}dt$ (fb^{-1})	$N_{B\bar{B}} \pm \sigma_{stat} \pm \sigma_{syst}(\times 10^6)$
1	20.371	$22.55 \pm 0.03 \pm 0.14$
2	61.313	$68.43 \pm 0.04 \pm 0.41$
3	32.278	$35.75 \pm 0.03 \pm 0.21$
4	99.574	$111.39 \pm 0.06 \pm 0.67$
5	132.336	$147.58 \pm 0.07 \pm 0.89$
6	78.317	$85.18 \pm 0.05 \pm 0.51$
Total	424.190	$470.88 \pm 0.12 \pm 1.32$

Table 4.1: The integrated luminosity (fb^{-1}) and the B *counting* result ($N_{B\bar{B}}$) for each Run of on-peak data processed for the analysis are listed. The total integrated luminosity exploited for this analysis is $L^{\text{onpeak}} = 424.190$ fb^{-1}.

Run	$\int \mathcal{L}dt$ (fb^{-1})
1	2.564
2	6.868
3	2.443
4	10.015
5	14.277
6	7.752
Total	43.919

Table 4.2: The integrated luminosity (fb^{-1}) for each Run of off-peak data processed for the analysis is listed in the table. The total off-peak integrated luminosity is $L^{\text{offpeak}} = 43.919$ fb^{-1}.

4.3.2 Monte Carlo samples

The Monte Carlo (MC) samples consist of two categories: the signal sample and the background samples.

For the former, a dedicated simulation has been provided by the BABAR official MC production and 687000 signal events of $\Upsilon(4S)$ decays in $B^0\bar{B^0}$ pairs have been generated. Each event consists of one B meson set to generic [3] decay, while its conjugate decays to the final-state signal signature ($pp\bar{p}\bar{p}$). The B meson decay model used for the mode $B^0(\bar{B^0}) \rightarrow pp\bar{p}\bar{p}$ is the *phase space* model, which assumes each momentum configuration with the same probability (flat in the phase space).

As concerns background samples, collections already implemented in BABAR MC have been studied: $B^0/\bar{B^0}$ to generic, B^+/B^- to generic and hadronisation from *continuum* processes $e^+e^- \rightarrow q\bar{q}$, $(q = u, d, s)$, which will be summarized from now on with the notation $e^+e^- \rightarrow uds$, and $e^+e^- \rightarrow c\bar{c}$. Samples simulating lepton pairs events have been added for describing the backgrounds composition seen in data: the Bhabha scattering $e^+e^- \rightarrow e^+e^-(\gamma)$, the μ and τ pairs production from continuum, $e^+e^- \rightarrow \mu^+\mu^-(\gamma)$, $e^+e^- \rightarrow \tau^+\tau^-(\gamma)$.

In Table 4.3 the number of generated and reconstructed events (N_{gen}, N_{reco}) for each sample studied is shown. The third column shows the weight applied to each reconstructed sample to scale the number of events to the data integrated luminosity for Run 1-6. It is worth noting that

---

[3] The B meson is allowed to decay to all known (and implemented in the EvtGen [17]) decay modes, according to their respective branching fraction and decay models.

Sample MC	N_{gen}	N_{reco}	Weight
Signal $B^0(\bar{B}^0)\to pp\bar{p}\bar{p}$	687000	292705	6.85415×10^{-5}
$B^0/\bar{B}^0$ to generic	$7.17995\ \times10^8$	13639	0.327913
B^+/B^- to generic	$7.08062\ \times10^8$	22169	0.332513
$e^+e^-\to uds$	$16.19504\ \times10^8$	75415	0.547425
$e^+e^-\to c\bar{c}$	$11.28544\ \times10^8$	28967	0.488636
$e^+e^-\to e^+e^-(\gamma)$	$3.44000\ \times10^8$	46	49.3244
$e^+e^-\to\ \tau^+\tau^-(\gamma)$	$3.12900\ \times10^8$	87	1.27433
$e^+e^-\to\ \mu^+\mu^-(\gamma)$	$3.981\ \times10^8$	0	–

Table 4.3: List of Monte Carlo samples used in the analysis. In the second column, the number of the generated events for each sample is provided and in the third column the corresponding number of reconstructed events is listed. The weights to scale the samples to the data integrated luminostity, 424.190 fb^{-1}, are reported in the last column. The scaling is proportional to the production cross section for each event type, while for the signal MC sample it is based on the assumption of a BF of 10^{-7}.

no event from $e^+e^-\to\ \mu^+\mu^-(\gamma)$ has fulfilled the reconstruction requirements and this sample will not be reported anymore in the next tables.

The weight w_i for each sample i is calculated normalizing the total number of generated events N^i_{gen} to the number of expected events for the given sample at the data integrated luminosity, which can be computed as the product $\sigma_i \cdot L^{onpeak}$, where σ_i refers to the production cross section for the sample i at the centre-of-mass energy of the $\Upsilon(4S)$:

$$w_i = \frac{\sigma_i \cdot L^{onpeak}}{N^i_{gen}}. \tag{4.2}$$

From now on, when mentioning the weighted m_{ES} distributions, I will refer to these weights applied to the analysed MC samples. The scaling for the signal MC sample requires an assumption on the expected branching fraction (BF) of the decay mode $B^0(\bar{B}^0)\to pp\bar{p}\bar{p}$, since the production cross section is not known. The weight applied to the signal MC sample is computed according to the formula $w_{sig} = BF \cdot N_{B\bar{B}}/N^{sig}_{gen}$. A rough estimate of the BF can be derived by the upper limit on the branching fraction of the baryonic mode $\bar{B}^0 \to \Lambda_c^+ p\bar{p}\bar{p}$ measured by Grünberg et al. [10], as described in Section 1.5. For this analysis, I assumed $BF = 10^{-7}$.

4.3.3 Expected number of events

The m_{ES} distribution of the reconstructed events from the signal MC sample is reported in Figure 4.1 (left). As expected for B candidates coming from signal events, the distribution of the beam-energy substituted mass shows a peak at the B meson mass. The experimental resolution is smaller than 2 MeV/c^2, as shown in the right plot in Figure 4.1. Only the *truth-matched* candidates are plotted. The *truth-matching* principle exploits the fact that in the MC simulation the true identity and properties of every particle are completely known. When the interaction with the detector is simulated and the reconstruction procedure is applied, the wrong identity might be assigned to a generated particle. A candidate is called *truth-matched* when its reconstructed identity and the identity of its reconstructed mother particle match the

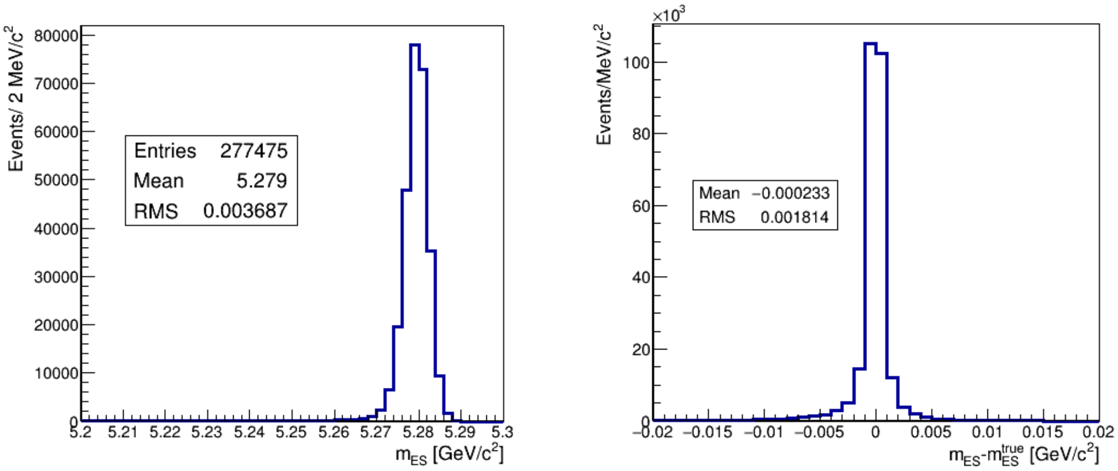


Figure 4.1: Left: the m_{ES} distribution of the truth-matched reconstructed events from the signal MC sample is shown. The plotted range correponds to the full reconstruction range for m_{ES}. The signal is contained as expected in the signal window $5.27 < m_{ES} < 5.29$ GeV/c^2. Right: the distribution of the difference between the m_{ES} value calculated from the reconstructed B meson candidate and the value assigned at MC generation stage. The width of the plotted distribution represents the experimental resolution for m_{ES}, around 2 MeV/c^2.

identities in MC generation. In this analysis, I will refer to a *truth-matched* signal B-meson candidate (or equivalently a truth-matched signal event) when, considering an event from the signal MC sample, all the four tracks detected as protons are *true* protons according to the MC generation assignments and they all come from the same mother particle whose *true* MC partner is a neutral B meson. This is a powerful method for investigating the sources of background on MC samples and testing the mis-identification probabilities. Applying the truth-matching to the reconstructed events from the signal MC sample (Table 4.3) gives $N^{true}_{reco} = 277475$, which results in a 5% rate of not-truth-matched signal events. Moreover, the difference between the value for m_{ES} calculated from the measured B meson's four-momentum and the value computed from the four-momentum assigned at the MC generation stage, ${m_{ES}}^{true}$, gives an idea of the expected experimental resolution for the measurement of this variable.

A preliminary estimate for the signal efficiency can be calculated by considering the ratio between the number of the truth-matched reconstructed events and the total number of generated events:

$$\epsilon^{sig}_{reco} = \frac{N^{true}_{reco}}{N^{sig}_{gen}} = 0.4039 \pm 0.0006. \tag{4.3}$$

This represents an upper limit for the signal efficiency that will directly enter the BF calculation. On the reconstructed events the selection for optimizing the background rejection is applied: it leads to tighten the requirements for accepting a B meson candidate and consequently to decrease the efficiency, in order to maximize the chosen *figure of merit*. In this analysis, since the goal is not precision, but maximizing the probability for measuring a new decay mode, I chose the

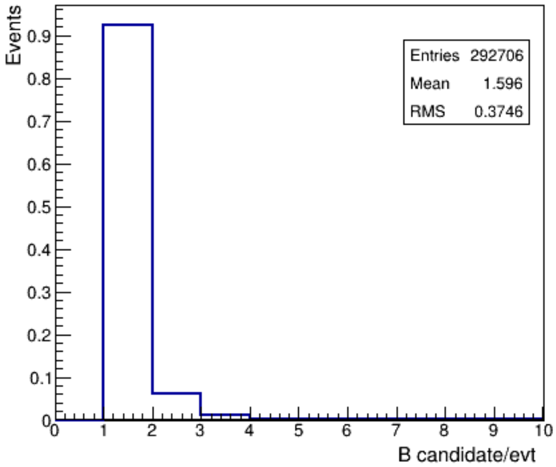


Figure 4.2: The normalized distribution of the number of reconstructed B meson candidates per event, for the signal MC sample. The mean B candidate multiplicity per event, shown in the box, is lower than 2.

significance, s, as figure of merit to optimize. The significance is defined as follows:

$$s = \frac{S}{\sqrt{S+B}} \tag{4.4}$$

where S and B are respectively the number of signal and background events counted in the signal window $5.27 < m_{ES} < 5.29$ GeV/c^2 of the m_{ES} distribution when the proper scaling for normalizing the MC samples to the full data luminosity is applied. This technique of counting the number of events in a defined region of a given variable distribution is the *cut-and-count* strategy. As mentioned in the analysis method overview, the signal yield extraction will be performed by a *fitting procedure* applied to the full reconstruction range of m_{ES}, exploiting shape information to separate the signal from the background events. However, for the selection optimization purpose, the cut-and-count technique will be used in the next chapters for evaluating the number of events selected on each MC sample after applying different sequential cuts.

The clean signature of 4 charged tracks identified as protons and coming from the same vertex is a very strong requirement in the event reconstruction: most times only one candidate per event is reconstructed. The plot of the B candidate multiplicity distribution for the signal MC sample is shown in Figure 4.2. In order to select only one candidate per event a ΔE selection is added to the reconstruction requirements. When more than one candidate is reconstructed in the same event, the one with the minimum absolute value for ΔE is kept, while the second is rejected, considering that ΔE peaks at zero for B-meson candidates coming from signal events. In this way, a one-to-one correspondence between reconstructed events and B meson candidates is ensured[4].

The number of reconstructed signal events expected at the full luminosity of 424 fb^{-1} in the

---

[4] The terms *candidate* and *event* can be equivalently exchanged in the efficiency calculation.

signal region is calculated:

$$N_{sig}^{exp}(424\ \text{fb}^{-1}) = 19 \tag{4.5}$$

The impact of the event selection on this number will be considered in detail in the next chapter.

Chapter 5

Event Selection and Validation

The event selection is the second step in the analysis strategy. Further requirements for accepting a B meson candidate as a signal event are added, in order to reject fake combinations that have been wrongly reconstructed. General kinematical constraints have been already imposed at the reconstruction level. Further cuts on track quality related variables eliminate the badly reconstructed tracks that can be more easily mis-identified in the candidate reconstruction. This is achieved thanks to the preselection, described in Section 5.1: events that passed both the reconstruction and the preselection steps undergo to the final background rejection phase.

This second part of the event selection mainly uses event shape variables. The aim of the event selection optimization is to maximize the significance s (Equation (4.4)) as figure of merit. There are several approches for implementing the background rejection: from optimization studies, in this thesis I decided to use a *multivariate analysis* method, which is quite common in High Energy Physics for classification purpose. All the tests have been performed on the signal and background MC samples, using the libraries implemented in the ROOT framework [23]. In multivariate analysis methods, event variables are combined to form a single discriminating selector, on which the final cut is applied. The goal of multivariate classification is to improve the significance with respect to sequential cuts on individual variables.

Different methods have been compared and different input combinations of the event variables have been tested. From these studies, the Boosted Decision Tree method, trained on a combination of 4 input variables, is the best perfoming technique with respect to the significance improvement. Its implementation is given in Section 5.3 and the BDT selection validation, which is fundamental for a realiable efficiency estimate, is discussed in Section 5.4.

5.1 The preselection

The goal of the preselection is to eliminate the fake combinations arising from badly-recontructed tracks and bad fit to the common vertex. Three requirements are added for accepting a reconstructed B meson candidate:

Sample	N_{passed}	$N^{exp}_{[s.r.]}(424 \text{ fb}^{-1})$
Signal $B^0(\bar{B}^0) \to p p \bar{p} \bar{p}$	233429	15.86 ± 0.03
$B^0/\bar{B}^0$ to generic	1497	77 ± 5
B^+/B^- to generic	3218	169 ± 7
$e^+e^- \to uds$	15581	1490 ± 30
$e^+e^- \to c\bar{c}$	2352	215 ± 10
$e^+e^- \to e^+e^-(\gamma)$	0	0
$e^+e^- \to \tau^+\tau^-(\gamma)$	0	0
Bkg_{TOT}	22648	1950

Table 5.1: For each analysed MC sample, the number of events passing the preselection step is reported in the table. Their statistical uncertainty is the Poisson error $\sqrt{N}$. The last column shows for each sample the expected events at data integrated luminosity in the signal window. These numbers are calculated applying the appropriate weights w_i (see Equation (4.2)) to each sample. The associated uncertainties take into account the scaling applied.

1. each proton candidate has to satisfy:

$$|doca^{XY}| < 1.5 \text{ cm}, \qquad\qquad |doca^{Z}| < 2.5 \text{ cm},$$

   where the signed distance of closest approach (*doca*) is referred to the point of the track helix which minimizes the distance in the xy transverse plane ($doca^{XY}$) and along the z coordinate ($doca^{Z}$) with respect to the origin (the interaction point);

2. the χ^2 probability of the fit to the common vertex is required to be larger than 0.001, in order to reject bad fits to random combinations of tracks;

3. the tracks have to be identified as protons by the *pKM* selector (see Section 2.3.9), available in six tightness modes. PID performance studies show that the best significance is achieved with the *SuperLoose* tightness level. In Figure 5.1 the characterization of the selector is provided by the plots for its efficiency and its mis-identification probability rates.

The m_{ES} distribution for the signal and the background MC samples, after the preselection application, is reported in the plot in Figure 5.2 (top). All the histrograms are scaled to the data integrated luminosity. The number of signal and background events that passed both the reconstruction and the preselection step are listed in Table 5.1. In the same table, the last column reports the number of signal and background events in the signal window $5.27 < m_{ES} < 5.29$ GeV/c^2 foreseen at the data integrated luminosity, for each analysed sample.

The QED backgrounds coming from $e^+e^- \to e^+e^-(\gamma)$ and $e^+e^- \to \tau^+\tau^-(\gamma)$ interactions are fully rejected by the preselection step. This is most likely due to the PID optimization, since the Decision Tree based selector has lower mis-identification probabilities for muons and electrons. Moreover the χ^2 probability cut eliminates all the combinations arising from the wrongly reconstructed vertices, which might gather charged tracks coming from the halo of the interaction point or generated by the particle interaction with the beam pipe walls.

The total number of background events expected in the signal region is $B = 1950 \pm 30$, while the signal events are $S = 15.86 \pm 0.03$. The errors associated do not correspond to the obvious statistical uncertainty $\sqrt{N}$, but they correctly take into account the scaling applied. From these values, the significance after the preselection step is:

$$s = \frac{S}{\sqrt{S+B}} = 0.357 \pm 0.003. \tag{5.1}$$

There is no hint for peaking background in the m_{ES} distribution from what is shown in Figure 5.2. Most background comes from the $e^+e^- \to uds$ continuum interactions, whose event topology shows a typical jet-like shape. Also the $sin\theta^*_B$ angular distribution of the B meson candidates in the $\Upsilon(4S)$ rest frame helps discriminate from continuum events $e^+e^- \to f\bar{f}$, distributed like $1 + cos^2\theta^*_B$. In *BABAR* analyses, information from phase-space distribution of the decay products has been quantified in different ways. The event shape variables exploited for background suppression in this analysis will be explained in the following sections.

The preselection has an impact on the signal efficiency, calculated as the number of the truth-matched passed events with respect to the total generated:

$$\epsilon_{sig} = \frac{N^{true}_{passed}}{N^{sig}_{gen}} = 0.3367 \pm 0.0006. \tag{5.2}$$

This number will decrease again because of the further background rejection applied, which aims at optimazing the significance.

5.2 Studies on the event selection

The first studies for the background rejection relied on sequential cuts on disciminating event variables. In the rest of the chapter, this strategy will be refered to as the cut-based selection: it is no more applied in the finalized event selection, but it is useful to show its performance in comparison to the BDT method (Section 5.3).

The choice of the variables and the cuts to impose on their distributions were determined by MC studies. I considered physically meaningful variables, expressing some kinematical constraints or some topologycal properties of the event, with a distribution consistent between the MC samples and data side band region events. Of course, the shape of the variable distribution has to be different for signal and background events to provide a good separation power.

Moreover, it is important to show that the chosen cut is not influencing the shape of the m_{ES} distribution, otherwise it has to be taken into account in the fitting procedure. This would mean to have a selection dependent background modeling, which would result in increasing the systematic uncertainty on the signal yield extraction.

5.2.1 The event shape variables

The phase-space distributions of both the B meson candidate and its decay products provide discriminating power between background and signal events, according on their shape. Lots of variables that quantify these properties have been implemented in the history of the B Factories.

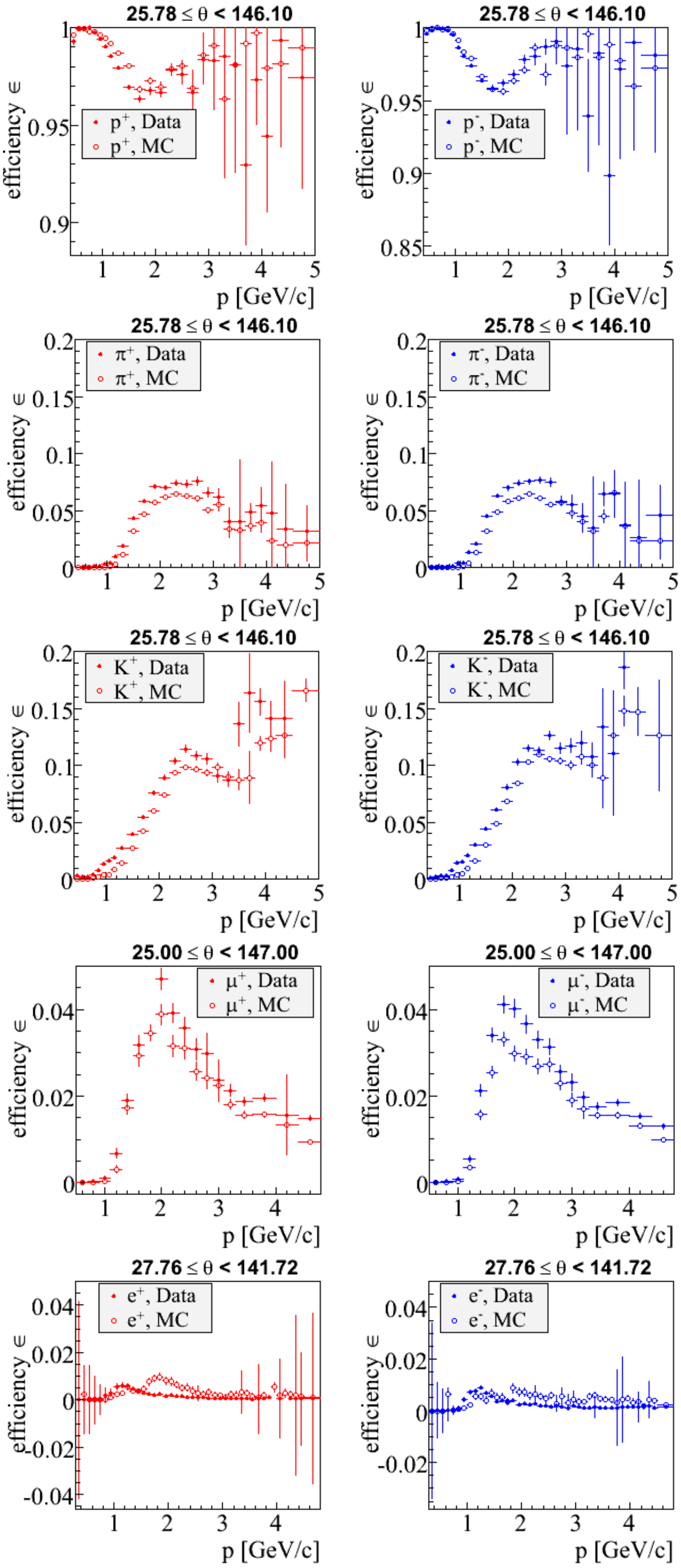


Figure 5.1: The *super loose* pKM selector efficiencies and mis-identification probabilities versus the particle momentum are shown. The first row represents proton and antiproton PID efficiency as a function of the particle momentum, in the azimuthal acceptance region covered by the detector, 25.78°-146.10°. Starting from the second row till the bottom, plots for mis-identification probability are reported: from top to bottom, pion, kaon, muon and electron fake rates are shown as a function of the particle momentum. The azimuthal acceptance slightly varies for the last two rows, depending on the sub-detectors involved in the particle identification.

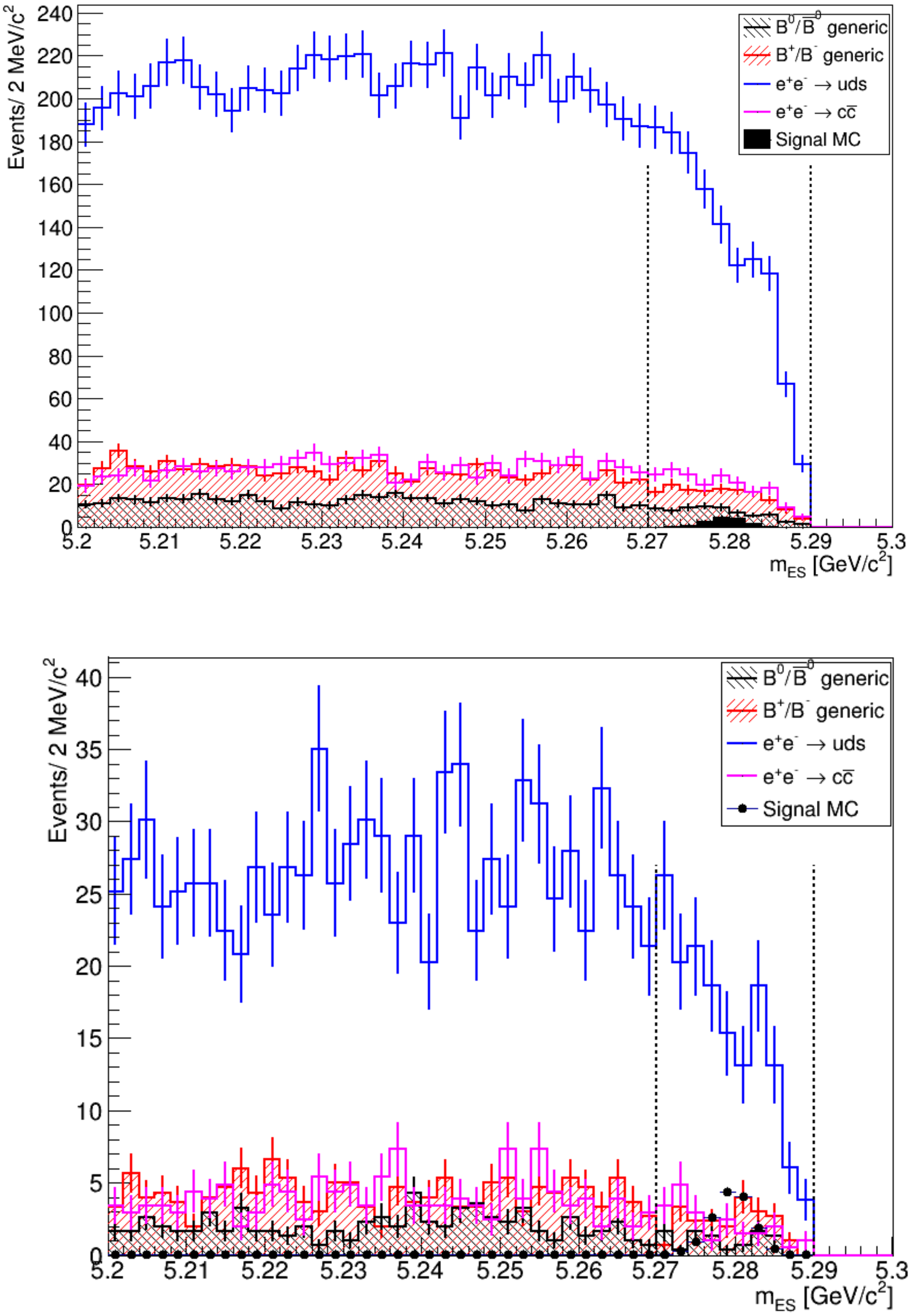


Figure 5.2: The weighted m_{ES} distributions for the analysed MC samples at different stages of the event selection are shown. *Top:* after the preselection; the signal MC sample is the black filled histogram. *Bottom:* the cut-based selection result (Section 5.2.2); the signal MC distribution is given by the black dots. All the histrograms have been scaled to the data integrated luminosity and the legend is given in the box on top right corner. The error bars show the correctly scaled Poisson uncertainty for each bin. The signal region is highlighted in the plots by the vertical dashed lines.

The most important for this analysis are:

- *The B meson direction*: angle θ^*_B, which is the angle between the direction of the B meson reconstructed momentum and the beam axis in the $\Upsilon(4S)$ rest frame, is expected to have a $sin^2\theta^*_B$ distribution for correctly reconstructed B meson candidates, in contrast with the $1 + cos^2\theta^*_B$ behaviour typical of $e^+e^- \to q\bar{q}$ events, with q a 1/2-spin particle;

- *The Thrust axis, T*: it is defined as the direction that maximizes the sum of the projection of the momenta $\vec{p_i}$ along the same axis, for a collection of N particles $(i = 1, 2, ..N)$; it is calculated maximizing the expression:

  $$T = \frac{\sum_{i=1}^{N} |\vec{T} \cdot \vec{p_i}|}{\sum_{i=1}^{N} |\vec{p_i}|}. \tag{5.3}$$

  The related variable exploited in the analysis is the absolute value $|cos\theta_{TH}|$, where θ_{TH} is the angle between the thrust axis of the momenta of the B meson candidate decay products with respect to the thrust axis of all the other reconstructed particles in the event (the *Rest Of Event*, ROE); given that the B meson decay products are isotropically distributed in the $\Upsilon(4S)$ rest frame, the signal events are expected to have a flat distribution in $|cos\theta_{TH}|$, while the back-to-back fragmentation of $q\bar{q}$ events produces a distribution peaking at values of $|cos\theta_{TH}|$ close to 1.

- *The Fox-Wolfram moments*: for a collection of N particles, the k-th order moment of the Fox-Wolfram method is given by the equation:

  $$H_k = \sum_{i,j} \frac{|\vec{p_i}||\vec{p_j}|P_k(cos\theta_{ij})}{E_i E_j}, \tag{5.4}$$

  where $E_i, \vec{p_i}$ are the energy and the three-momentum for a given particle i of the N-collection and P_k refers to the k-th order Legendre polynomial as function of the angle θ_{ij} between the momenta of particles i and j. The variable used in the analysis is $R_2 = H_2/H_0$, which is the second Fox-Wolfram moment normalized to the zero-th moment. It can be shown that this variable takes values closer to one for events with two strongly collimated jets, providing a good discrimination between events with different topologies.

5.2.2 The cut-based selection performance

Different set of cuts have been tested and for each configuration the significance s (as defined in Equation (4.4)) has been calculated in order to establish the best performing selection, which is the one providing the highest value for s. This is achieved with the optimized cuts (added to the preselection cuts):

- $|\Delta E| < 0.03$ GeV;

- $R_2 < 0.5$.

Sample	After reconstruction $_{[s.r.]}$(424 fb^{-1})	After cut-based selection$_{[s.r.]}$ (424 fb^{-1})
Signal $B^0(B^0)\to pp\bar{p}\bar{p}$	19.12 ± 0.03	14.64 ± 0.03
$B^0/\bar{B}^0$ to generic	662 ± 15	10 ± 2
B^+/B^- to generic	1146 ± 20	24 ± 3
$e^+e^-\to uds$	6570 ± 60	180 ± 10
$e^+e^-\to c\bar{c}$	2384 ± 34	24 ± 3
$e^+e^-\to e^+e^-(\gamma)$	640 ± 180	0
$e^+e^-\to \tau^+\tau^-(\gamma)$	10 ± 4	0
Bkg_{TOT}	11412 ± 190	238 ± 10

Table 5.2: Before and after the cut-based selection: for each MC sample, the number of events expected in the signal region $5.27 < m_{ES} < 5.29$ GeV/c^2 of the weighted m_{ES} distributions after the reconstruction are listed in the second column. In the third column, the events selected after the full selection cuts application (preselection + optimized cuts) on each MC sample are reported. The number of events is scaled to the full data luminosity of 424 fb^{-1}.

The number of expected events in the signal region at the data integrated luminosity, before and after the cut-based selection, are given in Table 5.2. The weighted m_{ES} distributions after the cuts application are plotted, for every analysed sample, in Figure 5.2 (bottom). The significance can be finally calculated as:

$$s = \frac{S}{\sqrt{S+B}} = 0.93 \pm 0.02, \tag{5.5}$$

where S, B refer to the number of signal and total background events obtained by integrating the weighted m_{ES} distributions in the range $5.27 < m_{ES} < 5.29$ GeV/c^2. The efficiency of the cut-based selection on the signal MC sample is calculated from the ratio of the selected signal events over the total generated, which gives $\epsilon_{sig}^{cuts} = 0.31$. It has been tested that tighter cuts would reduce the signal efficiency more then suppressing the left background. The optimized cut-based selection provides an expected significance around 1σ, which is quite far from even a claim (3σ deviation) of an observation. Since the cut-based analysis cannot exploit the correlation between variables, further studies aiming at improving the significance have been performed using multivariate analysis methods.

5.3 The multivariate analysis method

The Toolkit for Multivariate Analysis (TMVA) is a collection of multivariate classification techinques implemented for data analysis in the ROOT framework. The version used for these selection studies is TMVA 4.2.0. The toolkit provides a machine learning environment for the training, testing and performance evaluation of different classification methods on user-supplied datasets. True event classification must be known and weights must be provided to let the method train on signal and background separation. Further documentation can be found elsewhere [24]. The advantage of using these techinques for classifying the events consists in the possibility of exploiting also correlated variables. Several studies with different input variables and different methods have been performed, in order to maximize the significance.

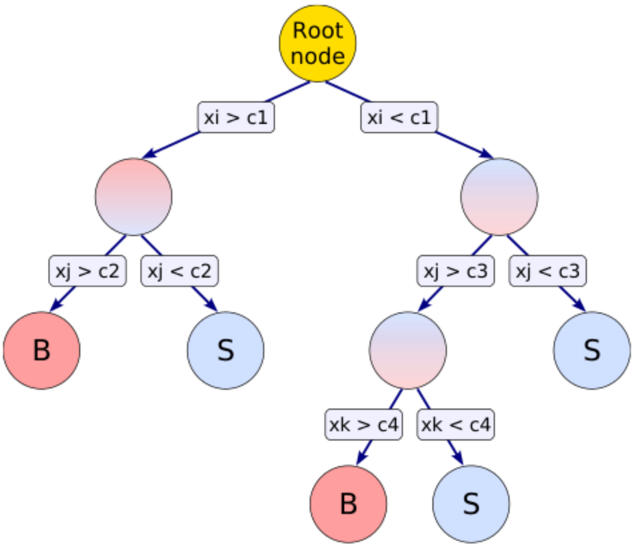


Figure 5.3: Schematic view of the structure of a decision tree. The sequence of splits applied to data at each node is represented by the rectangular labels. The binary classification is applied on the most discriminating variable x_i at each node. The final leaf nodes are classifed as signal (S) or background (B) according to the nature of most part of the events that end up in each leaf.

5.3.1 The decision tree method

A decision tree is a selector with a binary tree structure, as for instance the one visualized in Figure 5.3. Sequential splittings based on left/right decisions are implemented at each *node* until a stop criterion is fulfilled, typically a condition on the purity $\frac{S}{S+B} > 0.5$, and the phase space is divided into several regions, classified as signal or background depending on the majority of the training events that belong to the final *leaf* node.

The boosting of a decision tree consists in repeating the training for various trees, built from the same training ensemble by reweighting the events. The Boosted Decision Tree (BDT) is the single classifier obtained by averaging all the trained decision trees. The boosting technique is used to stabilize the response of the selector with respect to the statistical fluctuations of the training sample, which largely influence the response of an individual tree.

The BDT method implemented for this analysis exploits the default configuration set in the TMVA toolkit, with a total number of trained trees `NTrees`=800, `nCut`=20 corresponding to the number of grid points in the range of the input variables used to find the optimal splitting and `MinNodeSize`= 5% as the minimum percentage of the training events required at a leaf node.

The BDT training and its performance on data are explained in detail through the rest of the chapter.

5.3.2 The Boosted Decision Tree selection

At this point of the analysis, I split the MC samples into two equal subsets of equal size: one is exclusively used for the BDT training and the other for the testing phase, for estimating the efficiency and implementing the fitting procedure.

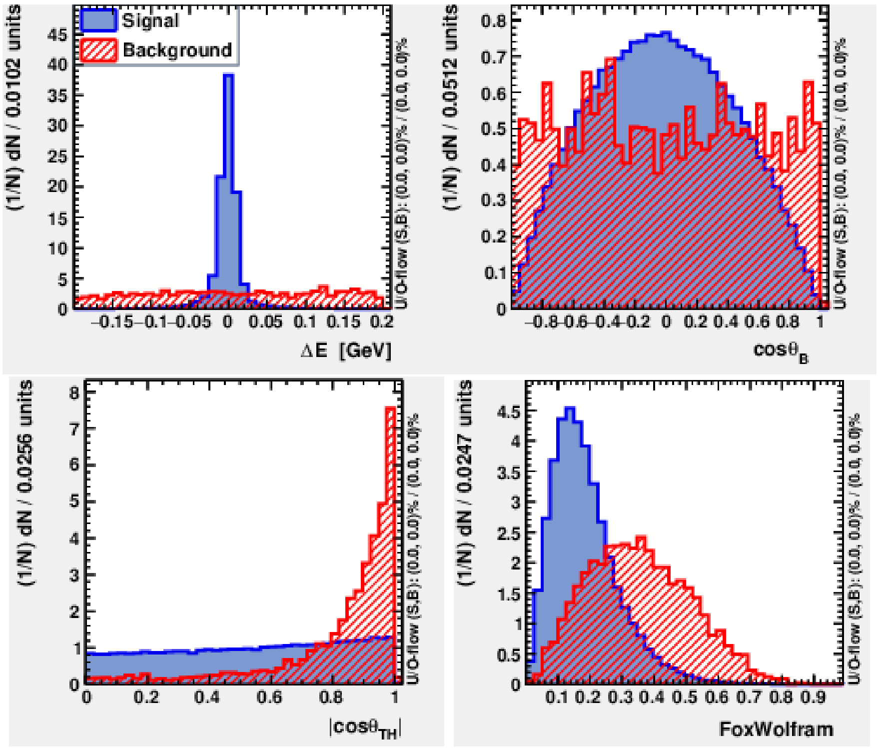


Figure 5.4: The distributions of the input variables used in the BDT selection are shown: from the top left corner, proceeding clockwise, the ΔE distribution and the three event shape variables discussed in Section 5.2. The plot showing the Fox-Wolfram distribution refers to the second normalized moment R_2. The signal (blue) and background (red) histograms refer to the MC samples; for the latter, the different types studied have been weighted according to their produuction cross sections in order to mimic the proper background composition expected in data.

The performed tests resulted in a selection optimized by the Boosted Decision Tree (BDT) method, trained on the four input variables ΔE, $cos\theta^*_B$, $|cos\theta_{TH}|$, R_2. Their distributions, separated between background (red) and signal (blue) events, are reported in Figure 5.4. The observed behaviour matches the expected shape as described in Section 5.2.1. The BDT selection has shown the best result when the preselection is already applied on the processed samples: removing the badly-recontructed candidates helps the method classify the background and signal events. The size of the training and testing samples is the same: half of the reconstructed signal and background MC events are used for the training phase and the half for testing and validation. During the training phase, the BDT learns from the sample the optimal splitting at each node on the input variables and it evaluates the BDT selection performance.

The better performance achieved by the BDT with respect to a selection based on sequential cuts on the same input variables is shown in Figure 5.5, where the background rejection $(1-\epsilon_{bkg})$ versus the signal efficiency (ϵ_{sig}) is plotted. The BDT achieves the highest background rejection for a given efficiency, resulting in a higher significance with respect to the selection based on sequential cuts on individual variables.

The set of input variables has been kept minimal, in order to limit the systematic uncertainty contribution coming from the BDT selection, which might be originated by differences in the shape of the input variable distributions between MC and data (Section 5.4).

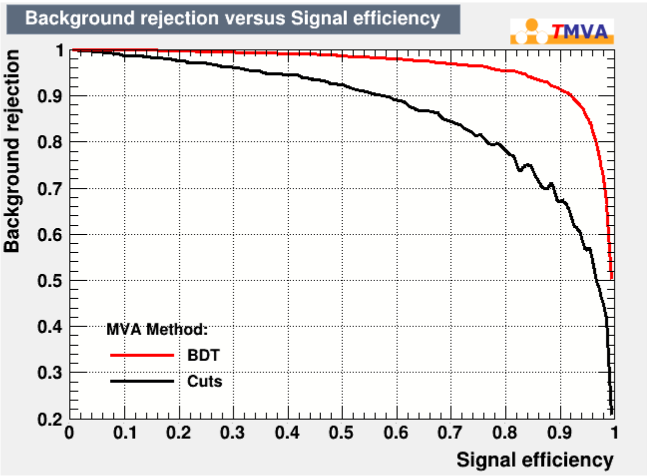


Figure 5.5: The background rejection ($1-\epsilon^{bkg}$) versus the signal efficiency (ϵ_{sig}) is plotted for the BDT method (red) and sequential cuts (black). As noted from the comparison of the red and black curve, the choice of the BDT selection is more desirable with respect to the sequential cut selection based on the same input variables, since at the same efficiency, the background rejection is higher for the BDT method.

For an efficiency and significance estimate, the number of expected signal and background events at data integrated luminosity in the signal region has to be provided. At this stage, after the preselection application, the MC-data comparison shows an excess in the total number of events in the background MC samples with respect to the number of events observed in the on-peak data side band region, $m_{ES} < 5.27$ GeV/c^2 , which is cured by scaling the sum of the MC samples in order to match the total amount of data in the side band region. The applied weight is:

$$w_{data/MC} = 0.447311 \tag{5.6}$$

The mis-matching might be explained by our ignorance about the precise mechanism of quark fragmentation and hadronisation into baryons in continuum processes $e^+e^-{\rightarrow}uds$. The JetSet generator in *BABAR* MC simulation relies on empiric tunable models of fragmentation, since the underlying theoretical models are not well known, which might cause the generation of more protons in continuum events than what observed in data. Evidence for the same issue has been observed also in other baryonic analyses at *BABAR*.

Applying the weight of Equation (5.6) to the total background events expected in the region $5.27 < m_{ES} < 5.29$ GeV/c^2, that passed the preselection (see last column of Table 5.1, $B = 1950$), the numbers for the BDT cut evaluation are given:

$$S^{beforeBDT} = 15.86, \qquad\qquad B^{beforeBDT} = 872. \tag{5.7}$$

From this configuration, on which the BDT selection should perform the background rejection, the optimal cut on the BDT output is evaluated. The efficiencies and the significance estimate for the above mentioned scenario are given in Table 5.3. Their behaviour as a function of the cut applied on the BDT output is reported in Figure 5.6 (left). The BDT response,

Optimal cut	$\frac{S}{\sqrt{S+B}}$	S	B	ϵ_{sig}^{BDT}	ϵ_{bkg}^{BDT}
0.1994	1.85492	9.579951	16.96447	0.604	0.01945

Table 5.3: The optimal cut on the BDT output, $BDT > 0.1994$, and the BDT performance evaluation estimated during the training phase. The expected significance ($\frac{S}{\sqrt{S+B}}$) and the number of the signal and background selected events are listed in the table. Finally, the BDT efficiency corresponding to the optimal cut is given, for both the signal and the background.

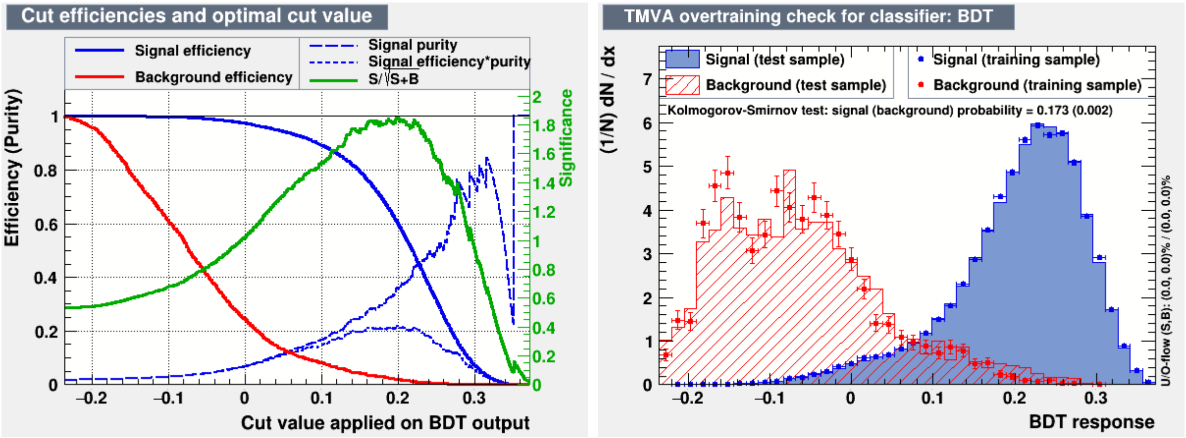


Figure 5.6: *Left*: the signal (blue line) and the background (red line) efficiency curves are plotted as function of the value of the BDT output cut applied; the purity (dashed blue line), defined as the ratio of the signal events over the sum of the signal and the background events, and the significance (green line) are also shown. The dotted-dashed blue line is the product of the signal efficiency times the purity. *Right*: the BDT response for the training (histograms) and testing (superimposed dots) processed samples, for both the signal (blue) and the background (red), is plotted in the picture.

for both the signal and the background events, is also shown in the same figure (right): the training and testing samples are compared in order to check for an *overtraining* of the method, which might happen when the statistical fluctuations of the specific training sample bias the BDT response. The good agreement of the training (solid histogram) and the testing (dots) distributions shows no evidence of overtraining.

The estimate for the significance expected after the BDT selection, $s = 1.85942$, shows a factor 2 improvement with respect to the result from the cut-based selection optimization. The optimized event selection is applied on both MC (testing sample only) and on-peak data side band region events and their consistency is assessed by the validation step.

5.4 The selection validation

The finalized event selection is summarized by:

- *preselection:*
  - 4 charged tracks satisfying *doca* requirements;
  - B vertex probability > 0.001;

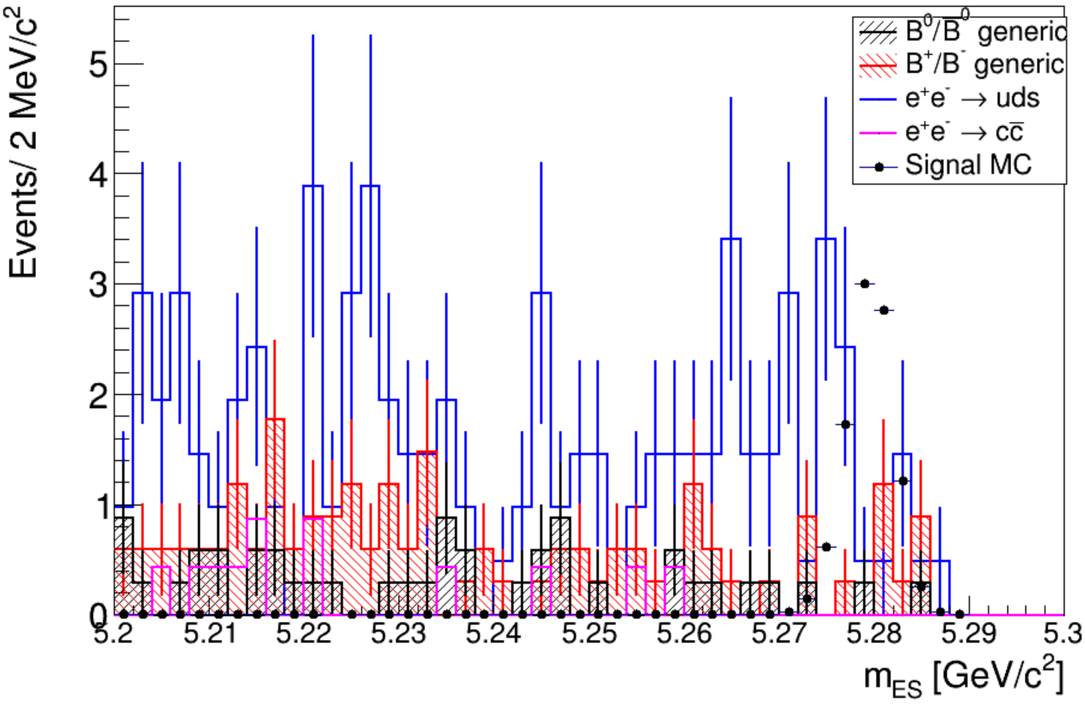


Figure 5.7: The weighted m_{ES} distributions for the analysed MC samples after the finalized event selection (preselection + BDT cut). All the histrograms have been scaled to the data integrated luminosity and the background samples corrected with the weight $w_{data/MC} = 0.443711$ for MC-data mis-matching. The signal MC sample is represented by the black dots and the signal region is highlighted in the plots by the vertical dashed lines. The error bars show the correctly scaled Poisson uncertainty for each bin.

  - 4 super loose protons from pKM selector;

- *BDT method:*
  - BDT output > 0.1994.

The weighted m_{ES} distributions, after the finalized event selection application to the signal and background MC samples, are plotted in Figure 5.7.

In Table 5.4, the event selection efficiency (ϵ) for each analysed sample is shown and the number of signal and background events, expected in the signal window at the data integrated luminosity and corrected for the MC-data mis-matching (background only), are listed in the last column.

The signal efficiency calculated after the BDT application is the final value that will enter the BF calculation. Since the truth-matching can not be exploited on data and since the efficiency can be evaluated only on the MC sample, the formula for ϵ_{sig} will not rely on the truth-matched candidates, but on the full number of selected events after the BDT cut application on the signal MC sample. The efficiency that will enter the BF formula is then given by:

$$\epsilon_{sig} = \frac{N_{sig}^{BDT}}{N_{gen}/2} = 0.2086 \pm 0.0007. \tag{5.8}$$

Sample	ϵ	After BDT$_{[s.r.]}$ (424 fb^{-1})
Signal $B^0(B^0) \rightarrow pp\bar{p}\bar{p}$	0.2086 ± 0.0007	9.80 ± 0.04
$B^0/\bar{B}^0$ to generic	$(1.1 \pm 0.2) \times 10^{-7}$	0.9 ± 0.5
B^+/B^- to generic	$(2.4 \pm 0.3) \times 10^{-7}$	3 ± 1
$e^+e^- \rightarrow uds$	$(1.8 \pm 0.2) \times 10^{-7}$	13 ± 2
$e^+e^- \rightarrow c\bar{c}$	$(0.21 \pm 0.06) \times 10^{-7}$	0
Bkg_{TOT}		17 ± 3

Table 5.4: The efficiency after the BDT selection are listed for each analysed MC sample. The number of expected events in the signal window at the data integrated luminosity, corrected for the MC-data mis-matching are provided. The associated errors refer to the proper scaled Poisson uncertainty which takes into account the applied weights.

where the $N_{gen}/2$ takes into account that the calculation is performed on the testing sample only. Among the selected signal events, after the BDT application, only the 0.2% are not truth-matched as signal in the MC generation and their contribution to the efficiency calculation is negligible. It is worth comparing the absolute efficiency calculated on the signal testing sample with the efficiency estimate provided by the BDT training phase in Table 5.3, which corresponds to the relative efficiency of the BDT method referred to the number of events provided for the training, which are half the signal events that passed the preselection. The conversion factor to apply is then the preselection efficiency $\epsilon^{presel} = 0.34$, evaluated from the Equation (5.2), and one can verify that the absolute efficiency obtained from the BDT training estimate is:

$$\epsilon_{sig}^{BDT} \cdot \epsilon^{presel} = 0.604 \cdot 0.34 = 0.21, \tag{5.9}$$

which shows a good agreement with the efficiency calculated from the testing sample after the BDT selection (Equation (5.8)). From the expected number of signal and total background events in the signal region, at the full luminosity of 424 fb^{-1}, the significance evaluation on the testing samples gives:

$$s = 1.89 \pm 0.09, \tag{5.10}$$

which corresponds to the expected significance from the BDT training estimate ($s = 1.86$).

5.4.1 The MC-data comparison

The last checks for the selection validation consists of two steps:

- comparing the MC sample and data distributions of all the input variables on which the BDT selection is trained;
- comparing the shape of the m_{ES} distributions before and after the BDT selection, for both MC and data.

In this section, only events in the side band region ($m_{ES} < 5.27$ GeV/c^2) for both MC and on-peak data are considered.

The MC samples and data distributions for the input variables of the BDT selection, before cutting on the BDT output, are reported in Figure 5.8. The preselection have already been applied. The MC distributions, weighted and corrected for the MC-data mis-matching, show a good agreement in the shape with respect to side band region data. The inset plot at the bottom of each distribution shows the ratio, bin per bin, of the number of observed data events, in the given variable, with respect to the weighted MC events in the same bin. If no significant difference in shape is observed, the ratio is expected to have a flat distribution centered within its uncertainty around $n_{data}/n_{MC} = 1$.

The largest fluctuations are observed in the tail of R_2 distribution (second Fox-Wolfram moment), but they are not affecting the BDT selection performance, since they are mainly in the region $R_2 > 0.5$. In Figure 5.4 comparing the signal and background distributions for R_2, it can be shown that most part of the signal is contained in the range $R_2 < 0.5$. Since the BDT method is trained on the events in the signal window only, it means that the background events entering the training are those belonging to the signal region of each studied variable and therefore the events with $R_2 > 0.5$ are almost totally excluded.

The same variables are checked also after the BDT selection, in order to confirm the same performance of the BDT method on MC samples and on real data. The input variable distributions after the BDT selection, for both the MC and data samples are given in Figure 5.9, where the usual scaling to the data luminosity and for the mis-matching correction is applied to the MC samples. The larger fluctuations shown in the plots are expected, given the reduced number of events. The MC-data shape agreement, within the statistical uncertainty of each bin content, is still satisfied, as shown in the inset plots. The m_{ES} distributions before (top) and after (centre) the BDT selection, for both MC and data, are plotted in Figure 5.10 and their shape does not show any evident change due to the BDT application. In the bottom plot, the response of the BDT method on the background MC testing samples and on data in the side band region is shown. The consistency between the BDT performance on MC and data is confirmed and the comparison to the expected BDT output from Figure 5.6 shows a good agreement with the training evaluation result.

As conclusion, the input variables chosen for the BDT method implementation ensure the optimal performance of the selection procedure and behave well both on MC and data samples, showing a good agreement for the resulting ditributions. The efficiency is estimated from the result of the BDT selection applied to the testing signal MC sample. The selected events are finally ready to undergo to the fitting procedure, treated in the next chapter.

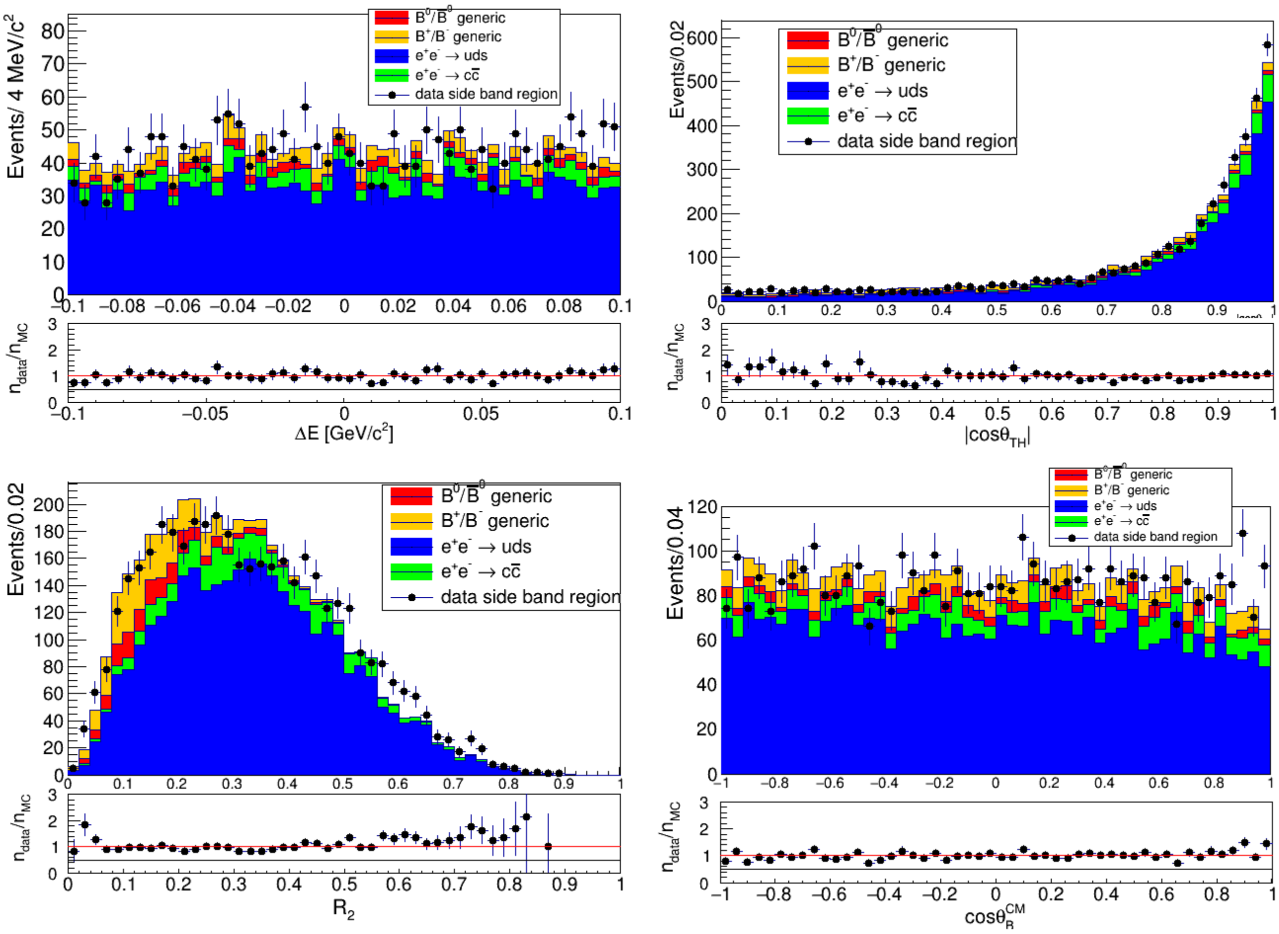


Figure 5.8: Comparison between on-peak data (dots) and MC (coloured stacked histrograms) in the side band region, before the BDT selection. In the inset plots, the ratio of the number of events in data with respect to the sum of all background MC events, for each bin of the plotted variable, is shown. A flat distribution around 1 is expected, if no significant shape difference is observed between data and MC.

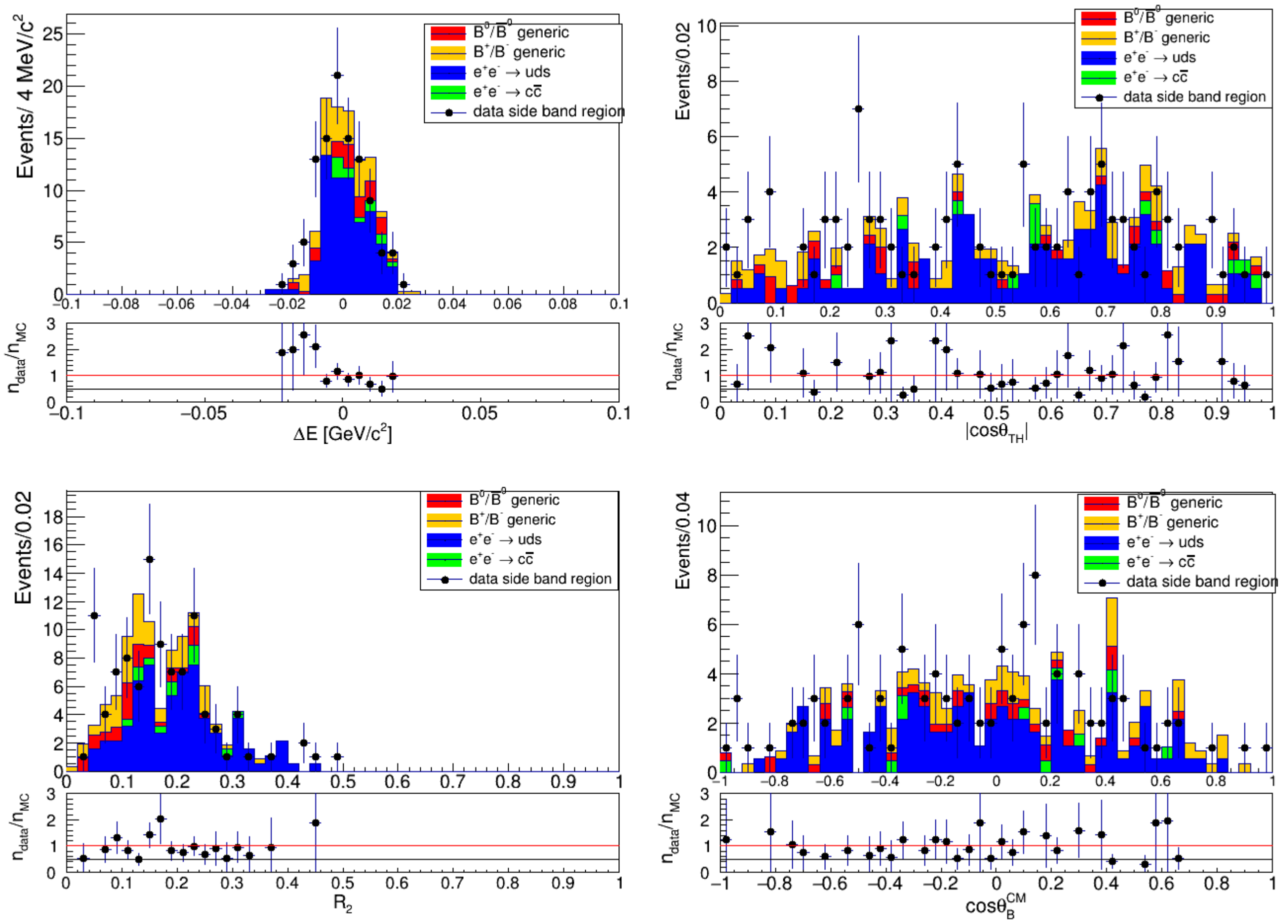


Figure 5.9: Comparison between on-peak data (dots) and MC (coloured stacked histrograms) in side band region, after the BDT selection. At the bottom of each plot, the ratio of the number of events in data with respect to the sum of all background MC events, for each bin of the plotted variable, is shown. As before, a flat distribution around 1 is expected in the case of no significant shape difference between data and MC.

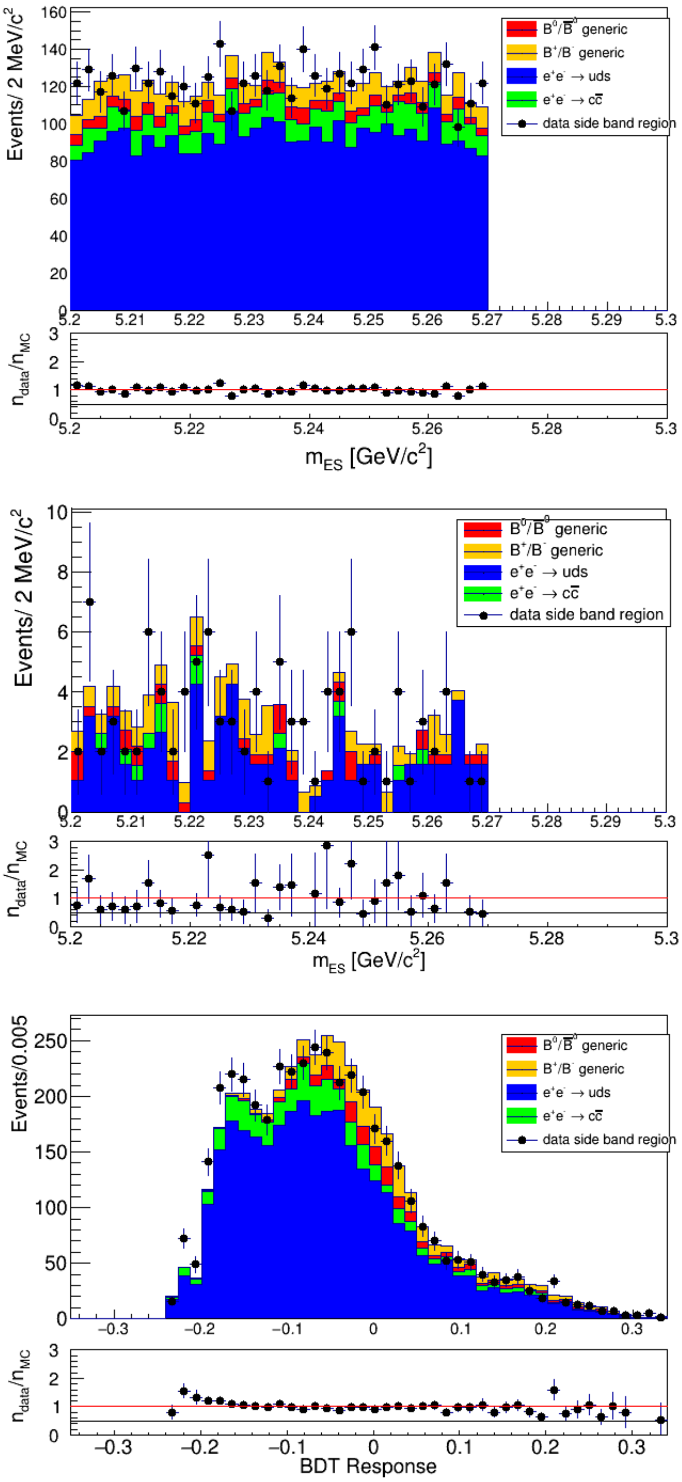


Figure 5.10: *Top and centre*: m_{ES} distributions for both MC (coloured stacked histograms) and side band region data (dots), respectively before and after the BDT selection. *Bottom*: the BDT method response evaluated on both backgrond MC samples (coloured stacked histograms) and side band region data (dots). In the inset plots, the ratio of the number of events in data with respect to the sum of all background MC events, for each bin of the plotted variable, is given. A flat distribution around 1 confirm no significant differences in shape between MC and data.

Chapter 6

Fitting Procedure

In this analysis, the signal yield will be extracted by performing an unbinned extended maximum likelihood fit [25] to the unblinded m_{ES} distribution of the on-peak data, after the BDT selection, in the whole reconstruction range which goes from 5.2 GeV/c^2 to 5.3 GeV/c^2. This technique has become popular for rare signal searches in High Energy Physics, since it correctly takes into account the Poisson error on the number of observed events. In order to apply this method, it is first necessary to define a probability density function (*pdf*) which associates at each event in the m_{ES} distribution the probability to occur. Since the observable to be fitted is a single variable, a one-dimensional *pdf* will be defined. The generic formula is given by $f(x;\vec{p})$, where the x variable refers to the fitted observable, m_{ES}, while the vector $\vec{p}$ indicates the parameters ensemble which the function depends on. In the specific case, we are interested in separating two components of the m_{ES} distribution, the signal and the background, which are known to be described by two different functions. Therefore the *pdf* for m_{ES} can be expressed as a linear combination of these two functions, where the coefficients are the signal and the background yields:

$$f(x;\vec{p},N_{sig},N_{bkg}) = \frac{1}{N_{sig}+N_{bkg}} \cdot (N_{sig} \cdot f_{sig}(x;\vec{p}_{sig}) + N_{bkg} \cdot f_{bkg}(x;\vec{p}_{bkg})) \tag{6.1}$$

Later on, the function described in Equation (6.1) will be referred to as the total *pdf.* The shape parameters for the signal component ($\vec{p}_{sig}$) are estimated on the signal MC sample, as explained in Section 6.1. For the estimate of the background shape parameters ($\vec{p}_{bkg}$) various possibilities are discussed and tested, as reported in Section 6.2. In the fit to unblinded data, the form of the total *pdf* is then completely defined and the values of both $\vec{p}_{sig}, \vec{p}_{bkg}$ are fixed. The only two floating parameters will be the signal and background yields (N_{sig}, N_{bkg}) and they will be determined by the extended maximum likelihood method. The unbinned extended maximum likelihood is given by:

$$L(N_{sig},N_{bkg}) = \frac{e^{-(N_{sig}+N_{bkg})}}{N!} \prod_{i=1}^{N} f(x;\vec{p},N_{sig},N_{bkg}) \tag{6.2}$$

where the total *pdf* is evaluated on each event i and the product is weighted by the Poisson probability for observing a total number of events $N = N_{sig} + N_{bkg}$. The fitted values for N_{sig}, N_{bkg} are those that maximise the likelihood L, or equivalently, minimise the expression:

$$-logL(N_{sig}, N_{bkg}) = -N_{sig} - N_{bkg} + \sum_{i}^{N} ln(f(x; \vec{p}, N_{sig}, N_{bkg})) \tag{6.3}$$

The logarithm of the likelihood is usually preferred in fitting algorithms, since it converts a product into a sum, saving computational effort.

Finally, toy MC studies have been arranged in order to test the impact of the fitting procedure on the signal yield extraction. Both the fitting procedure and the toy MC studies are implemented using the RooFit libraries (v3.60) [26], integrated in ROOT version 5.34/25 [23].

6.1 Signal modeling

The shape parameters for the signal *pdf* are estimated on the signal MC sample, after the BDT selection has been applied. Only a subset of 2183 events has been fitted, in order to help fit convergence. Too many data points require a very precise *pdf* to let the fit perform successfully, but only 10 events are expected after the BDT selection applied to on-peak data. Therefore, no high precision is necessary for this study, since large statistical fluctuations are expected to dominate over the systematic uncertainty on the signal yield coming from the shape parameters estimate. In light of this consideration, the following *pdf* candidates have been tested on the reduced data set:

1. a single Gaussian with mean m_0 and width σ,

$$singleGauss(x; m_0, \sigma) = N \cdot e^{-\frac{(x-m_0)^2}{2\sigma^2}} \tag{6.4}$$

2. the Crystal Ball Shape function [27] with parameters m_0, σ, α, n ,

$$CB(x; m_0, \sigma, \alpha, n) = N \cdot \begin{cases} e^{-\frac{(x-m_0)^2}{2\sigma^2}} & \text{for } \frac{x-m_0}{\sigma} > -\alpha \\ A \cdot (B - \frac{x-m_0}{\sigma})^{-n} & \text{for } \frac{x-m_0}{\sigma} \leq -\alpha \end{cases} \tag{6.5}$$

$$A(n, \alpha) = \left(\frac{n}{|\alpha|}\right)^n \cdot e^{-\frac{|\alpha|^2}{2}} \tag{6.6}$$

$$B(n, \alpha) = \frac{n}{|\alpha|} - |\alpha| \tag{6.7}$$

A, B are constants depending on the parameters n, α. In Figure 6.1 examples of the Crystal Ball shape function are plotted at varying of n and α. The difference with respect to an ordinary Gaussian is the power-law tail on the left side which models the side band of the distribution.

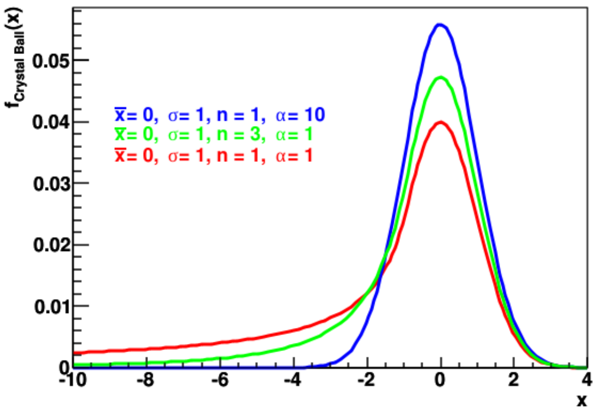


Figure 6.1: The Crystal Ball shape function. Various plots are shown for different values if n, α.

3. The sum of two Gaussians with the same mean m_0, but different widths σ_1, σ_2,

$$sumGauss(x; m_0, \sigma_1, \sigma_2, n) = n \cdot N_1 e^{-\frac{(x-m_0)^2}{2\sigma_1^2}} + (1-n) \cdot N_2 e^{-\frac{(x-m_0)^2}{2\sigma_2^2}} \tag{6.8}$$

   where n indicates the contribution of the Gaussian with width σ_1 to the total sum, $1-n$ is indeed the contribution of the Gaussian with width σ_2.

In all the three cases analysed, N, N_1, N_2 are normalization constants which are calculated from the *pdf* parameters. They ensure that the integral of the *pdf* over the reconstruction range $5.2 < m_{ES} < 5.3$ GeV/c^2 is equal to one.

Three fits to the reduced signal MC sample have been performed in the fit range $5.2 < m_{ES} < 5.3$ GeV/c^2, with floating shape parameters. The resulting plots and the pull distributions are shown in Figure 6.2. For every bin m of a binned data sample, pull values ($pull_m$) are defined as follows:

$$pull_m = \frac{N_m^{MC} - N_m^{fit}}{\sigma_m^{fit}} \tag{6.9}$$

where N_m^{MC} corresponds to the number of MC events in the bin m, while N_m^{fit} is the number of events predicted by the given *pdf* $f(x, \vec{p})$ at the centre of bin m. For an unbiased fit, in the hypothesis of N uncorrelated measurements of the variable x, the pull distribution is expected to be a Normal Gaussian, peaked at zero and with unit variance.

The goodness of each fit is also quantified and the $\chi^2 = \sum_m \left(\frac{N_m^{MC} - N_m^{fit}}{\sigma_m^{MC}}\right)^2$ for every tested *pdf* is reported in Table 6.1. In the last column, the associated probabilities for the calculated χ^2, given the number of degrees of freedom of each fit, are listed. For each *pdf* the number of degrees of freedom (dof) is defined as the difference $N_{bins} - n_{pars}$, where n_{pars} is the number of floating shape parameters of the fitting *pdf* and N_{bins} is referred to the number of bins in the signal window. The probability $P(\chi^2)$ is then used as figure of merit to optimize the choice of the signal *pdf*.

From these tests, the sum of the two Gaussians and the Crystal Ball function (*pdf* 2-3) showed the best agreement with the signal MC sample: they both describe in a more accurate

pdf	χ^2	N_{bins}	n_{pars}	$P(\chi^2, dof)$
singleGauss	54.52	50	3	0.2102
CB	31.53	50	5	0.9357
sumGauss	31.43	50	5	0.9375

Table 6.1: Tested *pdf* for the signal modelling: χ^2 , the number of bins, the number of parameters and the related probabilities are shown in the table.

Shape Parameter	Units	Initial Value	Fit Value
m_0	GeV/c^2	5.279	5.27952 ± 0.00005
σ_1	GeV/c^2	0.005	0.038 ± 0.028
σ_2	GeV/c^2	0.001	0.00252 ± 0.00004
n		0.1	0.0013 ± 0.0009

Table 6.2: Shape parameters are estimated from an unbinned maximum likelihood fit to signal MC sample, using as *pdf* the two Gaussians normalized sum (see Equation (6.8)).

way than the *singleGauss* the tails of the distribution, as their higher χ^2 probabilities indicate. A very low statistics is expected for the signal rate and since the precision of the measurement will be dominated by the statistical uncertainty, the impact of the signal *pdf* choice between the two selected candidates is not going to significantly influence the final yield extraction. For the same reason, the signal shape parameters will be fixed in the total *pdf* to the results of the MC modeling, since the low statistics cannot provide the sufficient sensitivity to simultaneously fit the shape and the signal yield.

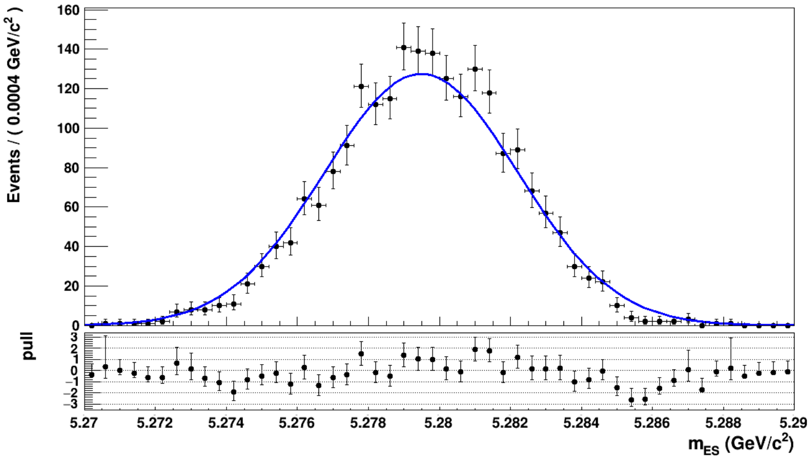
As a matter of simplicity, the sum of the two Gaussians is chosen to describe the signal component in the total *pdf*, which will be fitted to data to extract the signal yield. The shape parameters $(m_0, \sigma_1, \sigma_2, n)$ are estimated from the fit to the reduced signal MC sample, which is performed leaving all the parameters floating. The initial values of the parameters and the fitted values with the associated errors are shown in Table 6.2. The correlation matrix for the parameters of the fit is given in Appendix A.

6.2 Background description

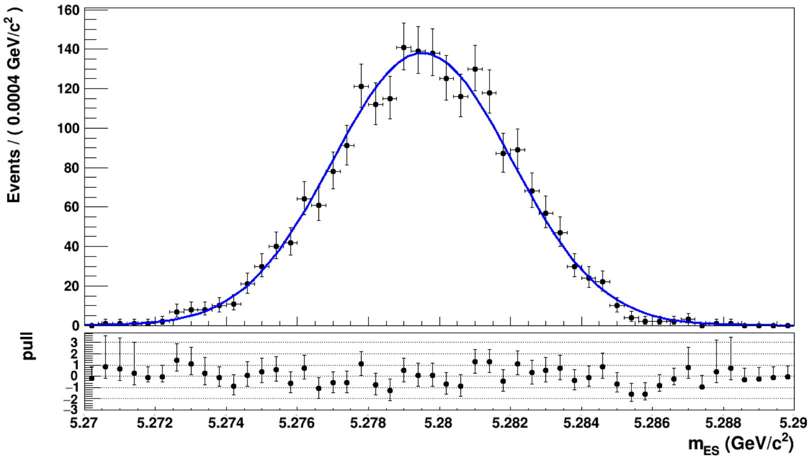
The contribution of the combinatorial background components to the invariant mass distribution of the reconstructed candidates for B meson decays can be modeled with a function originally developed by the Argus experiment [28]. The definition of the Argus function is the following:

$$f(x; a, b) = A \cdot x\sqrt{t(x)}e^{-\left(a^2/2 \cdot t(x)\right)}, \qquad t(x) = 1 - \frac{x^2}{b^2} \tag{6.10}$$

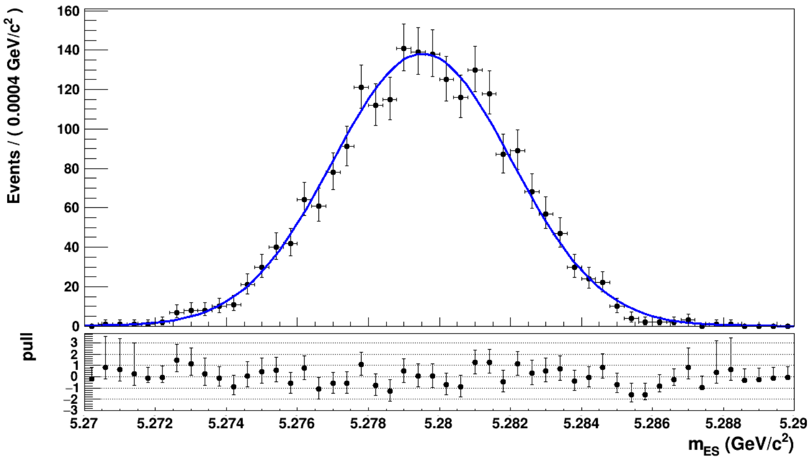
depending on two parameters, a and b. The reconstructed beam energy substituted mass (m_{ES}) corresponds to the variable x, which varies in the range $0 < x < b$. The parameter b is the cutoff value for x and it is completely determined by kinematics, while a is the curvature of the function, known as the Argus shape parameter. A is the normalizing constant calculated from



(a) *singleGauss*



(b) *CB*



(c) *sumGauss*

Figure 6.2: Tested *pdf* for the signal modeling: black dots show the reduced signal MC sample, the blue line is the fit performed. For each fit, in the inset plot the pull distribution is shown.

the parameter a.

Rather than testing different *pdf* as done for the signal, the modeling of the background is aimed at finding the best strategy for fitting the Argus shape parameter a and the cutoff b. According to the combination of two elements, that means the sample to fit and the tightness of selection applied before fitting, different strategies have been considered: for example, fitting data rather than the MC samples will prevent from studying the systematic uncertainty introduced by the simulation model and fitting before the BDT selection will gain in statistics. Every choice presents also some difficulties that have to be discussed. Given this consideration, four strategies have been considered and tested:

1. *Estimating the Argus shape parameter by fitting the background MC samples.*

   Possible differences in shape due to MC-data disagreement might require additional systematics studies. If the fit is performed before the BDT selection, the MC statistical power can be exploited, but it has to be shown that the BDT application is not changing the shape of the m_{ES} distribution.

2. *Estimating the shape from the off-peak data, after the BDT selection.*

   The advantage is that this sample is signal-free (off-peak data do not contain any real B decay since they were collected at a centre of mass energy 40 MeV below the invariant mass of the $\Upsilon(4S)$, below the threshold for $B\bar{B}$ production). On the other hand, a fit to less than 40 selected events is dominated by statistical uncertainty. In addition, the endpoint of the distribution is shifted as a consequence of the centre of mass energy shift and in order to have a fitted Argus shape consistent with the results from the other fits, b should be fixed.

3. *Fitting the Argus shape parameter on side band region of the on-peak data, after the BDT selection.*

   The fit result is not influenced by MC-data shape differences due to the decay and the fragmentation model in the simulation, which might introduce additional systematic uncertainties. Moreover, if a change in the shape caused by the BDT selection is expected, this is already included in the fit parameter estimate.

   Fitting the data after the BDT would require to blind the signal region and estimate the shape only from the side band of m_{ES}, in the range $m_{ES} < 5.27$ GeV/c^2, which might result in a systematic uncertainty in the shape estimate because of the exclusion of all the signal region. The cutoff parameter should be fixed and estimated from other fits, since the whole region sensitive to b, $5.27 < m_{ES} < 5.29$ GeV/c^2, is blinded.

4. *Performing a fit to on-peak data before the BDT selection.*

   The high statistics of the dataset before the final background rejection can be exploited to improve the shape sensitivity. The expected number of signal events before the BDT application, with the assumption of a branching fraction of order 10^{-7}, at the data integrated luminosity, is $S = 16$ (see Table 5.1), while for the sum of the background MC samples in the signal window $5.27 < m_{ES} < 5.29$ GeV/c^2 the number of expected events is $B = 872$, where the sum of background samples has been normalized to the total amount of background events selected in side band region of on-peak data, before the BDT application.

The hypothesis that the signal yield is negligible compared to the background yield, since $S< \sqrt{B} \sim 30$, is not enough to allow the implementation of this strategy for the Argus shape parameter estimate. In fact, it deals with a partial unblinding and the fit should be performed on the whole reconstruction range $5.2 < m_{ES} < 5.3$ GeV/c^2 before the BDT application, signal window included. Therefore, further validation studies, before implementing this procedure, are required.

6.2.1 The shape parameter estimate

Different shape estimate procedures related to the above discussed strategies 1, 2 and 3 have been tested. They are listed below and their initial configurations and results are summarized in Table 6.3. For all these tests the fitting *pdf* is the Argus function, as explained in the previous section. Tests are divided according to the samples used for the fit.

- Tests $(1-4)$ from strategy 1, fit to the background MC samples: both the selection configurations, before and after the BDT application, have been tested and for each configuration two fits have been performed, with all the parameters floating (Tests 1, 2) and with the cutoff b fixed (Test 3, 4).
- Tests $(5-7)$ from strategy 2, fit to the off-peak data: the off-peak sample before the BDT selection has been fitted only in the configuration with all the shape parameters floating (Test 5), while on the off-peak data after the BDT application two fits are performed, for testing the sensitivity of the shape parameter estimate to the cutoff;
- Tests $(8-9)$ from strategy 3, fit to the side band region of on-peak data: as discussed above, the cutoff b needs to be fixed; both the selection configurations, before and after the BDT application, have been tested but only the first fit strategy has converged.

List of Tests:

1. *all MC samples after BDT*: background MC, after the BDT selection, a and b floating;
2. *all MC samples before BDT*: background MC, before the BDT selection, a and b floating;
3. *all MC after BDT_{bfixed}*: background MC, after the BDT selection, a floating and b fixed;
4. *all MC samples before BDT_{bfixed}*: background MC, before the BDT selection, a floating and b fixed;
5. *off-peak Before BDT*: off-peak data, before BDT selection, a and b floating;
6. *off-peak After BDT*: off-peak data, after BDT selection, a and b floating;

7. *off-peak After* BDT_{bfixed}: off-peak data, after BDT selection, a floating and b fixed;
8. *on-peak Before BDT*: on-peak data, side band, before BDT selection, a floating and b fixed;
9. *on-peak After BDT*: on-peak data, side band, after BDT selection, a floating and b fixed;

Plots of the fits performed on MC samples are reported in Figure 6.3. In Figure 6.4 the fits to off-peak data and in Figure 6.5 to the on-peak data side band region are shown.

The option of fitting the Argus shape parameter a after the BDT selection on background MC samples (Test 1) should be discarded since it gives a relative uncertainty of around 50% and no improvement is gained by fixing the cutoff b of the Argus function (Test 3). The MC samples after the BDT application do not provide enough statistics and thus are not sensitive to a fit with the Argus shape parameter a left floating.

The cutoff b is estimated from the fit result of Tests 2 on MC samples before the BDT application, $b = 5.28959 \pm 0.00003$ GeV/c^2, which provides the most precise fitted value because of the large statistics available before the background rejection. The result is exploited in every test with fixed cutoff, except of course for the off-peak samples, which have a different endpoint. From Test 4, it is shown that fixing the cutoff b is not influencing the estimate of the Argus shape parameter a, which is consistent within the uncertainty with the result of Test 1.

From Tests $(5 - 7)$, the strategy of fitting off-peak data is discarded. In Test 6, since only 4 events have been found after the BDT application (assuming no weight is applied to scale the off-peak luminosity to the on-peak one), the fit result shows no sensitivity to the Argus shape estimate, with an uncertainty larger than 50%. Test 7 did not even converge, despite the several tests performed by varying intial values and fit ranges. Test 5 exploits the statistics available before the BDT selection and results in a fitted value for the Argus shape parameter a which is consistent with the estimate provided by the fit to the MC samples before the BDT (Test 2). Nonetheless, the off-peak estimate is less reliable because it might include a dependence of the shape on the shifted energy (40 MeV below the energy of the on-peak data) and furthermore it does not have enough statistics to be competitive with the MC estimate.

The results of Test 2 and Test 8 show consistency between the shape parameter estimates from the fit performed to the MC samples and to on-peak side band region data, before the BDT selection. In Test 8, the cutoff b is fixed to the value obtained from Test 2, since no sensitivity to the endpoint is possible due to the blinding. For Test 9, as already observed for off-peak data, the fit did not converge after the BDT selection.

The possibility to fit on-peak data, from side band tests results, seems to be very promising. However, in order to justify the partial unblinding for this procedure, further checks are required. The validation needs to show that, before the BDT selection, the signal yield expected in data is negligible compared to the background and that the background *pdf* properly fits the m_{ES} distribution of on-peak data before the background rejection and gives a correct estimate of the shape parameter a. For this investigation, specific toy Monte Carlo studies are needed. They will be discussed in the next Section, but they have not been implemented for this thesis.

Test	Sample	Selection	Parameters [units]	Initial Value	Fit result
1	MC	after BDT	a	-15	-15 ± 8
			[GeV/c^2] b	5.29	5.288 ± 0.001
2	MC	before BDT	a	-15	-25.9 ± 0.7
			[GeV/c^2] b	5.29	5.28959 ± 0.00003
3	MC	after BDT	a	-15	-12 ± 7
			[GeV/c^2] b fixed	5.28959	
4	MC	before BDT	a	-15	-24.9 ± 0.7
			[GeV/c^2] b fixed	5.28959	
5	off-peak	before BDT	a	-15	-23.8 ± 0.8
			[GeV/c^2] b	5.27	5.26920 ± 0.00006
6	off-peak	after BDT	a	-15	-21 ± 11
			[GeV/c^2] b	5.27	5.256 ± 0.001
7	off-peak	after BDT	a	-15	–
			[GeV/c^2] b fixed	5.26920	
8	on-peak	before BDT	a	-20	-25 ± 2
			[GeV/c^2] b fixed	5.28959	
9	on-peak	after BDT	a	-20	–
			[GeV/c^2] b fixed	5.28959	

Table 6.3: Results from the fitting strategies tested for the background shape estimate optimization.

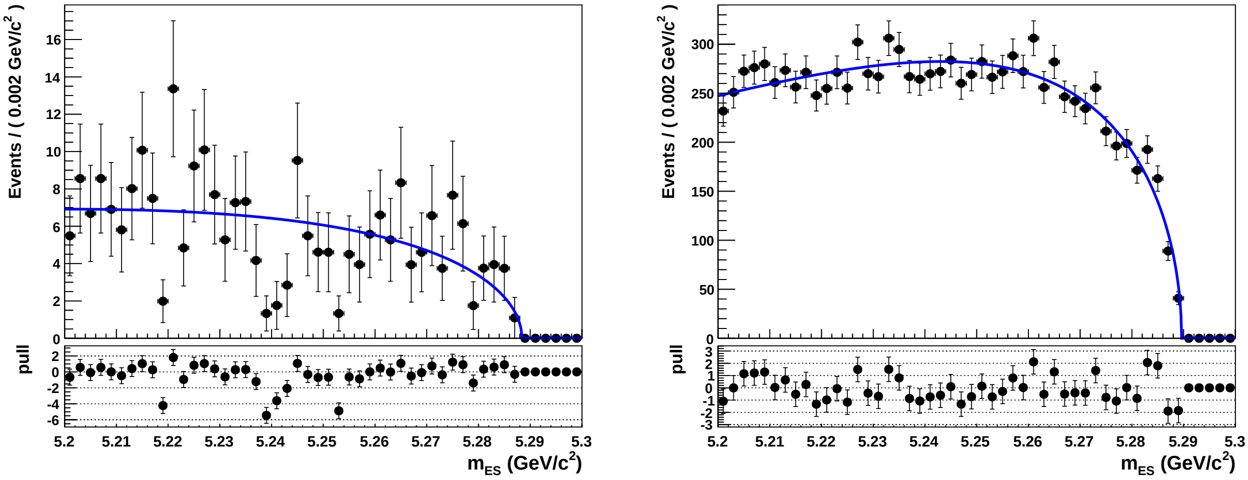
At the current status of the analysis, the performed tests allow to adopt the strategy of estimating the Argus shape parameters from the background MC samples, before the BDT selection. The values for a and b applied to the total *pdf* are the following:

$$a = -25.9 \pm 0.7 \ , \qquad b = 5.28959 \pm 0.00003 \ \mathrm{GeV/c^2} \tag{6.11}$$

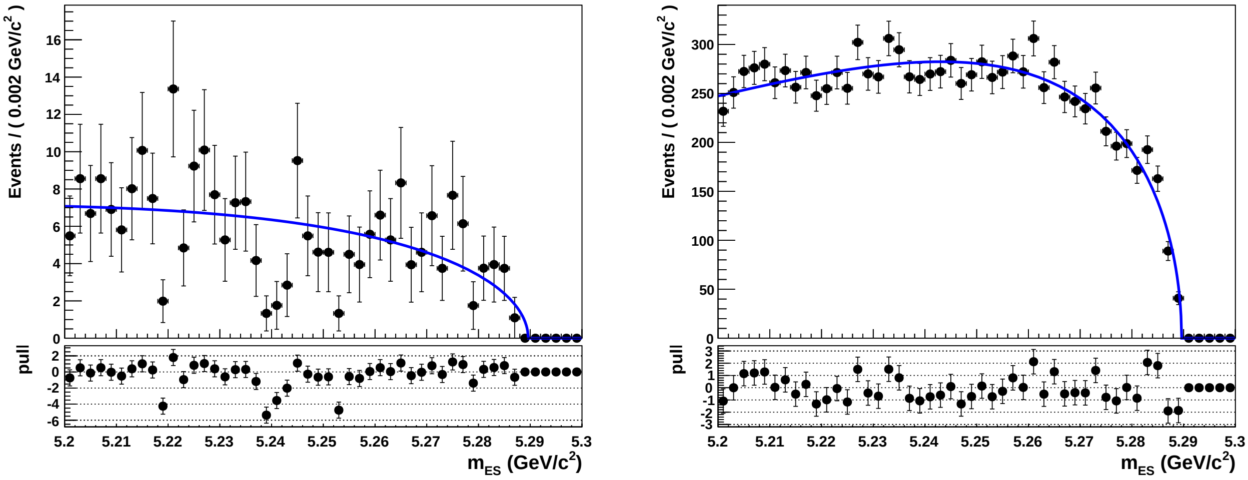
.

This is not the optimal fitting strategy, as explained before, and the systematic uncertainty eventually introduced by MC-data shape differences needs further investigation. Nevertheless, the fitted values will be updated as soon as validation for partially unblinding shows that the estimate of the background shape paramters from on-peak data is feasible.

The covariance matrices with the correlation coefficients resulting from the performed fits are given in Appendix A.

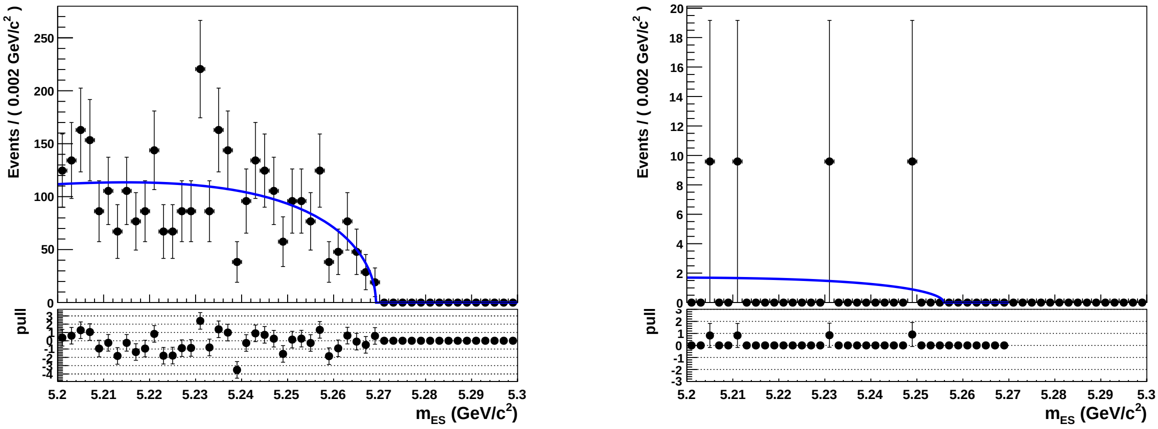


(a) Test 1 **(b)** Test 2

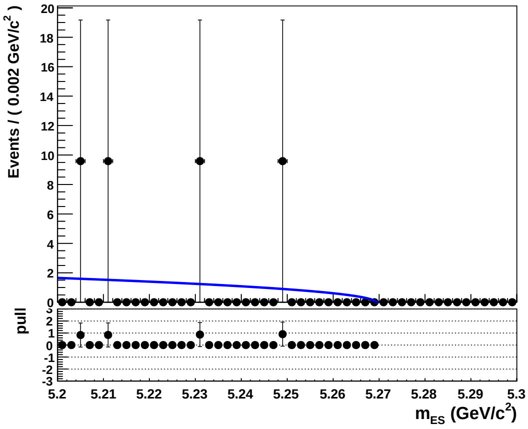


(c) Test 3 **(d)** Test 4

Figure 6.3: Fitting tests to background MC samples: black dots show the sum of all background MC samples, after (left) and before (right) the BDT selection. The contribution of each MC sample is weighted according to its production cross section and scaled to data integrated luminosity. The fit performed using the Argus function distribution is shown in blue line. The shape parameter a is always floating, the cutoff b is floating for fits shown in the top plots, while in the bottom plots is fixed. For each fit, in the inset plot the pull distribution is shown.

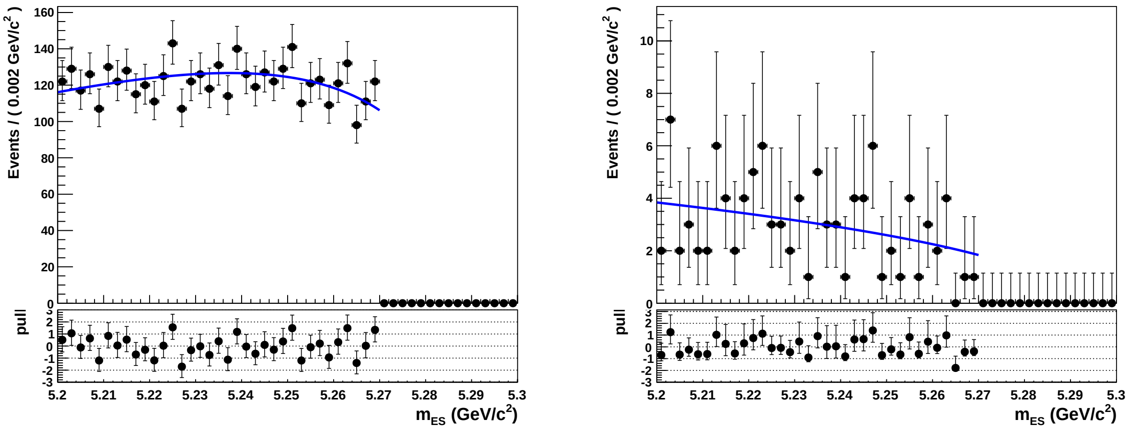


(a) Test 5 **(b)** Test 6



(c) Test 7

Figure 6.4: Fitting tests to off-peak data: black dots show the m_{ES} distribution for off-peak data, the blue line is the fit performed using Argus function. The shape parameters are always left floating in the fits, except for the bottom plot, where the cutoff b is fixed. For each fit, in the inset plot the pull distribution is shown. Off-Peak data are scaled to data integrated luminosity (weight applied $w = 9.58557$).



(a) Test 8 **(b)** Test 9

Figure 6.5: Fitting tests to on-peak data: black dots show the m_{ES} distribution in the side band region (m_{ES}< 5.27GeV/c^2) for on-peak data, before (left) and after (right) the BDT selection; the blue line is the projection of the fit performed using Argus function with the cutoff b fixed to the fitted value from *all MC samples before BDT*, a left floating. For each fit, in the inset plot the pull distribution is shown.

6.3 Toy Monte Carlo Studies

These studies are aimed to validate the fitting procedure and the shape estimate of the total *pdf*. Since this is a blind analysis, the signal yield is unknown and I want to study the bias on the extracted signal yield as a function of N_{sig}. In the RooFit framework, the implementation of toy MC studies for testing a fit procedure is already integrated. Once set up, this machinery automatically generates datasets from a given *pdf*, fits the datasets and shows plots for the mean of the fitted parameters, its error from the fit and the pull distribution. Two kinds of toy MC are needed:

1. Toy MC study for the validation of the signal yield extraction;

2. Toy MC study for the validation of the Argus shape estimate method.

In the first case, datasets are generated according to the total *pdf* and the same distribution is also used for fitting the generated datasets. Shape parameters are fixed to the default values estimated in the previous sections. The number of signal and background events are the only floating parameters in the fit. Their initial values are set according to the expected number of background and signal yields at data integrated luminosity, after the BDT application, in the whole reconstruction range $5.2 < m_{ES} < 5.3$ GeV/c^2, which means respectively:

$$N_{sig} = 10 \qquad\qquad N_{bkg} = 123 \tag{6.12}$$

where the signal yield is calculated on the assumption of an expected branching fraction of 10^{-7}. Then the generation routine has to mimic the extended maximum likelihoood method implemented in the fitting procedure, by assuming for each toy experiment a total number of observed events randomly drawn from a Poisson distribution with mean $\mu = N_{sig} + N_{bkg}$, where N_{sig} and N_{bkg} refer to the initial value setting. The generated events are then split into signal and background according to the the total *pdf*. From the distributions of the fit results for N_{sig} and N_{bkg}, it is possible to evaluate whether the fit procedure is biased and the errors are well estimated. This is verified by checking the pull distribution, which is expected to be a Normal Gaussian with mean $\mu = 0$ and width $\sigma = 1$. The study is repeated varying the initial value of N_{sig}, and therefore the total number of generated events, in order to estimate the impact on signal extraction and the bias to take into account for different signal yield configurations.

The second test has not been implemented for this thesis and it will be a further investigation to provide for the unblinding approval. It consists of toy MC studies aimed at justifying the procedure of the Argus shape parameter a estimate from on-peak data, before the BDT application. The generating *pdf* is the total *pdf* as before, but the generated datasets are fitted with the background *pdf* (the Argus function only), leaving the shape parameter a floating. Again, the pull distribution for the fitted variable will indicate whether the presence of the signal introduces a bias in the background shape estimate.

TEST	N_{sig}	σ_{sig}^{fit}	$mean_{pull}$	$bias$
1	10	4.54 ± 0.02	-0.13 ± 0.01	-0.6
2	0	3.11 ± 0.01	-0.21 ± 0.01	-0.7
3	30	6.48 ± 0.01	-0.08 ± 0.01	-0.5
4	$^{[N_{sig}>0]}\ 0$	3.34 ± 0.02	0.333 ± 0.006	$+1$

Table 6.4: Toy MC test results; the generated N_{bkg} is 123 for each test, while the generated N_{sig} varies according to the tested configuration. The fit mean uncertainty, the mean of the pull distribution and the bias introduced on N_{sig}^{fit} are shown. The last row shows the $N_{sig} > 0$ constraint effect on the fitting procedure and the introduced bias.

6.3.1 Validation of the signal yield extraction

In order to validate the fitting *pdf* and the procedure for extracting the signal yield, I set up toy MC tests. Each test generates 10000 toy experiments and for each experiment a number of signal and background events distributed according to the total *pdf*. Then the same *pdf* is used to fit the generated datasets and the resulting pull distributions for N_{sig}^{fit} and ${N_{bkg}}^{fit}$ give a feedback for the bias introduced by the fitting procedure. The initial value for N_{bkg} is always the same, N_{bkg}= 123, whereas for N_{sig} three configurations have been tested:

- Tests 1: default configuration with $N_{sig} = 10$, as expected;
- Test 2: no signal event detected, $N_{sig} = 0$;
- Test 3: the signal yield extracted is larger than the expected, $N_{sig} = 30$.

A last test (Test 4) is performed in order to show why the floating signal yield can not be constrained to be positive. The initial configuration is the same as Test 2, assuming no signal event has been detected, but the variation range for N_{sig} is now set to $(0, +\infty)$. As shown by the mean$_{pull}$, the restriction on N_{sig} introduces a positive bias on the fit result and the pull distribution is no more a Normal Gaussian.

The resulting plots from toy MC studies for Test$(1-3)$ are shown respectively in Figure 6.6, in Figure 6.7 and in Figure 6.8, while in Figure 6.9 the result of the constraint $N_{sig} > 0$ is shown. Pull results are summarized in Table 6.4, where the fit uncertainty (σ_{sig}^{fit}) and the mean of the pull distribution for N_{sig} are reported for each scenario. The expected bias on N_{sig} is:

$$bias_{sig} = \sigma_{sig}^{fit} \cdot mean_{pull}.$$

The fitted value for N_{bkg} are also reported in the pull box in the bottom plots in Figure 6.6, 6.7 and 6.8. Pull distributions, for both N_{sig} and N_{bkg}, are consistent with a Normal Gaussians of unit width. This implies that, as expected, the errors are symmetric and correctly estimated.

The maximum bias (-0.7) is expected in the case of $N_{sig} = 0$. The number of signal events will not be constrained to the range $(0, +\infty)$ in the fit, since it introduces a larger bias ($bias^{N_{sig}>0} = +1$) on the signal yield extraction, in the case no signal event is detected.

A rough estimate of the expected fit sensitivity for each tested configuration is evaluated from the ratio $N_{sig}/\sigma_{sig}^{fit}$. Toy MC study results (see Table 6.4) give a 2.2σ significance for Test 1, while for Test 3 more than a 4.5σ significance is achieved. It can be concluded that only with a detection of around 30 signal events or more a significant measurement[1] is expected, corresponding to a measured branching fraction (see Equation (7.4)) of $BF = (3.1 \pm 0.7) \times 10^{-7}$, where the associated uncertainty is only statistical. For signal yields below 30 events, an upper limit at a 90% confidence level on the expected BF is given (Section 7.1).

As conclusion, these studies validate the signal yield extraction strategy implemented with the total *pdf*. The fitting procedure is finalized fixing the shape parameters for both the signal and background components to the fit results on the MC samples. For the latter, the fit is performed before the BDT selection. The expected BF upper limits in different signal yield scenarios are analysed in the next chapter.

---

[1] A measurement providing a significance larger than 3σ, estimated from the ratio $N_{sig}/\sigma_{sig}^{fit}$. The fit error only refers to statistical uncertainty.

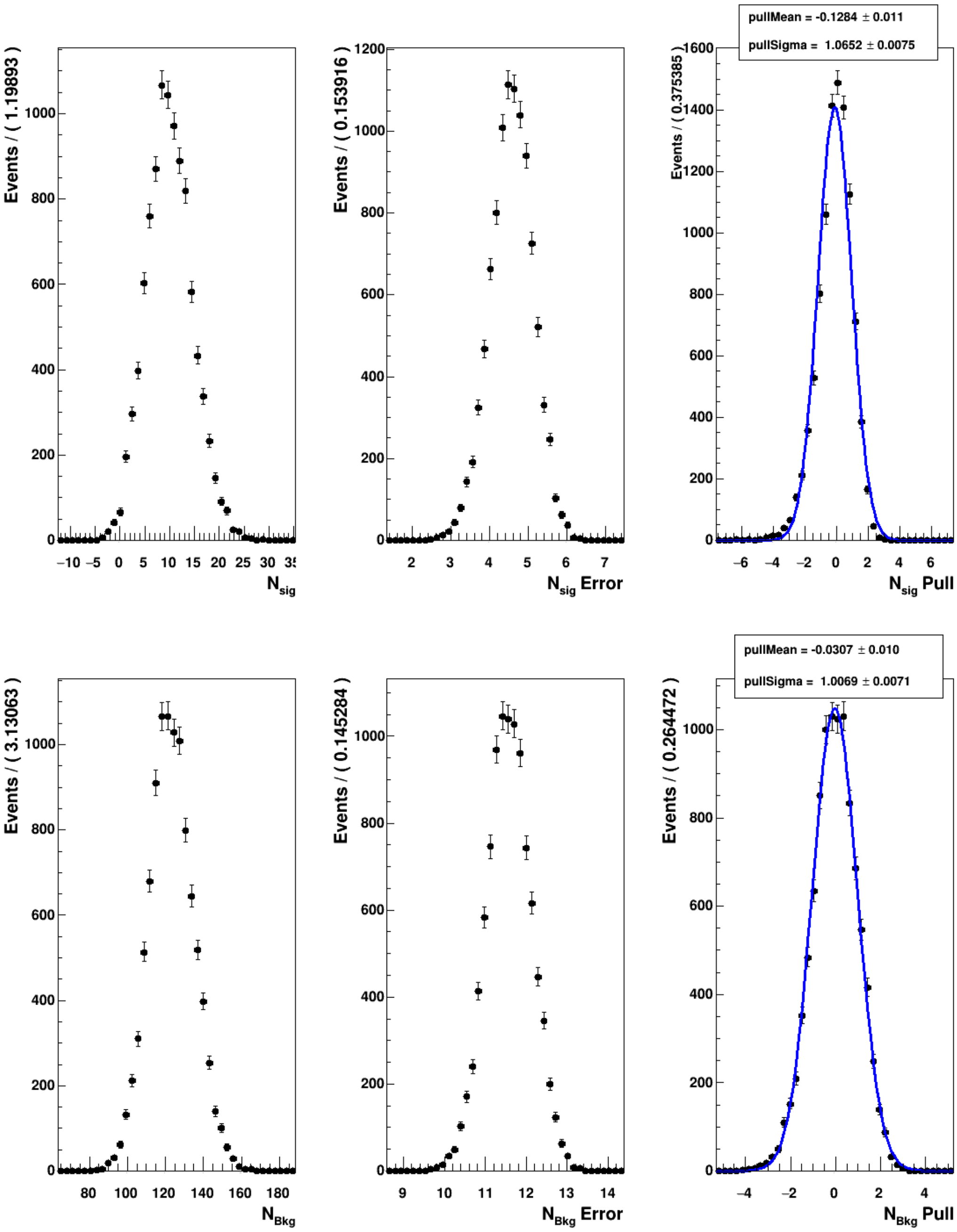


Figure 6.6: Test 1: for generation and fitting the total *pdf* is used, all shape parameters are fixed and the number of signal and background events are left floating. Their initial values are set to $N_{sig} = 10$, $N_{bkg} = 123$. From left to right, the first row shows the ${N_{sig}}^{fit}$, σ_{sig}^{fit} and pull distributions, while the second shows the same for the background.

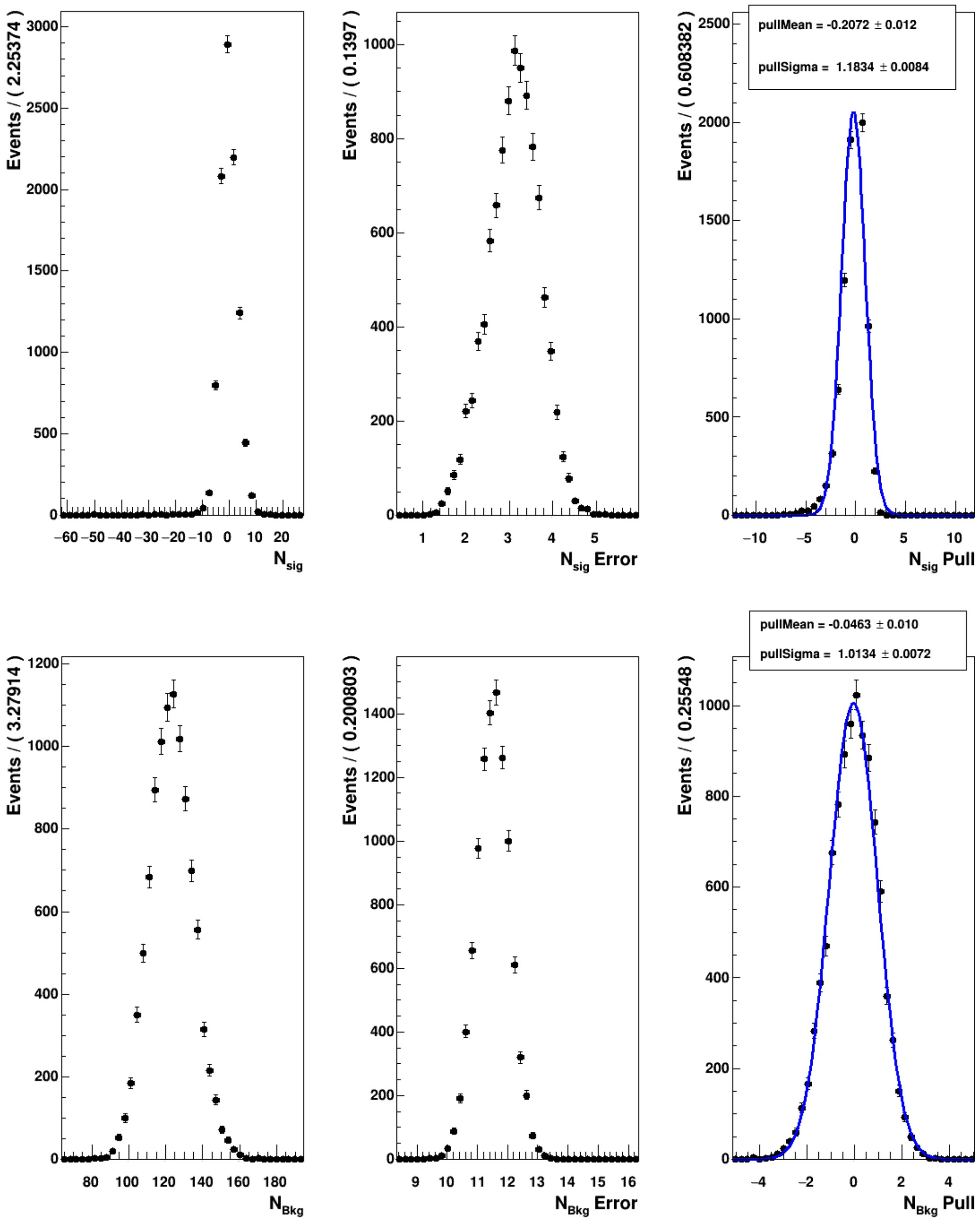


Figure 6.7: Test 2: for generation and fitting the total *pdf* is used, all shape parameters are fixed and the number of signal and background events are left floating. Their initial values are set to $N_{sig} = 0$, $N_{bkg} = 123$. From left to right, the first row shows the N^{fit}_{sig}, σ^{fit}_{sig} and pull distributions, while the second shows the same for the background.

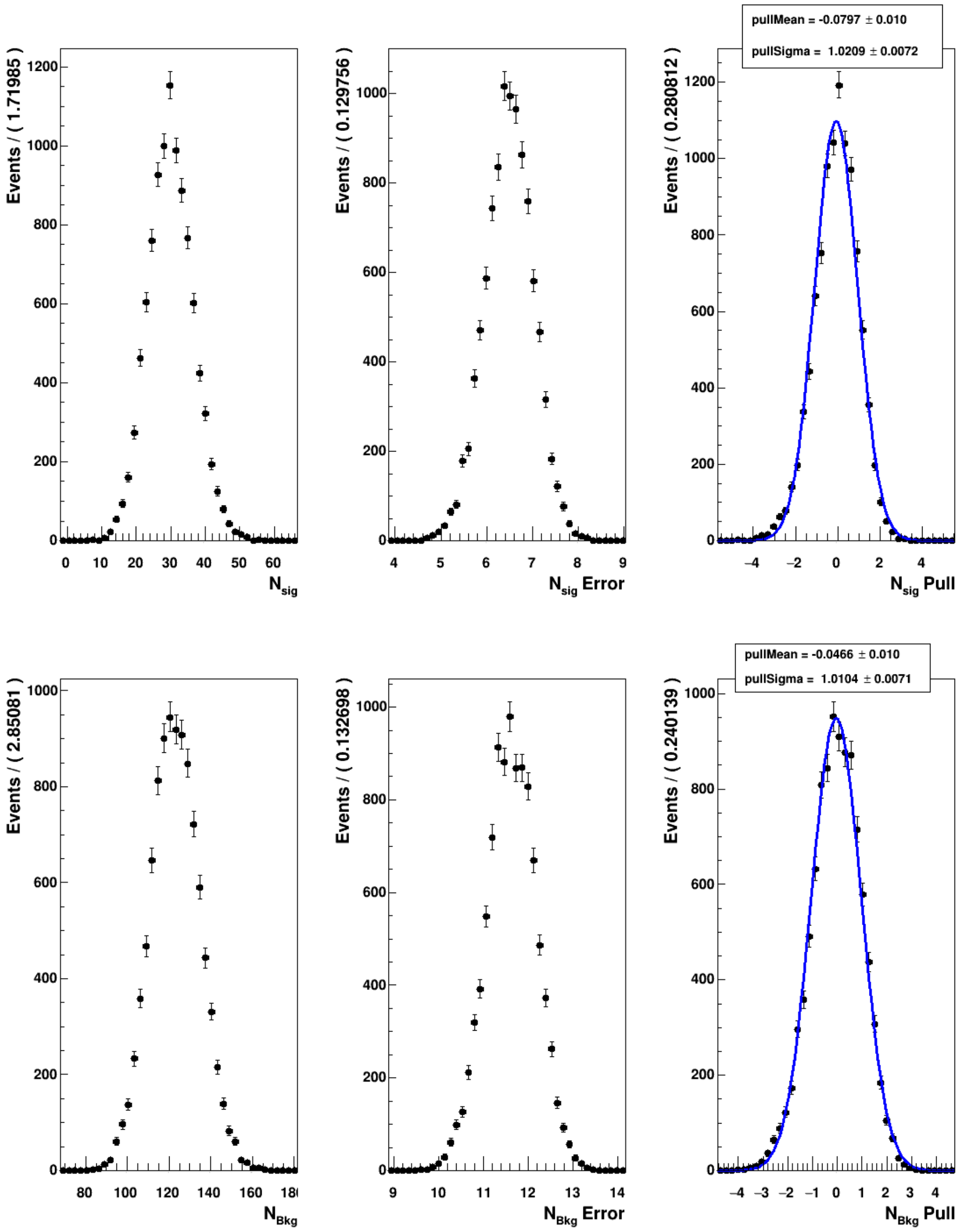


Figure 6.8: Test 3: for generation and fitting the total *pdf* is used, all shape parameters are fixed and the number of signal and background events are left floating. Their initial values are set to $N_{sig} = 30$, $N_{bkg} = 123$. From left to right, the first row shows the N_{sig}^{fit}, σ_{sig}^{fit} and pull distributions, while the second shows the same for the background.

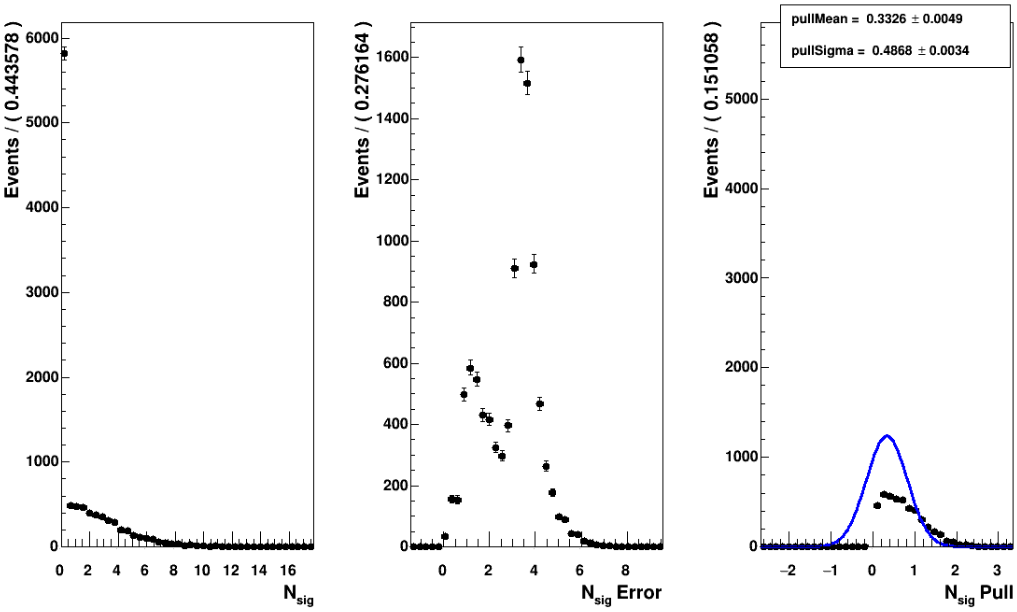


Figure 6.9: Test 4: for generation and fitting the total *pdf* is used, all shape parameters are fixed and the number of signal events is constrained to vary in the range $(0, +\infty)$. Initial values are set to $N_{sig} = 0, N_{bkg} = 123$ and from left to right, the N_{sig}^{fit}, σ_{sig}^{fit} and pull distribution are shown.

Chapter 7

Results and Future Prospects

Results from on-peak data will not be presented in this thesis, as it will be possible to look into the signal region only once the unblinding has been allowed. In this chapter, the application of the fitting procedure to the on-peak data, which will provide the number of observed signal events after the unblinding, and the procedure for calculating the measurement result will be explained. In Section 7.1, it is shown how to get the branching fraction from the fitted values and its associated statistical uncertainty is discussed. An overview of the expected contributions to the systematic uncertainty is given in Section 7.2 and the expected branching fraction upper limit, as a function of the number of observed events, are provided (Section 7.3). Finally, the last part of the chapter will introduce this measurement in the high luminosity scenario of the Belle II/SuperKEKB upgrade (Section 7.4).

7.1 The Branching Fraction estimation

Considering that the Branching Fraction $BF = \frac{\Gamma(B \to pp\bar{p}\bar{p})}{\Gamma_{tot}}$ is the probability of a B to decay in the specific final state $pp\bar{p}\bar{p}$, given the total number of neutral B mesons produced in $\Upsilon(4S)$ decays, $N_{B^0\bar{B}^0}$, the number of signal events is calculated by:

$$N_{sig} = BF(1-BF)\cdot N_{B^0\bar{B}^0} + (1-BF)BF\cdot N_{B^0\bar{B}^0} + BF\cdot BF\cdot N_{B^0\bar{B}^0} = \left(2BF - BF^2\right) N_{B^0\bar{B}^0}, \tag{7.1}$$

which is justified by the assumption that the total number of signal events, assuming a 100% signal efficiency, is given by the sum of the three cases:

1. the B^0 meson decays to the the signal final state but the $\bar{B}^0$ does not;

2. the B^0 meson does not decay to signal final state, the $\bar{B}^0$ does;

3. both the B mesons decay to the signal final state.

Since the assumption made for the expected branching fraction is $BF \sim 10^{-7}$, the contribution of the third case to the signal rate should be of order 10^{-14} and therefore, neglecting the second order term in BF, the branching fraction is:

$$BF = \frac{N_{sig}^{obs}}{\epsilon \cdot 2N_{B^0\bar{B}^0}}, \tag{7.2}$$

where N_{sig}^{obs}/ϵ stands for the true number of signal events N_{sig}, given by the ratio between the number of observed signal events and the signal selection efficiency. The number of neutral B mesons $N_{B^0\bar{B}^0}$ is not known and the measured quantity at the B Factories is the total number of B meson pairs ($N_{B\bar{B}}$). Taking into account the measured ratio $r = \frac{f^{+-}}{f^{00}} = 1.059 \pm 0.027$ [29] for the fraction of the charged B meson pairs with respect to the neutral ones in the $\Upsilon(4S)$ decay, imposing that $N_{B\bar{B}} = N_{B^0\bar{B}^0} + N_{B^+B^-}$, the number of neutral meson pairs can be written as:

$$N_{B^0\bar{B}^0} = \frac{1}{1+r} N_{B\bar{B}} = 0.486 \cdot N_{B\bar{B}}. \tag{7.3}$$

Replacing this expression for $N_{B^0\bar{B}^0}$ in Equation (7.2) the branching fraction formula becomes:

$$BF = \frac{N_{sig}^{obs}}{\epsilon \cdot 0.97 \cdot N_{B\bar{B}}} \tag{7.4}$$

where the factor 0.97, being calculated from the measured quantity r [29], has a propagated error, 0.97 ± 0.01, which is to take into account as a systematic uncertainty contribution of 1%.

Three input parameters are necessary to compute the BF: the signal yield, the selection efficiency and the number of $B\bar{B}$ pairs. The estimates for those parameters and their associated uncertainties are discussed in the remaining part of the section.

7.1.1 The signal yield extraction

After the shape parameter estimate (Section 6.2.1), the total *pdf* is set up to perform the unbinned extended maximum likelihood fit to the m_{ES} distribution of on-peak data, after the BDT selection. The fit will provide the background and signal yields, in the full range $5.2 < m_{ES} < 5.3$ GeV/c^2. All the shape parameters need to be fixed to the MC modeling results, since the low number of events expected after the finalized selection do not provide a good shape sensitivity in the same fit that extracts the signal yield.

Both the statistical and the systematic uncertainties have to be evaluated on the extracted signal yield. The expected statistical uncertainty, provided by the fit, is estimated from the toy MC tests (Table 6.4): with an expected signal yield of $N_{sig} = 10$, the average error provided by the fit is $\sigma_{sig}^{fit} = 4.5$, from which the relative statistical uncertainty on the extracted signal yield is

$$\frac{\sigma_{sig}^{fit}}{N_{sig}} = 4.5/10 = 0.45\ . \tag{7.5}$$

The significance is then around 2.2σ, which confirms the estimate from the BDT selection validation (Section 5.4). Toy MC results also give the estimation for the expected bias on the signal yield extraction, which in principle can be treated in two ways: by subtracting the bias to the signal yield or by assigning a further systematic uncertainty to account for it. For this analysis, the performed studies show that the bias can be negleted: even in the worst case, with no detected signal event, it would be a correction of order -0.7, neglible if compared to the expected fit uncertainty $\sigma_{sig}^{fit} = 3$.

The total relative uncertainty on the BF will be the sum in quadrature of each contribution (signal yield, efficiency, $N_{B\bar{B}}$): the fit result shows that, if only 10 events are seen, the precision of the measurement is statistically limited at 45%. The contribution of the systematic uncertainty is expected to be far lower and it is discussed in Section 7.2.

7.1.2 The selection efficiency

The selection efficiency on the signal can be evaluated only from MC studies. As described in Section 5.4, the final result for the signal efficiency is:

$$\epsilon_{sig} = 0.2086 \pm 0.0007. \tag{7.6}$$

where the uncertainty is statistical only. One advantage of calculating the efficiency on the MC samples is to exploit the full statistics of the total number of generated events to minimize the statistical uncertainty, coming from the limited size of the MC samples. From the binomial variance formula, the relative uncertainty on the efficiency is calculated as:

$$\frac{\sigma_\epsilon}{\epsilon} = \left(\sqrt{\frac{\epsilon \cdot (\epsilon - 1)}{N_{gen}}}\right) / \epsilon = 0.0033. \tag{7.7}$$

Further studies are needed for proper evaluation of the systematic uncertainty on the event selection efficiency, which is discussed in Section 7.2.

7.1.3 The number of $B\bar{B}$ pairs

The BF can be determined only if the number of $B\bar{B}$ events is known, which is achieved thanks to the B *counting* [22] method. This technique is based on the subtraction of the number of hadronic events in continuum interactions from the total number of hadronic events in the on-peak dataset:

$$N_{B\bar{B}} = \frac{(N_H - N_\mu \cdot R_{off} \cdot k)}{\epsilon_{B\bar{B}}}, \tag{7.8}$$

where N_H and N_μ are respectively the number of hadronic and dimuon events selected in on-peak data; $R_{off} = \frac{N_H^{off}}{N_\mu^{off}}$ is the ratio of selected hadronic events to dimoun events in off-peak data. Finally, k and $\epsilon_{B\bar{B}}$ take into account production cross-section corrections and selection efficiencies:

- $k = \left(\epsilon_\mu^{'} \sigma_\mu^{'} / \epsilon_\mu \sigma_\mu\right) \cdot \left(\Sigma_i \epsilon_i \sigma_i / \Sigma_i \epsilon_i^{'} \sigma_i^{'}\right)$ corrects for the cross-section energy dependence, at varying energy around the $\Upsilon(4S)$ threshold, with primed quantities referred to off-peak events. Since the muon pairs production cross-section only slightly varies with the CM energy, k is a factor close to 1;

- $\epsilon_{B\bar{B}} = 0.940$ is the efficiency for $B\bar{B}$ pairs to satisfy the hadronic event selection, assuming an equal branching ratio for $\Upsilon(4S) \to B^+B^- = \Upsilon(4S) \to B^0\bar{B}^0 = 0.5$ (possible differences between them are included in the systematic uncertainty).

The total systematic uncertainty on the full nominal dataset collected at $\Upsilon(4S)$ energy, $N_{B\bar{B}} = (471.0 \pm 2.8) \times 10^6$, is 0.6%, where the most relevant contribution comes from the efficiency for the hadronic selection and for low multiplicity event selection (dimuon events). In the total uncertainty it is also included the uncertainty in the modeling of the total energy distribution, exploited for the k factor evaluation.

7.2 The systematic uncertainties

A careful evaluation of the systematic uncertainty is to be provided before the unblinding and it is the most important next step for concluding this analysis. As for the relative statistical uncertainty, also the final relative systematic uncertainty on the BF (Equation (7.4)) will be the sum in quadrature of the three contributions from $N_{B\bar{B}}$, from the signal yield and from the selection efficiency. The systematic uncertainty on $N_{B\bar{B}}$ is known (Section 7.1.3) and it is 0.6% of the total σ^{syst}_{BF}/BF.

For the signal yield, the main sources of systematic uncertainty are the choice of the *pdf* for modeling the signal and background shape in the fit and the shape parameter estimate. The former can be investigated comparing the fit results from using alternative *pdf* models, for example, by replacing the sum of the two Gaussians in the signal modeling with a single Gaussian. The difference in the fit results between the different *pdfs* will give an estimate for the systematic uncertainty. Given the low statistics foreseen in data after the BDT selection, the signal yield extraction is not expected to be sensitive to this contribution, as already discussed for the choice of the signal *pdf* in Section 6.1. The shape estimate contribution to the systematic uncertainty can be evaluated by varying the shape parameters in the range defined by their associated fit uncertainties; the sensitivity to the shape parameters will be provided by comparing the fit results.

To assess the systematic uncertainty associated to the selection efficiency various contributions have to be taken into account:

1. the difference in particle identification performances between MC and data;
2. the BDT method uncertainty, related to the differences in the input variable distributions between MC and data;
3. the track finding efficiency;
4. the decay model used in the generation of the signal MC sample.

The differences in data and MC are evaluated on events from high purity channels, which are the control samples (CS) for a given particle type. The results are tabulated in the *BABAR* database and the PID efficiency plots for each type of particle are produced according to those measurements. It is worth noting that the PID performance contributes in two different ways to the systematic uncertainty on the selection efficiency, with the PID efficiency and with the mis-identification probability rate for a given particle type. The latter is sensitive to the background composition and, when there is a large fraction of mis-identified events, its contribution becomes relevant. In this analysis, having only 0.2% not truth-matched signal events after the BDT selection and assuming this number as the upper limit for the fraction of mis-identited events, the mis-identification contribution can be neglected. A way to evaluate the contribution from PID efficiency is by comparing the selection efficiency resulting from two configurations: when the PID efficiency on the MC samples is not weighted and after reweighting the MC PID efficiency with the weight $\frac{\epsilon_{CS,Data}}{\epsilon_{CS,MC}}$, resulting from PID performance evaluation on both MC and data control samples. From the PID efficiency plot for the proton selector used in the analysis, shown in Figure 5.1, for 1 GeV/c momentum tracks, the difference $\alpha = \epsilon_{MC} - \epsilon_{Data} \approx 0.7\%$ gives a rough estimate for the probability for not identifying in data a track identified as true proton in the MC sample. The probability for not correctly identifying only one proton track over four in real data events is:

$$\sigma_{notPID} = 4 \cdot \alpha(1-\alpha)^3 = 2.7\%, \tag{7.9}$$

which results in a systematic uncertainty contribution less than 3%. Of course, further studies should be performed to validate this preliminary estimate.

For the BDT method uncertainty evaluation, no indication can be drawn from Grünberg's analysis [10] for the $\bar{B}^0 \to \Lambda_c^+ p\bar{p}\bar{p}$ decay search. The relative systematic uncertainty on the selection efficiency was estimated to be 1.42% by comparing efficiency results from different selection strategies, *e.g.*, selecting the signal events in the signal region defined in the m_B vs m_{ES} rather than in the m_{ES} vs ΔE plane. This comparison accounts for the dependence of the efficiency on the cut variables. However, Grünberg's approach cannot be applied to this analysis both because a multivariate method, different from sequential cuts, has been implemented and because several input variables are used in the selection. One way for evaluating this uncertainty is to reweight each bin of the input variable distributions in the MC samples, according to the result of the MC-data comparison before the BDT application, and then evaluate the efficiency on the reweighted MC samples. Their difference gives an estimate of the sensitivity of the BDT selection to the MC-data shape differences. As shown in the BDT validation (Section 5.4), the difference in shape between MC and data is not large and a contribution of roughly the same order of the other uncertainty sources is expected. Studies will be performed to evaluate this uncertainty.

The track reconstruction systematic is related to the track finding efficiency at *BABAR* as explained in [30]. The systematic uncertainty contribution coming from track reconstruction varies as a function of the particle momentum and it is estimated to be of 1.54% per single track with transverse momentum lower than 180 MeV/c, while for high momentum particles it is 0.234% per track. Having four high momentum protons as signal signature and calling $\alpha = 0.00234$ the measured probability for loosing one track in the reconstruction, a rough estimate for the probability of not detecting one out of the 4 tracks in the signal event can be calculated as:

$$\sigma_{notTracked} = 4 \cdot \alpha(1-\alpha)^3 = 0.93\% \tag{7.10}$$

Repeating the calculation with the track inefficiency probability for low momenta, 1.54%, gives the conservative estimate of a systematic uncertainty around 6%, which can be considered as the upper bound for the systematic uncertainty coming from track reconstruction. However, given the mean proton momentum of 1 GeV/c for the simulated decay, the most likely estimate for track reconstruction uncertainty in this analysis is about 1%.

Finally, no theoretical model is so far known which describes the decay mode $B^0(\bar{B^0}) \rightarrow pp\bar{p}\bar{p}$ and no assumption was made on its dynamics. The employed decay model just assumes a flat momentum probability distribution for each proton in the phase space and it indirectly influences the efficiency by determining the kinematics of the signal events. Since I avoided to use strongly model-dependent variables in the selection, such as the proton momentum, the proton energy or the proton angular distribution, the systematic uncertainty coming from the decay model choice is expected to be small.

From the above considerations, statistics is expected to be the main source of uncertainty for this measurement.

7.3 The upper limit estimate

Fitting on-peak data will not be allowed before the final unblinding and the branching fraction can not be calculated yet.

In addition, the expected signal yield is most likely not enough to provide a significance higher then 3σ. The expected significance from the BDT validation studies (Section 5.4) is around 2σ. Therefore an upper limit estimate for the BF will be provided. Two different approaches, Bayesian and frequentist, have been implemented to get an estimate of the expected upper limit (UL) as a function of the number of observed event (n) in the signal region [31]. The estimate is based on the following hyphotesis:

- it is a counting experiment where events are assumed to be distributed according to the Poisson statistics;
- B is the number of expected background events in the signal region $5.27 < m_{ES} < 5.29$ GeV/c^2;
- n is the number of observed events in the signal region.

In the Bayesian case, the upper limit for the number of signal events S_{up}, once the credibility for the interval (CI) is fixed, can be extracted by the formula [31]:

$$\frac{\int_0^{S_{up}} \left[(S+B)^n e^{-(S+B)}/n!\right] P(S)dS}{\int_0^{\infty} \left[(S+B)^n e^{-(S+B)}/n!\right] P(S)dS} = CI \tag{7.11}$$

Under the assumption of a flat prior $P(S)$ for a signal yield $S > 0$, equation (7.11) can be simplified to the ratio:

$$\frac{\sum_{m=0}^{n}\left[(S_{up}+B)^m e^{-(S_{up}+B)}/m!\right]}{\sum_{m=0}^{n}\left[(B)^m e^{-B)}/m!\right]} = 1 - CI \tag{7.12}$$

where S_{up} is interpreted as the upper bound on the model parameter range for which the probability of getting n or less observed events is equal or lower than $1 - CI$.

The frequentist formula [31] for the S_{up} estimate differs from the Bayesian formula just for the absence of the denominator:

$$\sum_{m=0}^{n}\left[(S_{up}+B)^m e^{-(S_{up}+B)}/m!\right] = 1 - CL. \tag{7.13}$$

In Equation (7.12) the denominator takes into account that the number of background events is always less or equal to the number of observed events, while this constraint is no more implemented in the frequentist approach (Equation (7.13)), resulting in a possible difference between the two upper limit estimates, especially in the case of low signal events. Both approaches are implemented for estimating S_{up}, in different configurations of n. The Bayesian method resulted in a more conservative estimate.

Fixed at 90% the credibility interval CI (Bayesian approach) or the confidence level CL (in the frequentist hypothesis) for the expected upper bound S_{up}, three scenarios have been studied. The expected number of background events in the signal region (B) of on-peak data cannot be accessed because of the blinding. From the BDT selection results on both MC samples and on-peak side band region data, a good estimate is provided by the number of MC events found in the signal window, scaled with the weight $w_{data/MC} = N_{data}^{side-band}/N_{MC}^{side-band}$. With this technique (see also Table 5.4) the number of background events in the signal region is found to be $B = 17$. The number of observed events n changes according to the hypothesis made for the signal yield:

- *Scenario 1:* $B = 17, n = 17$; it is assumed that no signal event is detected;
- *Scenario 2:* $B = 17, n = 27$; in the assumption of detecting the expected number of signal events after the BDT selection, S=10;
- *Scenario 3:* $B = 17, n = 47$; three times the expected number of signal event after the BDT selection is detected, S= 30. This last hypothesis aims at predicting the resulting upper limit on the BF in the case the signal yield is much larger than the expected.

The results for the above mentioned configurations, with both statistical approaches, are summarized in Table 7.1. From the S_{up} estimate, using Equation (7.4), it is possible to compute the branching fraction upper limit expected for each scenario and the resulting upper limit estimates for the BF($B^0(\bar{B^0})\rightarrow pp\bar{p}\bar{p}$), at a 90% confidence level, are reported in the last row of Table 7.1.

Approach	S_{up}		
	Scenario 1	*Scenario 2*	*Scenario 3*
Bayes	8.18	17.99	40.06
Frequentist	6.60	17.95	40.06
$BF^{UL}(B^0(\bar{B}^0)\to pp\bar{p}\bar{p})$ $[Bayes/Frequentist]$	$[\times 10^{-7}]$ 0.9/0.7	$[\times 10^{-7}]$ 1.9/1.9	$[\times 10^{-7}]$ 4.3

Table 7.1: Upper limit estimates for the signal yield at three different scenarios: no detected signal event (*Scenario 1*), a signal yield corresponding to the espected $N_{sig} = 10$ (*Scenario 2*) and a larger number of detected signal events $N_{sig} = 30$ (*Scenario 3*). The S_{up} estimate is used in the BF formula (7.4) for calculating the expected upper limit.

With my analysis, I am able to provide the first measured upper limit on the decay mode $B^0(\bar{B}^0)\to pp\bar{p}\bar{p}$ of 1.9×10^{-7} at a 90% confidence level, for the case of an extracted signal yield $N_{sig} = 10$.

7.4 Prospects at Belle II

The measurement described in this analysis is mainly dominated by the statistical uncertainty. The same measurement at Belle II will reduce this uncertainty because of its almost 100 times larger integrated luminosity with respect to the *BABAR* full dataset. The improvement foreseen at the SuperKEKB nominal integrated luminosity, at the end of its data taking period, is briefly discussed in this section, which provides a summary of the main issues that the extrapolation needs to consider in further investigation.

Let's assume two possible scenarios, called A and B.

Scenario A

The expected performance of the implemented analysis is assumed to be the same for *BABAR* as for Belle II. The main relative uncertainty on the BF is then given by:

$$\left(\frac{\sigma_{BF}}{BF}\right)^2 = \left(\frac{\sigma_{sig}}{N_{sig}}\right)^2 = \frac{k}{L}, \tag{7.14}$$

where L stands for the integrated luminosity and k is a factor depending on the signal selection efficiency and therefore on the detector performance. In fact, assuming a Poisson statistics and expliciting the signal yield as the efficiency times the production cross section for that type of events (σ^{prod}) multiplied by the integrated luminosity L, the square relative uncertainty on the signal yield is

$$\left(\frac{\sigma_{sig}}{N_{sig}}\right)^2 = \frac{1}{N_{sig}} = \frac{1}{\epsilon_{sig} \cdot \sigma^{prod} \cdot L} \tag{7.15}$$

which compared to Equation (7.14), gives $k = 1/(\sigma^{prod} \cdot \epsilon_{sig})$. The physics described in σ^{prod} does not change for different detectors. The hypothesis made for this scenario implies that

$\epsilon_{sig}^{BaBar} = \epsilon_{sig}^{BelleII}$ and therefore k is eliminated in the ratio between the statistical relative uncertainty on the BF measured by the two experiments, which then goes like the square root of the ratio of the integrated luminosities:

$$\sigma_{BF}^{BelleII}/\sigma_{BF}^{BaBar} = \sqrt{\frac{L_{BaBar}}{L_{BelleII}}} = \sqrt{\frac{0.424 \text{ ab}^{-1}}{50 \text{ ab}^{-1}}} = 0.092 \tag{7.16}$$

Using the nominal value for the foreseen full integrated luminosity at Belle II, a factor of around 10 is gained on the statistical precision of the measurement, passing from roughly 45% to 4.5% for the case of 10 detected signal events.

Scenario B

The second scenario is a more realistic case in which the BDT selection performance might present some differences beween *BABAR* and Belle II. In that case, when comparing the relative statistical uncertainties,

$$\sigma_{BF}^{BelleII}/\sigma_{BF}^{BaBar} = \sqrt{\frac{k'}{k} \cdot \frac{L_{BaBar}}{L_{BelleII}}} = 0.092 \cdot \sqrt{k'/k}, \tag{7.17}$$

the ratio $\sqrt{k'/k}$ goes as $\sqrt{\epsilon_{sig}^{BaBar}/\epsilon_{sig}^{BelleII}}$, and in the assumption that the Belle II detector upgrade (Sections 2.4 and 3) will provide at least an equal or better performance with respect to the *BABAR* detector, this ratio is equal or lower than 1 and it will further improve the statistical precision of the measurement. A rough evaluation of the factor $\sqrt{k'/k}$ can be estimated by considering the dependence of the efficiency on three main detector-related features:

- the beam energies and the resulting boost in the laboratory frame;
- the angular acceptance of the detector;
- the PID efficiencies for protons at 1 GeV/c momentum.

At Belle II the reduced asymmetry in the beam energies will result in a reduced Lorentz boost $\beta\gamma = 0.28$ of the laboratory frame, roughly half the boost of *BABAR*, which will provide a lower shift for the momentum distributions of the B decay products. This analysis is not exploiting the shape of the momentum spectra of the protons and moreover, in the case of 1 GeV/c momentum track, the shift will be negligible.

The Belle II detector will have an extended azimuthal coverage compared to the *BABAR* detector. Assuming the same geometrical acceptance for Belle II as for *BABAR* is then a conservative estimate applied in this extrapolation.

Finally, the least obvious part of the extrapolation regards the PID efficiency comparison. The two detectors, as illustrated in Sections 2.3 and 2.4, exploit different devices to provide PID information. Moreover, very different PID selectors were used at *BABAR* (Section 5.1) and Belle for particle identification. Belle II inherits Belle Likelihood-based PID selectors and their performance in the high luminosity environment of SuperKEKB is still under investigation.

For protons at 1 GeV/c momentum, the MC preliminary studies on Belle II PID performance give an efficiency (mis-identification probability) in proton/pion discrimination of 0.96 (0.02) and in the proton/kaon case 0.94 (0.02). These are not the final numbers and further improvement is expected for the PID performance at Belle II.

Also the factor 2 improvement on the impact parameter resolution expected at Belle II has to be taken into account in the comparison of the track reconstruction efficiency for the two experiments (See the end of Section 3.6). A better vertexing resolution will be provided by the improved quality of the SVD and the measured B meson vertex distance from the interaction point will result in a more efficient constraint on the combination of four charged tracks coming from the same vertex to form a B meson candidate. Preliminary studies have shown that requirements on the distance of closest approach are important for eliminating fake candidates, therefore further studies will be desirable for better understanding this impact.

Of course, more detailed studies have to be arranged for giving a reliable estimate of the ratio $\sqrt{k'/k}$, which is expected to be of order of 1, but this can be achieved only with a more definitive PID efficiency evaluation for Belle II performance and only once the systematic uncertainty studies on the *BABAR* measurement are complete.

Summary and Conclusions

In this thesis, the analysis method for the measurement of the branching fraction of the decay mode $B^0(\bar{B}^0) \to pp\bar{p}\bar{p}$ at *BABAR* was developed. The event reconstruction and selection have been optimized on the official MC samples and validated on the side band region of on-peak data, without looking at the signal region, since it is a blind analysis. Kinematical constraints and cuts on the quality of the reconstructed tracks and on the probaility of the fit to the common vertex have been exploited to reject wrong combinations of tracks. The finalized event selection relies on the BDT method which mainly exploits event shape variables to suppress backgrounds from continuum events. After the final selection is applied on the testing samples, consisting in half the generated MC samples, the efficiency of the event selection is evaluated on the testing signal MC sample to be 21% and the BDT performance is validated by comparing the results on the background MC samples and on data in the side band region. The probability density functions used for the signal and background modeling are studied on the testing MC samples and dedicated toy MC studies evaluate the fit sensitivity in the configuration of various extracted signal yields and also the bias introduced by the fitting procedure.

From the tested values in toy MC studies, only in the case of 30 detected signal events (or more) a significant measurement is expected, corresponding to a branching fraction of $BF = (3.1 \pm 0.7) \times 10^{-7}$, where only the statistical uncertainty is provided. In the default hypothesis of a branching fraction of 10^{-7} and 10 detected signal events, a sensitivity of 2.2σ is expected, which is not enough to have a significant measurement. The corresponding upper limit on the branching fraction, estimated from MC studies (Table 7.1), results in 1.9×10^{-7} at a 90% confidence level.

Before the unblinding, further studies are needed for estimating the systematic uncertainty, which is expected to be of order of a few percent. A proper evaluation of the BDT method contribution to the total systematic uncertainty, through the error associated to the selection efficiency, will be studied by reweighting the input variable distributions used for the BDT training.

After the approval for the unblinding, the on-peak data will be fitted with the unbinned extended maximum likelihood fit implemented in this thesis and from the extracted signal yield, the first measurement of the branching fraction upper limit for the baryonic mode $B^0(\bar{B}^0) \to pp\bar{p}\bar{p}$ will be provided. Moreover, in the case of more detected signal events than the expected 10 events, toy MC tests show that a good sensitivity can be achieved and in that scenario it would be possible to calculate the central value for the branching fraction with a significance of

$BF/\sigma_{BF} \approx 4.5\sigma$ (estimated in the case of 30 signal events extracted by the fitting procedure). The sensitivity expected for this analysis is consistent with the upper limit on the baryonic mode $\bar{B}^0 \to \Lambda_c^+ p\bar{p}\bar{p}$, previously measured by Grünberg et al. to be 2.8×10^{-6} at a 90% confidence level.

The measurement is dominated by the statistical uncertainty at BABAR data integrated luminosity of 424 fb^{-1} and the prospect of repeating this analysis at Belle II is to reduce this contribution by a factor 10, exploiting the full nominal data set of 50 ab^{-1}, which will be collected by the end of 2024 and which might allow to access scenarios where a significant meaurement is feasible. It is worth noting that the current analysis sensitivity is dominated by the remaining backgrounds in the signal region (17 background events are expected): a possible way of improvement might come from further background suppression based on vertexing requirements. It has been shown in the optimization studies that the distance of closest approach of the reconstructed tracks is a good parameter for rejecting the fake combinations of B meson candidates, which seem to be the main cause of the remaining combinatorial background, since the rate of mis-identified particle is negligible (0.2%). Therefore, a first next step could be to study in more detail the impact of the better vertexing performance foreseen for the SVD at Belle II.

In this thesis, the assembling and the electrical testing of the forward and backward modules of the Belle II SVD, performed at the INFN High Technology Laboratories in Pisa, were also presented. All the phases of the module production in which I actively participated, gaining experience with the complex and delicate procedures, have been illustrated in detail. The overall quality of the produced modules is excellent both on mechanical and on electrical sides. The relative positions of all the components are within the established mechanical tolerances and less than 0.5% strips per module show AC defects. The next phase of the SVD construction, as soon as the ladder assembling is complete, will be the SVD ladder mounting, taking place at KEK laboratory during the next months. As a final remark, the excellent quality of the Belle II SVD is fundamental for an optimal vertex reconstruction, which is also expected to improve the combinatorial background rejection, estimated to be the major source of background for this analysis.

Appendices

Appendix A

Fitting procedure: covariance matrices

The covariance matrices from the fit results discussed in Chapter 6 are reported.

The fit with *sumGauss* to the reduced signal MC sample (see Section 6.1) produced the correlation coefficients listed in Table A.1.

The covariance matrices for the various fits tested in studying the background modeling (Section 6.2) are given in Table A.2, A.3,A.4,A.5.

Parameter	Correlation coefficients					
	Global	n	m_0	N_{sig}	σ_1	σ_2
n	0.23207	1.000	0.005	0.000	-0.229	-0.041
m_0	0.00512	0.005	1.000	0.000	-0.002	0.000
N_{sig}	0.000	0.000	0.000	1.000	0.000	0.000
σ_1	0.23016	-0.229	-0.002	0.000	1.000	0.027
σ_2	0.04462	-0.041	0.000	0.000	0.027	1.000

Table A.1: Correlation coefficients for the shape parameters estimated from fit *sumGauss* on the reduced signal MC sample.

Parameter	Correlation coefficients			
	Global	a	b	N_{bkg}
a	0.33	1.00	0.33	0.00
b	0.33	0.33	1.00	0.00
N_{bkg}	0.00	0.00	0.00	1.00

Table A.2: Correlation coefficients for *Test 1*: all background MC samples after the BDT selection. Both a, b floating.

Parameter	Correlation coefficients			
	Global	a	b	N_{bkg}
a	0.07	1.00	0.07	0.00
b	0.07	0.07	1.00	0.00
N_{bkg}	0.00	0.00	0.00	1.00

Table A.3: Correlation coefficients for *Test 2*: all background MC samples before the BDT selection. Both a, b floating.

Parameter	Correlation coefficients			
	Global	a	b	N_{bkg}
a	0.27	1.00	0.27	0.00
b	0.27	0.27	1.00	0.00
N_{bkg}	0.00	0.00	0.00	1.00

Table A.4: Correlation coefficients for *Test 5*: off-peak data before the BDT selection. Both a, b floating.

Parameter	Correlation coefficients			
	Global	a	b	N_{bkg}
a	0.56	1.00	0.56	0.00
b	0.56	0.56	1.00	0.00
N_{bkg}	0.00	0.00	0.00	1.00

Table A.5: Correlation coefficients for *Test 6*: off-peak data after the BDT selection. Both a, b floating.

Ringraziamenti

Se sono arrivata anche io alla fine, devo ringraziare parecchie persone. Il primo è il mio relatore, il professore Francesco Forti, che ha sempre avuto una pazienza infinita con me, sia nell'insegnarmi, sia nel correggermi, ma soprattutto nello spronarmi sempre ad allargare lo sguardo, senza la paura di vedere le cose nelle loro dimensioni reali. Sono le persone come te che mi fanno ricordare perchè ho scelto di studiare Fisica.

E come lui, tutti i membri della collaborazione Belle II a Pisa. Un grazie a voi che lo scorso gennaio mi avete accolta come laureanda e siete stati per me, in questi mesi, una seconda famiglia. Grazie a Giuliana che mi ha ospitata ogni giorno nel suo ufficio, grazie ad Antonio e Giulia che non si sono mai tirati indietro quando ho chiesto loro una mano, fosse per un turno in laboratorio, un fit o un collaboration drink. Specialmente nell'ultimo caso, grazie al supporto immancabile di Alberto C., Alberto M. e Giacomo. Grazie ad Eugenio e Stefano e al mio maestro Jedi, Thomas. Senza di loro questi ultimi mesi non sarebbero stati gli stessi e il Giappone sarebbe rimasto un sogno.

Anche la California sarebbe stata solo un miraggio, se non fosse stato per *BABAR*: grazie alla collaborazione con cui ho iniziato tutto questo, in particolare a Tina e a Fabio.

Grazie agli amici che da vicino hanno condiviso con me esaltazione e patimenti, a quelli che da casa non mi hanno mai fatto mancare il loro tifo, agli amici di sempre, a Chantal, Veronica e Marta: lo sapete che senza voi tre avrei gettato la spugna mille e mille volte.

Ma soprattutto non so come avrei fatto senza avere accanto ogni giorno te, Francesco, che sei stato la vittima di tutti i miei sbalzi di umore, delle mie paure e delle mie manie. Sappi, però, che tu sei stato anche la mia cura e ogni volta mi hai dato tutta la forza di cui avevo bisogno.

Un ringraziamento speciale va alla mia famiglia che in ogni modo e in ogni momento mi ha supportata e amata, se possibile ancora di più da quando sono partita di casa. So bene che le nonne hanno consumato quantità industriali di ceri per ogni esame o compito di questi cinque anni, che i cugini mi spacciano per chissà che studiosa nei loro temi, che gli zii non hanno esistato un secondo a farsi trecento chilometri pur di passare una giornata con me.

Non trovo le parole, infine, per dire grazie a Matteo e ai miei genitori.
A voi devo ogni cosa.

Bibliography

[1] J. Bernstein, *Spontaneous symmetry breaking, gauge theories, the Higgs mechanism and all that*, Rev. Mod. Phys. **46** (Jan, 1974) 7–48. `http://link.aps.org/doi/10.1103/RevModPhys.46.7`.

[2] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10** (Jun, 1963) 531–533. `http://link.aps.org/doi/10.1103/PhysRevLett.10.531`.

[3] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49** (1973) 652–657.

[4] A. J. Bevan et al., *The Physics of the B Factories*, The European Physical Journal C **74** (2014) no. 11, 431–438. `http://dx.doi.org/10.1140/epjc/s10052-014-3026-9`.

[5] H. Albrecht et al., *Measurement of inclusive baryon production in B meson decays*, Zeitschrift für Physik C Particles and Fields **56** (1992) no. 1, 1–6.

[6] V. Chernyak and I. Zhitnitsky, *B-meson exclusive decays into baryons*, Nuclear Physics B **345** (1990) no. 1, 137 – 172. `http://www.sciencedirect.com/science/article/pii/055032139090612H`.

[7] X.-G. He, T. Li, X.-Q. Li, and Y.-M. Wang, *Calculation of* $\mathrm{BR}(\overline{B}^0 \to \Lambda_c^+ + \overline{p})$ *in the perturbative QCD approach*, Phys. Rev. D **75** (Feb, 2007) 034011. `http://link.aps.org/doi/10.1103/PhysRevD.75.034011`.

[8] K. Abe et al., Belle Collaboration, *Observation of* $B^{\pm} \to p\overline{p}K^{\pm}$, Phys. Rev. Lett. **88** (Apr, 2002) 181803. `http://link.aps.org/doi/10.1103/PhysRevLett.88.181803`.

[9] T. Leddig, *B decays to baryons*, Pramana **79** (2012) no. 5, 1251–1254. `http://dx.doi.org/10.1007/s12043-012-0434-y`.

[10] J. P. Lees et al., BaBar, *Search for the decay* $\bar{B}^0 \to \Lambda_c^+\bar{p}p\bar{p}$, Phys. Rev. **D89** (2014) no. 7, 071102, `arXiv:1312.6800 [hep-ex]`.

[11] K. A. Olive et al., Particle Data Group, *Review of Particle Physics*, Chin. Phys. **C38** (2014) 090001.

[12] P. ODDONE, *An Asymmetric B-Factory Based on PEPa*, Annals of the New York Academy of Sciences **578** (1989) no. 1, 237–247. `http://dx.doi.org/10.1111/j.1749-6632.1989.tb50614.x`.

[13] T. Abe et al., Belle-II, *Belle II Technical Design Report*, `arXiv:1011.0352 [physics.ins-det]`.

[14] *Belle II web page*, 2016. `http://www-superkekb.kek.jp/`. accessed on 7/2016.

[15] B. Aubert et al., BaBar, *The BaBar detector*, Nucl. Instrum. Meth. **A479** (2002) 1–116, `arXiv:hep-ex/0105044 [hep-ex]`.

[16] R. Frühwirth, *Application of Kalman filtering to track and vertex fitting*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **262** (1987) no. 2, 444 – 450. `http://www.sciencedirect.com/science/article/pii/0168900287908874`.

[17] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. A 462 (2001) 152–155.

[18] T. Sjöstrand, *High-energy-physics event generation with PYTHIA 5.7 and JETSET 7.4*, Computer Physics Communications **82** (1994) no. 1, .

[19] S. Jadach, B. F. L. Ward, and Z. Wacs, *Coherent exclusive exponentiation for precision Monte Carlo calculations*, Phys. Rev. D **63** (May, 2001) 113009. `http://link.aps.org/doi/10.1103/PhysRevD.63.113009`.

[20] S. Agostinelli et al., *Geant4—a simulation toolkit*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **506** (2003) no. 3, 250 – 303. `http://www.sciencedirect.com/science/article/pii/S0168900203013688`.

[21] W. D. Hulsbergen, *Decay chain fitting with a Kalman filter*, Nuclear Instruments and Methods in Physics Research A **552** (Nov., 2005) 566–575, `physics/0503191`.

[22] G. D. McGregor, *B Counting at BaBar*, arXiv:0812.1954v1 [hep-ex] (2008) .

[23] R. Brun and F. Rademakers, *ROOT - An Object Oriented Data Analysis Framework, version 5.34/25*, in *AIHENP'96 Workshop, Lausane*, vol. 389, pp. 81–86. 1997.

[24] *TMVA-Toolkit for Multivariate data analysis web page*, 2016.

[25] I. Narsky, *Estimation of upper limits using a Poisson statistic*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **450** (2000) no. 2–3, 444 – 455. `http://www.sciencedirect.com/science/article/pii/S0168900200003053`.

[26] W. Verkerke and D. Kirkby, *The RooFit toolkit for data modeling*, ArXiv Physics e-prints (June, 2003) , `arXiv:physics/0306116`.

[27] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonsances.*, . DESY-F31-86-02, Appendix E.

[28] H. Albrecht et al., *Search for hadronic* $b \to u$ *decays*, Physics Letters B **241** (1990) no. 2, 278 – 282. `http://www.sciencedirect.com/science/article/pii/037026939091293K`.

[29] The Heavy Flavor Averaging Group, *HFAG averages: End Of Year 2014/PDG 2014*, `http://www.slac.stanford.edu/xorg/hfag/index.html`, 2016.

[30] T. Allmendinger et al., *Track Finding Efficiency in BaBar*, Nucl. Instrum. Meth. **A704** (2013) 44–59, `arXiv:1207.2849 [hep-ex]`.

[31] S. D. Biller and S. M. Oser, *Another Look at Confidence Intervals: Proposal for a More Relevant and Transparent Approach*, Nucl. Instrum. Meth. **A774** (2015) 103–119, `arXiv:1405.5010 [physics.data-an]`.