GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

Fakultät für
Physik

Master's Thesis

Sensitivitätsstudie für eine zeitabhängige $B \to K_S \ell\ell$ Analyse mit Belle II

Sensitivity study for a $B \to K_S \ell\ell$ decay-time dependent analysis at Belle II

prepared by

Dennis Rodermund

from Herzberg am Harz

at the II. Physikalisches Institut

Thesis period:	5th August 2025 until 31st October 2025
First referee:	Prof. Dr. Ariane Frey
Second referee:	Prof. Dr. Thibaud Humair

Abstract

Diese Masterarbeit präsentiert eine Sensitivitätsstudie für eine zeitabhängige $B \to K_S \ell\ell$ Analyse mit Belle II. Es wird gezeigt, dass mit dem derzeitig nutzbaren Datensets von Belle und Belle II eine Sensitivität von 0.49 für S_{CP} und 0.40 für C_{CP} erreicht werden kann, wobei die verbleibende Unsicherheit durch die geringe Statistik dominiert wird. Die Sensitivität wird mithilfe einer MVA-gestützten Event Selection unter Verwendung von BDTs sowie einem mehrdimensionalem maximum Likelihood Fit abgeschätzt.

Stichwörter: Physik, Teilchenphysik, B Physik, BDT, MVA, Belle, Belle II, Sensitivität

Abstract

This thesis presents a Sensitivity study for a $B \to K_S \ell\ell$ decay-time dependent analysis at Belle II. It will be shown that with the currently accessible datasets of Belle and Belle II a sensitivity of 0.49 for S_{CP} and 0.40 for C_{CP} can be achived. The remaining uncertainty is dominated by the low statistics. The sensitivity is estimated using a MVA based event selection with the usage of BDTs followed by a multidimensional maximum likelihood fit.

Keywords: Physics, Particle Physics, B Physics, BDT, MVA, Belle, Belle II, Sensitivity

Contents

1. Introduction

One of the most important open questions in modern physics is the origin of the imbalance of matter over antimatter. To allow for an imbalance, three conditions need to be met: there must have been a baryon number violating process in the early universe so that more baryons were formed than antibaryons and there must have been a departure from thermal equilibrium to avoid that a baryon number violating process is balanced by its inverse reaction [1]. Furthermore, CP invariance has to be violated, otherwise for every reaction which creates a net number of baryons there would exist a CP conjugate reaction generating the same net number of antibaryons. Therefore, the violation of CP invariance plays a fundamental role in the evolution of our universe.
The origin of CP violation can be explained by the CKM mechanism [2], however, the CKM mechanism alone is not sufficient to explain the huge imbalance of matter over antimatter. Thus, new sources of CP violation need to be identified.
Recently there has been an observation of anomalies in branching fractions and angular observables in processes involving $b \to s\ell^+\ell^-$ transitions, with the $B \to K_S\,\ell^+\ell^-$ decay channel being one example of those processes [3]. The $b \to s\ell^+\ell^-$ transitions constitute a flavour changing neutral current process and are thus heavily suppressed in the Standard Model, therefore making the transition sensitive to new physics. The observed anomalies could be connected to a new source of CP violation. Due to the low branching fraction of $\mathbf{BR}(B \to K_S\,\ell^+\ell^-) \approx 10^{-7}$ [4] the decay channel however suffers from low statistics limiting the precision a future analysis can achive. In this Thesis, a sensitivity study for a $B \to K_S\ell\ell$ decay-time dependent analysis using the combined Belle and Belle II datasets is presented.
In Chapter 2, an overview of the Standard Model of particle physics is given. In Chapter 3 and 4, the Belle II detector is described and an overview of the used analysis techniques is provided. Chapter 5 describes the event selection applied, followed by a presentation of the maximum likelihood fit to extract the CPV parameters S_{CP} and C_{CP} as well as the results in Chapter 6. Chapter 7 finally discusses the presented analysis and gives an outlook toward the future of a possible full $B \to K_S\,\ell^+\ell^-$ analysis.

2. The Standard Model

The Standard Model of Particle Physics (SM) describes the known elementary particles and their interactions with each other and is currently one of the most successful scientific theories ever created [5, 6, 7, 8, 9]. After a brief introduction to the SM in general in Chapter 2.1 and the weak interaction in Chapter 2.2, the Cabibbo mechanism, the CKM mechanism and CP violation are explained in Chapter 2.3, 2.4 and 2.5, respectively. The limitations of the SM are discussed in Chapter 2.6 and the $B \to K_S \ell^+\ell^-$ decay channel is introduced in Chapter 2.7.

2.1. Fundamentals

All particles can by divided into two groups: Fermions and bosons. While fermions have a half numbered spin, bosons have an integer spin. On the scale of elementary particles, fermions are the particles which make up the matter while the fundamental forces, electromagnetism, the strong interaction and the weak interaction are described by the exchange of (gauge) bosons. Fermions are further divided into two subgroups: the quarks, which interact via the strong force, and the leptons, which do not. An example for leptons are the electrons and their antiparticles, the positrons, which are used for collisions at the SuperKEKB accelerator in Japan. Within the SM, the fermions are grouped into three different generations in order of increasing mass. The particles of the SM are shown in Figure 2.1.

All electrically charged particles interact via the electromagnetic force which is described by the exchange of photons. The electric charge of a particle describes its sensitivity to the electromagnetic force. Furthermore, the quarks interact via the strong force which is described by the exchange of gluons, the gluons also interact via the strong force themself. The charge of the strong force is called the colour charge which has, in contrast to the electric charge, three possible values: blue, green and red as well as the corresponding anticolour charges. Due to colour confinement, a phenomenon which states that quarks can only exist in a state with neutral colour charge, quarks cannot exist as free particles,

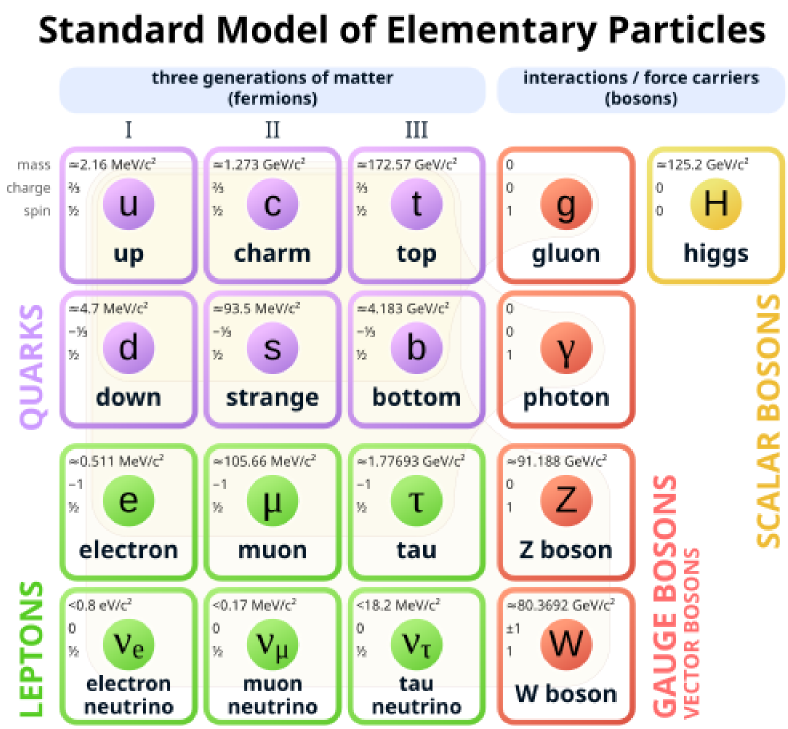

Figure 2.1.: Shown are the elementary particles of the SM together with their mass, their electric charge and their spin. The quarks are coloured in purple, the leptons are coloured in green and the gauge bosons are coloured in red, the Higgs boson is coloured in yellow [10].

but only in bound states, called hadrons. Hadrons can be divided into two groups: The baryons, which have a half numbered spin, and mesons, which have an integer spin. Examples for mesons are the B mesons, which are of special interest for this Thesis. The B mesons were discovered in 1983 by the CLEO Collaboration and consist of a bottom antiquark together with an up, down, strange or charm quark.

2.2. The Weak Interaction

The weak interaction is of special interest for this thesis. All particles except for the photons and the gluons interact via the weak force, which is described by the exchange of W and Z bosons. The sensitivity of a particle to the weak force is described by its weak isospin.

In 1957, Chien-Shiung Wu showed that the weak interaction is parity violating [11]. The W boson is maximally parity violating such that only left-handed particles and right-handed antiparticles couple to the W boson. For the Z boson however the coupling

strength of left- and right-handed particles differ depending on the particle type. Furthermore, given that there is no observation of a right-handed neutrino but of a right-handed antineutrino, the weak interaction also violates the charge inversion.
All leptons have the same coupling strength to the weak gauge bosons. This is known as Lepton Universality (LU). For quarks however, this universality seems to be violated. This violation can be explained by the Cabibbo Mechanism and its extension, the Cabibbo–Kobayashi–Maskawa Mechanism (CKM Mechanism).

2.3. The Cabibbo Mechanism

The Cabibbo Mechanism, published by Nicola Cabibbo in 1963 [12], can explain the missing quark universality for the weak interaction. As described in Chapter 2.1, quarks can only exist in bound states called hadrons, therefore quarks can only be observed in their mass eigenstate. The weak interaction however couples to the weak eigenstate of the quarks. Mass and weak eigenstates of the quarks are rotated against each other, the rotation is described by the Cabibbo matrix, given by

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} \cos(\theta_\mathrm{c}) & \sin(\theta_\mathrm{c}) \\ -\sin(\theta_\mathrm{c}) & \cos(\theta_\mathrm{c}) \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix} \tag{2.1}$$

with the Cabibbo angle θ_c, which denotes the strength of the rotation, experimentally determined to be $\sin(\theta_\mathrm{c}) = V_\mathrm{us} \approx 0.22$ [4].
For the calculation of matrix elements and transition amplitudes, one has to take the rotation of mass and weak eigenstates into account. If, for example, a W boson decays into an up quark and an antidown quark, the W boson couples to the weak eigenstates of the quarks. Therefore the matrix element has to include the rotation factor $\cos(\theta_\mathrm{c})$ in order to take the rotation into account, as shown in Figure 2.2.

The Cabibbo mechanism only accounts for the first two generations of quarks. To describe all three generations of quarks, the CKM mechanism is needed.

2.4. The CKM Mechanism

The CKM mechanism was established in 1973 and is the extension of the Cabibbo mechanism to three generations of quarks [2]. The principle remains the same as for the Cabibbo mechanism: the mass- and the weak eigenstates of the quarks are rotated against each other, the rotation is described by a 3x3 rotation matrix, given as

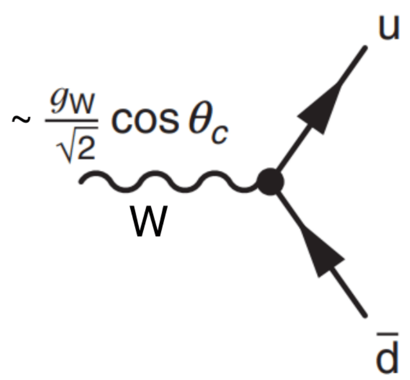

Figure 2.2.: A positively charged W boson decays into an up quark and an antidown quark. The corresponding matrix element involves the Cabibbo angle θ_c.

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{\mathrm{ud}} & V_{\mathrm{us}} & V_{\mathrm{ub}} \\ V_{\mathrm{cd}} & V_{\mathrm{cs}} & V_{\mathrm{cb}} \\ V_{\mathrm{td}} & V_{\mathrm{ts}} & V_{\mathrm{tb}} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{2.2}$$

This matrix is called the CKM matrix. The matrix element V_{ij} is the probability that the up-type quark i decays into the down-type quark j, therefore the CKM matrix has to be unitary. For the calculation of matrix elements and transition amplitudes, one has to take the rotation of the weak- and the mass eigenstates of the quarks into account by applying the corresponding element of the CKM matrix to the vertices, similar as for the Cabibbo mechanism. An example of this is shown in Figure 2.3.

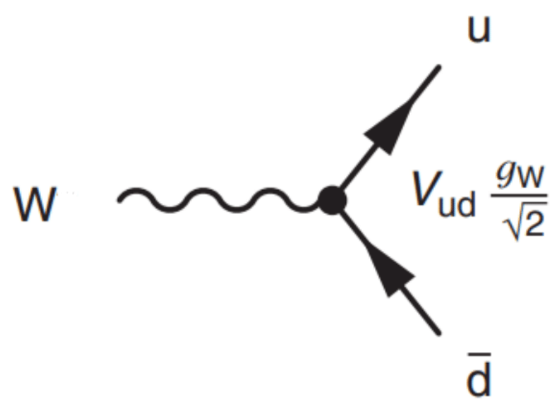

Figure 2.3.: A positively charged W-boson decays into an up and an antidown quark. For the calculation of the matrix element, the CKM matrix element V_{ud} has to be taken into account.

The CKM matrix has 9 complex entries, which corresponds to 18 real parameters. The unitarity of the CKM matrix yields nine constraints. Furthermore, for six quarks five relative phases can be adjusted, resulting in four free parameters, three rotation angles and one complex phase δ, which fully parametrise the CKM matrix. This parametrisation is given by

$$\begin{pmatrix} V_{\mathrm{ud}} & V_{\mathrm{us}} & V_{\mathrm{ub}} \\ V_{\mathrm{cd}} & V_{\mathrm{cs}} & V_{\mathrm{cb}} \\ V_{\mathrm{td}} & V_{\mathrm{ts}} & V_{\mathrm{tb}} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \times \begin{pmatrix} c_{13} & 0 & s_{13}\,e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}\,e^{-i\delta} & 0 & c_{13} \end{pmatrix} \times \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{2.3}$$

In this parametrization, the parameters $c_{12} = \cos(\theta_{12})$, $c_{13} = \cos(\theta_{13})$, $s_{12} = \sin(\theta_{12})$ and $s_{23} = \sin(\theta_{23})$ describe rotation parameters, while δ is a complex phase which models the strength of CP violation within the SM. Since the matrix elements of the CKM matrix are not predicted by theory, but have to be measured, the complex phase could also be equal to zero, in which case CP violation would not manifest itself. A value of δ which is not equal to zero, however, results in CP violation due to differences in transition amplitudes originating from different contributions of conjugated and non conjugated CKM matrix elements. Current measurements of the CKM matrix elements yield [4]

$$\begin{pmatrix} V_{\mathrm{ud}} & V_{\mathrm{us}} & V_{\mathrm{ub}} \\ V_{\mathrm{cd}} & V_{\mathrm{cs}} & V_{\mathrm{cb}} \\ V_{\mathrm{td}} & V_{\mathrm{ts}} & V_{\mathrm{tb}} \end{pmatrix} \approx \begin{pmatrix} 0.974 & 0.221 & 0.004 \\ 0.221 & 0.975 & 0.041 \\ 0.008 & 0.042 & 1.010 \end{pmatrix}. \tag{2.4}$$

The measurements show that the CKM matrix is almost diagonal. Therefore, the rotation of the weak eigenstates to the mass eigenstates is relatively small. Because of this, quarks tend to decay within their own generation. A top quark for example will almost always decay into a b-quark, emitting a W-boson in the process.
The unitarity of the CKM matrix yields the following conditions on the CKM matrix elements:

$$\begin{aligned} \sum_{j=1}^{3} V_{ij}\,V_{kj}^{*} &= 0,\ \forall i \neq k \in \{1,2,3\} \\ \sum_{j=1}^{3} V_{ji}\,V_{jk}^{*} &= 0,\ \forall i \neq k \in \{1,2,3\} \end{aligned} \tag{2.5}$$

Since the matrix elements are complex, the matrix elements and the products of those can be interpreted as vectors in the complex plane. The unitarity conditions therefore correspond to six closed triangles within the complex plane, also called unitarity triangles. One of the unitarity triangles is shown in Figure 2.4. The height of the triangles are given by the complex phase δ. If the complex phase is equal to zero, the unitary triangle would collapse into a straight line, which corresponds to a vanishing CP violation. A measurement of the unitarity triangles therefore gives the opportunity to validate the unitarity of the CKM matrix and to measure the degree of CP violating effects.

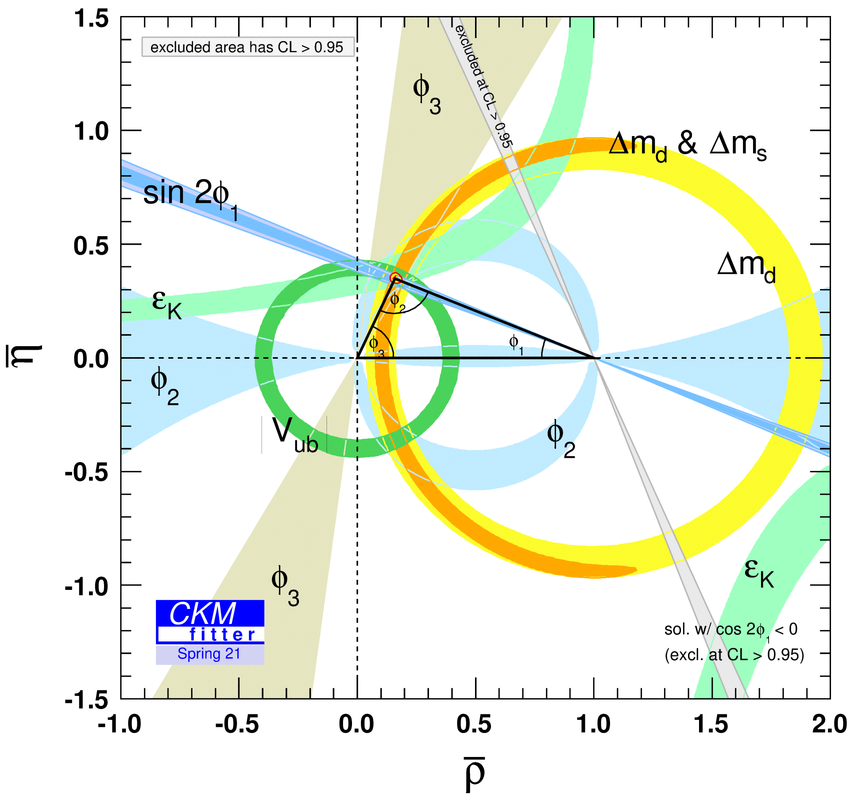

Figure 2.4.: Shown is one of the unitarity triangulars of the CKM matrix together with measurement results of its angles ϕ_1, ϕ_2 and ϕ_3 [13].

2.5. CP Violation

Three different types of CP violation are known: indirect CP violation, direct CP violation and CP violation in the interference of direct and indirect CP violation.
Indirect CP violation occurs when $\Gamma(f \rightarrow \bar{f}) \neq \Gamma(\bar{f} \rightarrow f)$, where Γ refers to the decay width, f is a particle and $\bar{f}$ its corresponding antiparticle. It can therefore only occur in oscillating systems like the neutral Kaon system, where it was first observed in 1964 [14], which also was the first observation of any kind of CP violation.
Direct CP violation occurs when $\Gamma(I \rightarrow f) \neq \Gamma(\bar{I} \rightarrow \bar{f})$, where I is an initial state particle and f its decay product. As with indirect CP violation, the direct CP violation was first observed in the neutral Kaon system. The first observation was made by the

NA48 and KTeV experiments in 2002 [15, 16].
CP violation in the interference of direct and indirect CP violation refers to the phenomena $\Gamma(I \rightarrow f) \neq \Gamma(\bar{I} \rightarrow f)$, where I is an initial state particle and f its final state after a decay. In contrast to direct CP violation, in this case booth the initial state particle I and its antiparticle $\bar{I}$ decay into the same final state particle f, which therefore needs to be an eigenstate of the CP operator. CP violation in the interference was first observed in the neutral B meson system by the Belle and BaBar experiments in 2001 and 2002, respectively [17, 18].
Similar to the neutral Kaon, the neutral B meson can oscillate into its own antiparticle. This process is illustrated in Figure 2.5.

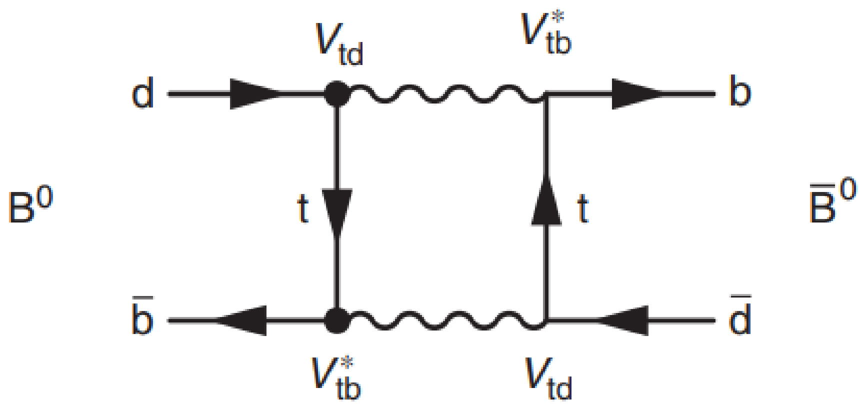

Figure 2.5.: A neutral B meson oscillates into its own antiparticle.

The B^0 meson can decay into a final state f_{CP} with an amplitude $A_{\text{f}_{\text{CP}}}$. However, the B^0 meson can also first oscillate into its antiparticle $\bar{B}^0$ and then decay into the same final state f_{CP} with an amplitude $\bar{A}_{\bar{\text{f}}_{\text{CP}}}$. CP violation in the interference refers to a difference in this process and the same process, where the roles of B^0 and $\bar{B}^0$ are interchanged.

2.6. Limitations of the SM

Despite the huge success of the SM, it has its limitations. For example, the SM cannot describe the nature of gravity on the quantum scale, the nature of dark matter or the huge imbalance between matter over antimatter. This imbalance is of the order $\frac{n_B - n_{\bar{B}}}{n_\gamma} \approx 10^{-9}$, so for every 10^9 antibaryons there must have been 10^9+1 baryons which annihilated to 10^9 photons, leaving one baryon [19, 1].
The violation of CP invariance described in chapter 2.5 may deliver a first source for the imbalance between matter and antimatter in the universe. However, the amount of CP

violation predicted by the SM is too low in order to account for the huge disproportionality between matter and antimatter. Physics which tries to extend the SM in order to give explanations for those phenomena is called physics Beyond the Standard Model (BSM). One example for a theory from the field of BSM is the Supersymmetry (SUSY), in which the assumption is made that every particle of the SM has a "Superpartner" with higher mass [20]. For example, within SUSY, the superpartners of neutrinos, B^0 and π^0 would be candidates for dark matter.

2.7. The $B \to K_S \ell^+\ell^-$ Decay Channel

One way to look for New Physics (NP) which extends the SM is to conduct precision measurements of processes suppressed in the SM. Any observable Q can be expressed as the sum of an amplitude A_{SM} coming from the SM and an amplitude A_{NP} coming from NP:

$$Q \approx |A_{\mathrm{SM}} + A_{\mathrm{NP}}|^2. \tag{2.6}$$

Thus, in order to look for NP, a promising approach is to consider processes which have a very low A_{SM}, thus allowing any NP effects to influence the observable Q via A_{NP} to a noticeable extend.
A Feynman diagram of the $B \to K_S \ell^+\ell^-$ decay is depicted in Figure 2.6. A neutral B meson decays into a K_{S} and two leptons. The decay into the lepton pair occurs via a Flavour Changing Neutral Current (FCNC) and is thus heavily suppressed in the SM due to the loop involved in the decay, resulting in a low branching ratio of $\mathbf{BR}(B \to K_S \ell^+\ell^-) \approx 10^{-7}$ [4]. The $B \to K_S \ell^+\ell^-$ decay involves a $b \to s\ell^+\ell^-$ transition, those transitions are sensitive to new physics due to their suppression in the SM. Within a $B \to K_S \ell^+\ell^-$ decay, the short distance electroweak effects decouple from the long distance QCD effects, which makes the approach of an Effective Field Theory (EFT) applicable. In the context of EFT, the Lagrangian which describes the $B \to K_S \ell^+\ell^-$ decay can be written as a sum of effective coupling constants, called Wilson Coefficients C_i, times Operators O_i:

$$L_{\mathrm{eff}} = \frac{4G_{\mathrm{F}}}{\sqrt{2}} V_{\mathrm{tb}} V_{\mathrm{ts}}^* \frac{g_{\mathrm{e}}^2}{4\pi} \sum_i C_i O_i, \tag{2.7}$$

where G_{F} is referring to the Fermi Constant, V_{tb} and V_{ts}^* are elements of the CKM matrix and g_{e} is the electromagnetic coupling constant. The effects of heavy particle fields like the top quark are integrated out into the Wilson coefficients C_i, lighter fields appear in the operators O_i. If any yet unknown heavy particle is involved in the $b \to s\ell^+\ell^-$ transition, it would therefore reflect itself in a shift of a Wilson coefficient from its SM expectation. Recently there has been an observation of anomalies in branching fractions and angular observables in processes involving $b \to s\ell^+\ell^-$ transitions [3], those anomalies can be explained by a shift of the Wilson Coefficient C_9 from its SM expectation $C_9 \approx 4.9$. Recent measurements indicate that the shift of C_9 is up to a significance of $6.2\,\sigma$ [21]. A shift of the kind $C_9 \to C_9 + \delta C_9$ could allow C_9 to take a complex value. A non vanishing imaginary part of C_9 could be connected to a new source of CP violation. The value of C_9 is constant across all processes involving a $b \to s\ell^+\ell^-$ transition since the short distance electroweak effects are the same for all processes involving a $b \to s\ell^+\ell^-$ transition. Therefore, a time-dependent analysis of the $B \to K_S \ell^+\ell^-$ decay yields the possibility to nail down the coupling constant C_9, thus determining if indeed NP processes are involved in the $b \to s\ell^+\ell^-$ transitions.

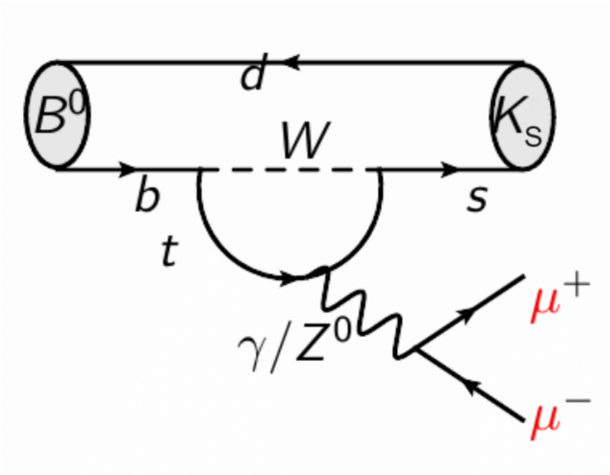

Figure 2.6.: The $B \to K_S \ell^+\ell^-$ decay. A neutral B meson decays into a K_S and two leptons. The decay into the lepton pair occurs via a FCNC and is thus heavily suppressed in the SM, resulting in a low branching ratio.

3. The Belle II Experiment

3.1. The SuperKEKB Accelerator

The SuperKEKB accelerator [22, 23] is an asymmetric positron-electron collider, located at the High Energy Accelerator Research Organisation (KEK) in Tsukuba, Ibaraki Prefecture, Japan. SuperKEKB was constructed by upgrading the original KEKB accelerator [24] to achive higher luminosity.
SuperKEKB operates at a center-of-mass energy of $\sqrt{s} = 10.58$ GeV, which corresponds to the $\Upsilon(4S)$ resonance, thus enabling the SuperKEKB accelerator to produce a high amount of B mesons via the transition $e^- e^+ \rightarrow \Upsilon(4S) \rightarrow B^0 \bar{B}^0$. SuperKEKB is therefore a so-called "B Factory". A schematic of the accelerator is shown in Figure 3.1. The SuperKEKB accelerator consists of an electron-positron linear accelerator, a positron damping ring and an electron and a positron ring. The Belle II detector is placed at the collision point.
SuperKEKB is an asymmetric collider, the energies of the electrons and the positrons are 7 GeV and 4 GeV. This energy asymmetry allows for a boost of the produced B mesons and thus enabling time-dependent measurements, this concept will be discussed further in Chapter 4.2.

3.2. The Belle II Detector

The Belle II detector [25] is an upgrade of the Belle detector [26] used at the KEKB accelerator. Like most particle detectors, it has an onion-like structure, where the different layers of the detector are responsible for measuring the various kinematic and particle identification properties of interest from the incoming particles. A schematic of the Belle II detector is shown in Figure 3.2. The detector consists of a vertex detector, a central drift chamber, a time of propagation chamber, an aerogel ring-imaging Cherenkov detector, an electromagnetic calorimeter as well as a K_{L} and muon detector system.

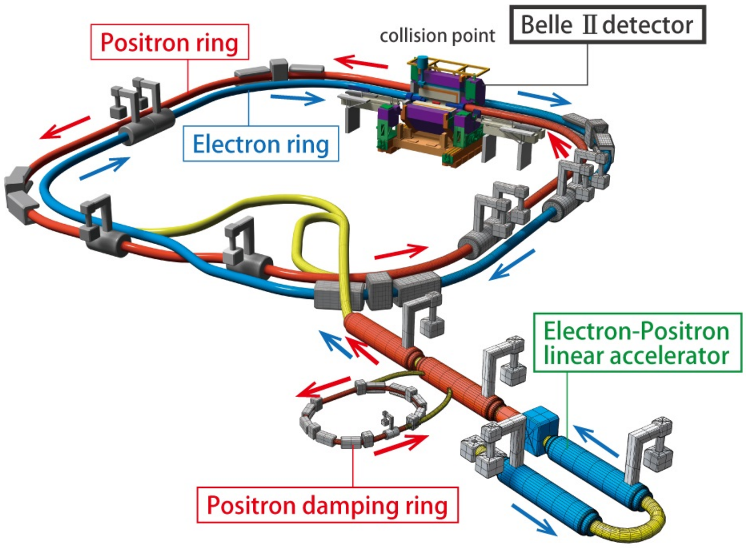

Figure 3.1.: The SuperKEKB accelerator, located at the KEK in Japan. It consists of an electron-positron linear accelerator, a positron damping ring and an electron and a positron ring. The Belle II detector is placed at the collision point [22].

3.2.1. Vertex Detector

The innermost part of the Belle II detector consists of the vertex detector. The vertex detector consists of two parts: The Pixel Detector (PXD) and the Silicon Vertex Detector (SVD). The PXD is responsible for measuring the position of each particle going through it with high precision. A particle passing through the PXD will leave a track due to ionisation. Every detected ionisation process is called a "hit" in the detector. By combining the propagation paths of the various particles it is possible to determine where the "vertex" of those particles is located, where "vertex" is referring to the position of the mother-particle decay. The PXD is based on the DEPFET (DEPleted Field Effect Transistor) technology which allows for very thin (50 μm) sensors. Those sensors are needed close to the interaction point since strip detectors are not capable of dealing with the high beam background at SuperKEKB due to their high occupancy.
The PXD is followed by the SVD which has the task to measure the B vertex information, same as for the PXD. It comprises of double sided silicon microstrip sensors with strips widths down to 50 μm to avoid the huge channel count of pixels. In the SVD, the vertex position is measured via the combination of two 1D positions: the horizontal and the vertical position on the sensor. This has the advantage that the data size is much smaller compared to a 2D position measurement and thus data readout can be much faster. However, close to the interaction point where the particle density is the highest, this set up

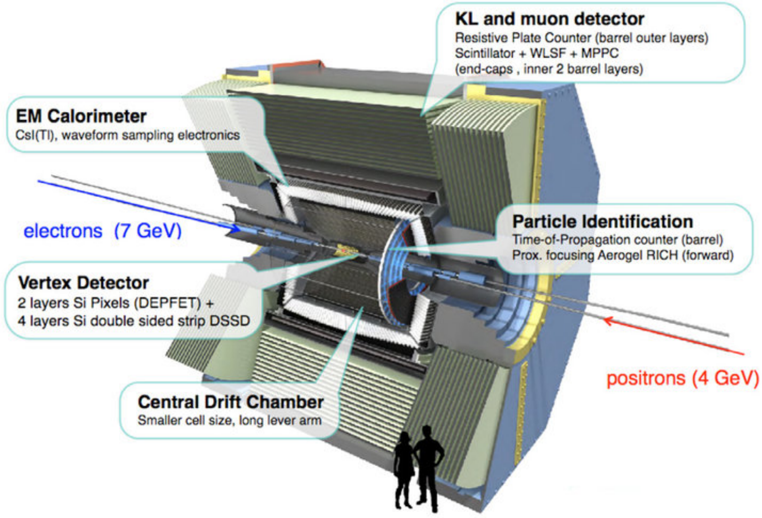

Figure 3.2.: The Belle II detector. The detector is located at the interaction point of the SuperKEKB accelerator and consists of a vertex detector, a central drift chamber, a time of propagation chamber, an aerogel ring-imaging Cherenkov detector, an electromagnetic calorimeter as well as a K_{L} and muon detector system [27].

can lead to problems identifying the correct combination of particles. For this reason, the vertex detector at Belle II consists of a PXD at the inner part and a SVD at the outer part.

3.2.2. Central Drift Chamber

The Central Drift Chamber (CDC) is the main particle track system of Belle II. The aim of the CDC is to reconstruct charged tracks and to measure their momenta as well as their energy. By combining measurements of the momenta with measurements of the energy loss within the gas volume of the CDC, the CDC provides particle identification information. The CDC is filled with a He-C_2H_6 gas, any charged particle which passes through this gas will cause an ionisation charge which then starts to drift to nearby sense wires, where they cause a gas amplification. By inferring the time of each hit within the CDC it is possible to reconstruct the trajectory of a particle through the CDC.
The CDC consists of 14,336 sense wires with a diameter of 30 μm arranged into 56 layers.

3.2.3. Time of Propagation Detector

The Time of Propagation Detector (TOP) aims to provide particle identification information in the barrel region of Belle II by utilizing Cherenkov radiation. A schematic of the TOP is shown in Figure 3.3. The TOP consists of quartz bars. If a charged particle flies through the detector, it will emit Cherenkov radiation. The Cherenkov angle θ_c depends on the velocity of the particle via $\cos(\theta_c) = \frac{1}{n\beta}$ with n being the refraction index of the quartz bars and β the relativistic velocity of the particle. Combining the velocity with the momentum measurements of the CDC and the vertex detector enables a mass measurement and thus an identification of the particle.

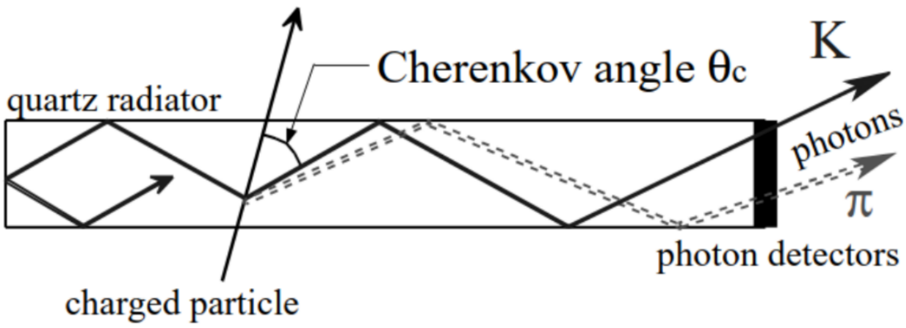

Figure 3.3.: Schematic view of the TOP at Belle II. The TOP consists of quartz bars, a charged particle flying through the detector will create Cherenkov radiation [25].

3.2.4. Aerogel Ring-Imaging Cherenkov Detector

The Aerogel Ring-Imaging Cherenkov (ARICH) detector is based in the forward endcap of Belle II and focusses on providing discrimination power between kaons and pions as well as to separate pions, muons and electrons below 1 GeV. The ARICH detector uses an aerogel to produce Cherenkov radiation if a charged particle passes through the subdetector, very similar to the way the TOP operates. The Cherenkov radiation is forming a cone of light around the particle track, the radius of the cone depends on the velocity of the particle. There is no particle identification system located in the backward endcap of Belle II due to the asymmetric energy of SuperKEKB and the resulting boost for any created particle towards the forward region of Belle II.

3.2.5. The Electromagnetic Calorimeter

The task of the Electromagnetic Calorimeter (ECL) is to measure the electromagnetic energy of photons and to provide particle identification information. If a photon or an electron reaches the ECL, a particle shower evolves within the ECL due to electromagnetic effects like pair production and Bremsstrahlung. A schematic of an arising particle shower is shown in Figure 3.4. The resulting particles are detected using 8736 CsI(Tl) scintillation crystals of 29 distinct shapes which create scintillation light when a particle flies into them. The amount of light created is proportional to the energy deposited within the crystal, the ECL is thus providing a way of measuring the energy of an incoming electron or photon. Muons are minium-ionising particles, thus travelling through the ECL without creating a shower.

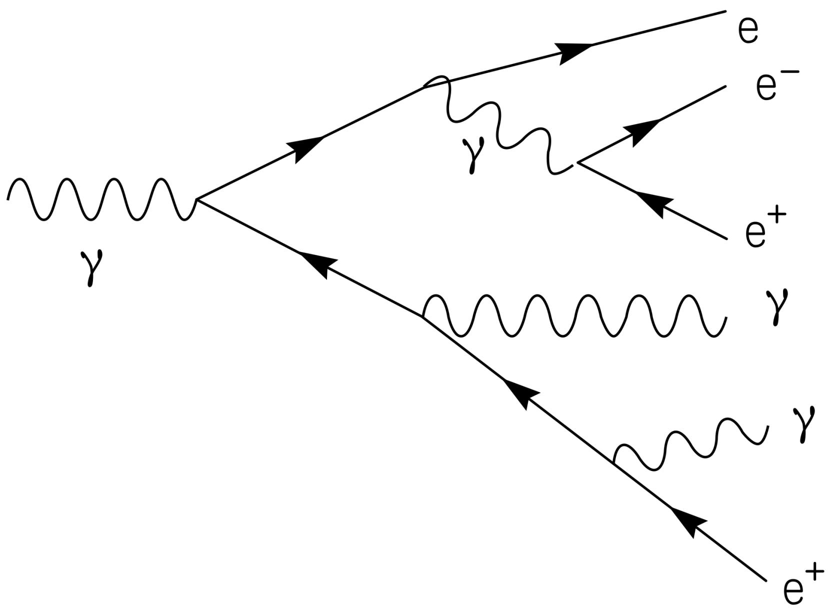

Figure 3.4.: Schematic view of an arising particle shower inside the ECL. New particles are created due to electromagnetic effects like pair production and Bremsstrahlung.

3.2.6. The K_L and Muon System

The K_L and muon detector (KLM) is the outermost layer of the Belle II detector and provides muon identification information and it also reconstructs K_L mesons from the collision. The KLM consists of an alternation between 4.7-cm thick iron plates and active detector elements, functioning as a hadronic calorimeter for the K_L. The KLM is based on glass-electrode resistive plate chambers (RPC). A charged particle travelling through the KLM ionizes the gas molecules along its path. An electric field is applied to the RPCs which then accelerates the electrons toward the anode and the ions toward the cathode. The electrons initiate secondary ionizations, the resulting current indicates a charged particle flying through the KLM.

3.2.7. Detector Parametrization

The Belle II detector is described through a cylindrical coordinate system. The z-axis corresponds to the beam line, with positive values pointing toward the forward direction of Belle II. The x and y axis are perpendicular to the beam line. Furthermore, an azimuthal angle ϕ and a polar angle θ are used. A schematic of the coordinate system used for Belle II is shown in Figure 3.5.

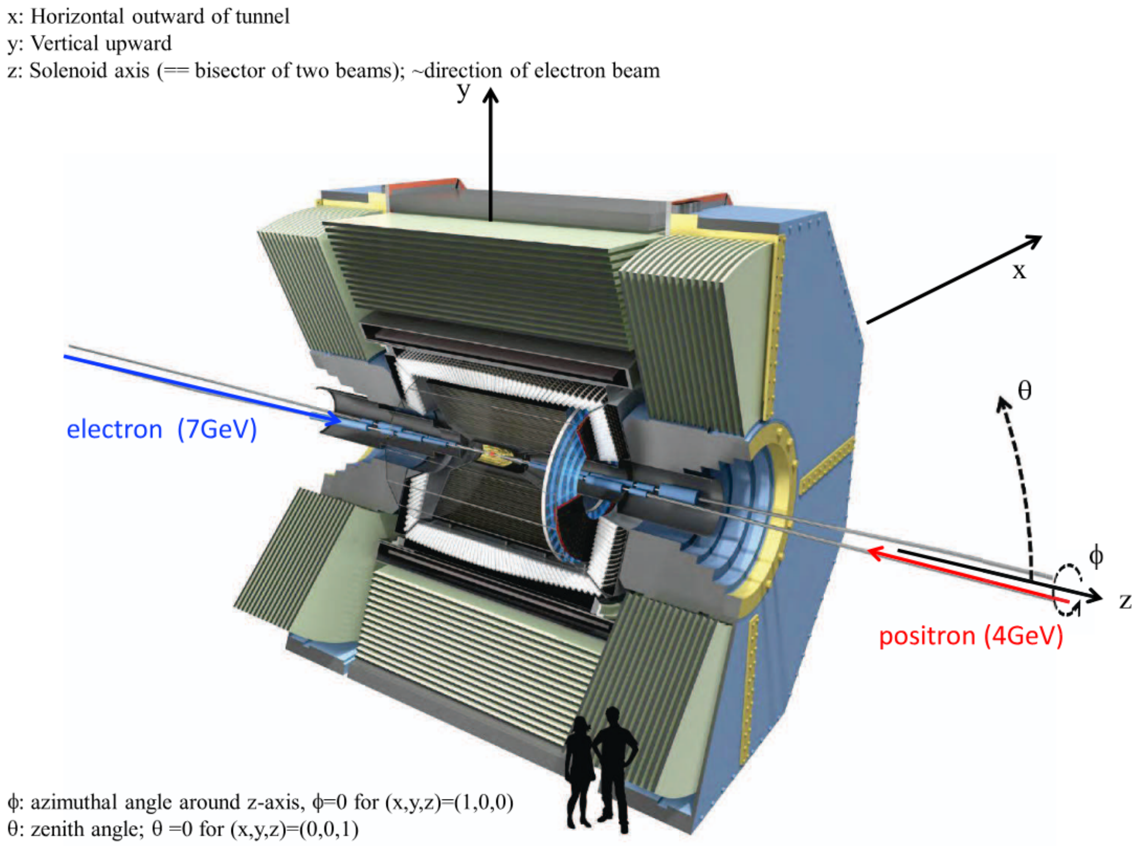

Figure 3.5.: Schematic view of the Belle II detector together with the used coordinate system [28].

4. Analysis Techniques

In this Chapter, an overview of the analysis techniques used for the $B \to K_S \ell\ell$ sensitivity study is provided. In Chapter 4.1, the flavor tagger used at Belle II is described. In Chapter 4.2, a description of how a time-dependent analysis at Belle II is performed is given, followed by an introduction into decision trees and boosted decision trees in Chapter 4.3 and an introduction to maximum likelihood fits in Chapter 4.4.

4.1. Flavor Tagger

In order to determine the Time Dependent CP Violation (TDCPV) of the $B \to K_S \ell^+\ell^-$ decay, the flavor of the tag side B meson needs to be identified. For this, Belle II uses an MVA based flavor Tagger (FT) [29]. The FT uses information from the ECL and KLM clusters as well as tracks in the Rest of Event (ROE) and feeds those informations into an MVA model. The goal of the FastBDT is to determine the flavor specific channel into which the tag B meson decayed, thus identifying the flavor of the B meson. For example, in the decay $b \to Xe^-(b \to ce^-\bar{\nu})$ the charge of the electron tags unambiguously the flavor of the B Meson. The signatures of the flavor specific decay channels can be grouped into 13 different categories.

As a first step, electron, muon, Kaon, pion and proton candidates are reconstructed and their variables are taken as input for a category specific multivariate method. For each category, a FastBDT is applied which outputs the probability that the considered event is part of the corresponding category, thus delivering a probability of a certain flavor. In a second step a Graph Neural Network (GNN) is applied taking the probabilities of the 13 different categories as input, the output of this GNN is a variable qr, which is the product of the flavor $q \in \{+1, -1\}$ of the tag side B meson and the dilution factor $r \in [0, 1]$ indicating the confidence of the FT with respect to its flavor output q. The dilution factor is defined as $r = 1 - 2w$, where w corresponds to the fraction of wrongly tagged events. An overview of the workflow of the FT is shown in Figure 4.1.

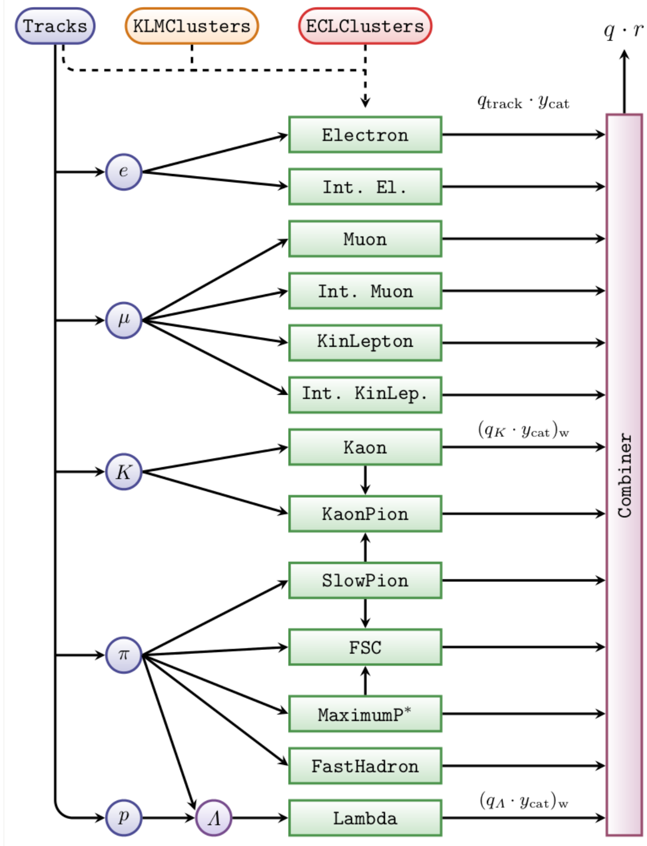

Figure 4.1.: Overview of the workflow of the Belle II FT. The FT takes ROE tracks, KLM clusters and ECL clusters as input and groups the reconstructed particles into 13 different categories. A GNN is applied to combine the results of the different categories, delivering the variable qr as output of the FT [29].

4.2. Time Dependent Analysis

Figure 4.2 shows a sketch of the positron-electron collision at SuperKEKB with the subsequent decay chain. The electron and the positron are colliding, creating a $\Upsilon(4S)$ meson in the process, which immediately decays into a pair of quantum-entangled B^0 mesons. The first B^0 meson to decay is considered the tag-side B^0 meson used to identify the flavor of the signal-side B^0 meson with the FT discussed in Chapter 4.1, exploiting the entangled state of the two B^0 mesons. The decay of the tag-side B^0 meson is happening per definition at $\Delta t = 0$, the signal-side B^0 meson continues its flavor oscillation as described in Chapter 2.5. The probability to observe a flavor specific decay of the B^0_{sig} after the time Δt is given via

$$P(\Delta t, q_f|\tau_{B^0}, \Delta m_d) = \frac{e^{-|\Delta t|/\tau_{B^0}}}{4\tau_{B^0}}[1 + q_f \cos(\Delta m_d \Delta t)], \qquad (4.1)$$

where τ_{B^0} refers to the lifetime of the B^0 meson, Δm_{d} is the mass difference of the mass eigenstates of the B meson and q_f refers to the same flavor state ($q_f = +1$) or opposite flavor state ($q_f =$-1) of the B^0_{sig} meson with respect to the tag-side B^0 meson. For the decay of a B^0 meson into a CP eigenstate, the corresponding PDF is given via

$$P(\Delta t, q_{\mathrm{tag}}) = \frac{e^{-|\Delta t|/\tau_{B^0}}}{4\tau_{B^0}}\{1 + q_{\mathrm{tag}}[-C\cos(\Delta m_d \Delta t) + S\sin(\Delta m_d \Delta t)]\}, \tag{4.2}$$

where q_{tag} refers to the flavor of the tag-side B meson and S and C are the mixing-induced and direct CP violation parameters. Using the wrong tag fraction w of the FT with Δw being the wrong tag fraction difference $\Delta w = w_{B^0_{\mathrm{Tag}}} - w_{\bar{B}^0_{\mathrm{Tag}}}$ and the detection efficiency asymmetry $\mu = \frac{\epsilon_{B^0_{\mathrm{Tag}}} - \epsilon_{\bar{B}^0_{\mathrm{Tag}}}}{\epsilon_{B^0_{\mathrm{Tag}}} + \epsilon_{\bar{B}^0_{\mathrm{Tag}}}}$, equation (4.2) takes the form

$$\begin{aligned} P(\Delta t, q_{\mathrm{tag}}) = \frac{e^{-|\Delta t|/\tau_{B^0}}}{2\tau_{B^0}}\{1 - q_{\mathrm{tag}}\Delta w + q_{\mathrm{tag}}\mu(1-2w) \\ +[q_{\mathrm{tag}}(1-2w) + \mu(1 - q_{\mathrm{tag}}\Delta w)][S\sin(\Delta m_d \Delta t) - C\cos(\Delta m_d \Delta t)]\}. \end{aligned} \tag{4.3}$$

The goal of a time-dependent analysis is to determine the values of S and C using a maximum likelihood fit, which will be further described in Chapter 4.4. In the SM, the expectation on S and C are $S = \sin(2\phi_1)$ and $C = 0$, with ϕ_1 being one of the angles of the unitarity triangles of the CKM matrix which were introduced in Chapter 2.4. Any deviation from the SM expectation would be a clear sign of BSM physics. The time-dependent fit is performed in seven different $|qr|$ bins, where qr refers to the output of the FT. The different bins are defined as $|qr| \in$ (0.0, 0.1), (0.1, 0.25), (0.25, 0.45), (0.45, 0.6), (0.6, 0.725), (0.725, 0.875) and (0.875, 1.0). The values of the FT parameters are determined in each $|qr|$ bin individually by performing an unbinned fit to the background-subtracted Δt distribution of events in which the fully reconstructed B meson decays as $B^0 \to D^{(*)-}\pi^+$, for which the Δt distribution is given via equation (4.1).

The PDF described in (4.3) does not take into account any resolution effects on Δt. For this purpose, a resolution function R is introduced [30]. The resolution function is responsible for taking the residual, defined as $d\Delta\tau = \Delta t - \Delta t_{\mathrm{MC}}$, into account, where Δt_{MC} refers to the value of Δt on truth level. The resolution function consists of a core component modeled by a Gaussian function to account for time resolution uncertainties originating from uncertainties in the reconstruction of the B meson decay vertices. To increase the resolution of the B meson vertices, SuperKEKB is build as an asymmetric collider as discussed in Chapter 3, thus boosting the B mesons and allowing for a higher separation of the two vertices. The resolution function furthermore consists of a tail component modeled by a weighted sum of a Gaussian and two exponentially modified

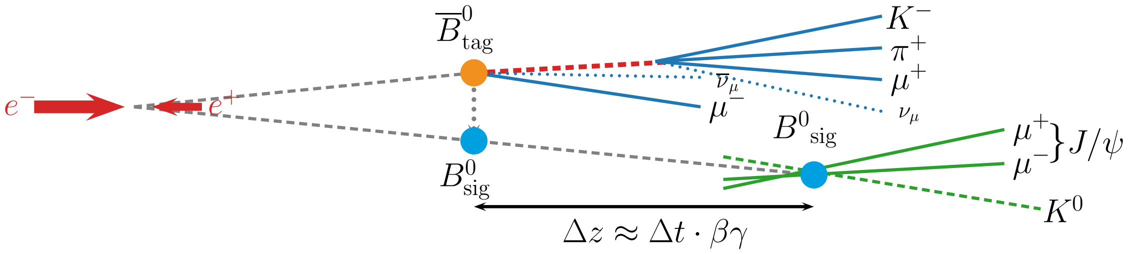

Figure 4.2.: A sketch of the positron-electron collision with subsequent decay of the $\Upsilon(4S)$ into a pair of two B mesons. The first decaying B meson is considered the tag-side B meson used to identify the flavor of the signal-side B meson at the time $\Delta t = 0$. The signal-side B meson later on decays into a neutral Kaon and a J/ψ meson.

Gaussian functions and an outlier component modeled by a Gaussian function to account for poorly determined B_{tag} vertices due to intermediate charm mesons yielding displaced secondary vertices. The Resolution function is thus given by

$$\begin{aligned} R(d\Delta\tau, \sigma_{\Delta t}) &= f_{\text{core}} G(d\Delta\tau; m_{\text{core}}\sigma_{\Delta t}, s_{\text{core}}\sigma_{\Delta t}) \\ &+ f_{\text{tail}} R_{\text{tail}}(d\Delta\tau, m_{\text{tail}}\sigma_{\Delta t}, s_{\text{tail}}\sigma_{\Delta t}, c/\sigma_{\Delta t}, f_>, f_<) \\ &+ f_{\text{OL}} G(d\Delta\tau, 0, \sigma_{\text{OL}}), \end{aligned} \tag{4.4}$$

with G being a Gaussian distribution of mean m and width s and R_{tail} is the sum of a Gaussian distribution and two exponentially modified Gaussian distributions, $\sigma_{\Delta t}$ refers to the uncertainty of Δt. The fractions have to satisfy $f_{\text{core}} + f_{\text{tail}} + f_{\text{OL}} = 1$. The shape parameters are obtained from calibration studies on MC data. Figure 4.3 shows the resolution function together with the residual $d\Delta\tau$ for a subset of the Belle MC sample, clearly showing that the resolution function can model the Δt residual. The total signal PDF for the Δt distribution is obtained by convolving (4.3) with the resolution function.

4.3. Multivariate Analysis

For the classification of events into signal and background events, Multivariate Analysis (MVA) techniques are used [31]. MVA is a statistical tool which has the goal to create algorithms that are capable of learning. Learning means that the considered algorithm gets a set of training data as input and practices to make predictions or decisions without an external user telling the program how to do so. Those predictions are based on underlying relations between the input variables which are hard to take into account manually. The model used to classify events in this thesis is a Boosted Decision Tree (BDT).

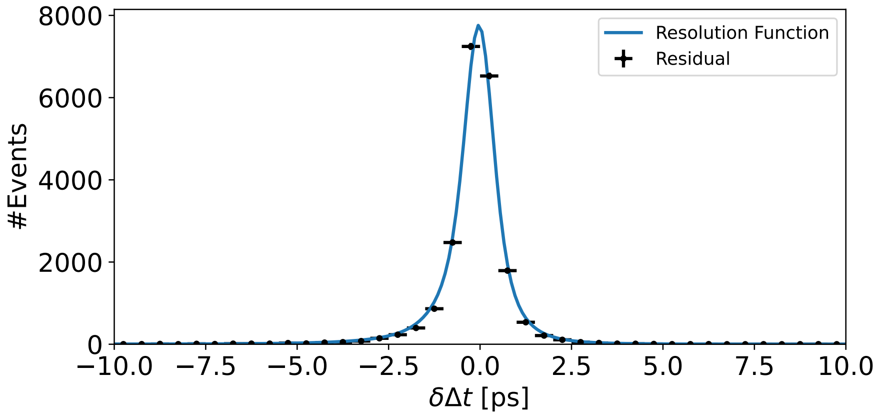

Figure 4.3.: The resolution function together with the Δt residual for a subset of the Belle MC. The resolution function is shown in blue, the residual is shown in black.

4.3.1. Decision Trees

Decision Trees (DTs) are an MVA model developed in the 80's which can be used for classification tasks like a signal-background classification of events within high-energy particle physics [32]. DTs consist of different nodes which can be distinguished into the root node, the inner nodes and the leaf nodes. The root node is the head of the DT, while the leaf nodes are the final nodes of the DT which are used to assign a label to a certain data instance. An example of a simple classification DT used to decide if a passenger of the Titanic survived can be seen in Figure 4.4. The root node of the DT considers the gender of the passenger. Based on the outcome of this variable the DT will go down a certain path. In the case of "female", the DT will go down the right path and reach a leaf node. The corresponding label "survived with a probability of 73%" is assigned to the considered passenger.
If the root node outputs "male", the DT will go down the path to the left so that it reaches the inner node where the age of the passenger is considered. This procedure is iterated until a leaf node is reached and the corresponding label is assigned to the passenger.

4.3.2. Boosted Decision Trees

Boosted Decision Trees (BDTs) are a subclass of DTs [34]. The idea behind boosting is to combine a group of simple weak DTs into a single strong learner by iteratively building new classification models. In the beginning, one simple weak DT is set up and trained on the dataset. After the training, the performance of the weak DT is evaluated on an

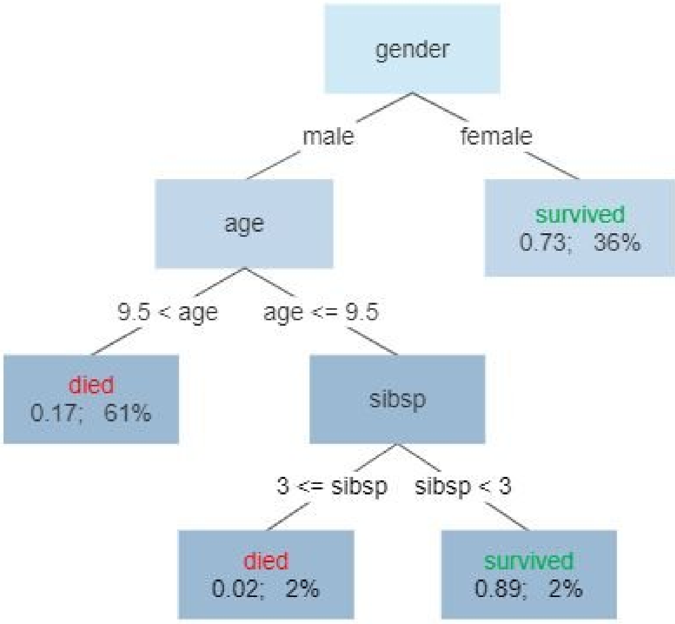

Figure 4.4.: Sketch of a simple DT used to decide if a passenger of the Titanic survived. The numbers in each node are referring to the probability of survival and the percentage of observations in the leaf. The DT makes it decision based on the gender, the age and the number of spouses or siblings on board, referred to as "sibsp" [33].

evaluation dataset. For the next iteration, another simple DT is set up and the training set is reweighed with respect to the former iteration, where the wrongly assigned training instances get a higher weight compared to the previous iteration, thus enabling the DT to correct the errors of the former DT. This procedure is repeated until a certain amount of simple DTs are created. For the final classification of a data instance, the weak DTs are combined into a singular classifier by averaging their respective outputs into one final output. Via this procedure a simple DT which misclassified a certain data instance is outvoted by the following DTs which are supposed to correct the wrong assignment of the first DT.

4.3.3. Construction of Decision Trees

The most relevant question with respect to DTs is how to build the best DT for a given classification task. For this task multiple algorithms exist, but all of them follow the same strategy, a DT is build by iteratively finding the attribute of the data with the highest separation power. Take a data sample S, this sample will have a certain degree of impurity referring to the fact that not all instances within this data sample will have the same class. To measure this impurity, one can define the entropy E of the data sample as

$$E(S) = -\sum_{c=1}^{m} p_c \log(p_c), \tag{4.5}$$

where p_c is defined as $p_c = \frac{|S_C|}{|S|}$ and c being the different classes, p_c the corresponding probability of a randomly drawn instance to be of the class c. The goal of any algorithm building a DT is now to maximise the information gain G by selecting a certain attribute a of the data set. The information gain is defined as

$$G(S,a) = E(S) - \sum_{v \in V(a)} \frac{|S_{a=v}|}{|S|} E(S_{a=v}) \tag{4.6}$$

for each attribute of the data sample. The attribute yielding the highest information gain is chosen as the next attribute on which the DT will base its decision on. For the example shown in Figure 4.4, the attribute with the highest information gain is the gender of the passenger, followed by the age of the passenger and the amount of spouses and siblings on board. The procedure of choosing the best attribute is iterated until no attribute remains or other termination criteria like a certain depth of the DT or a lower limit for data instances within a node is reached. Different algorithms for building a DT mostly differ in their choice of the impurity measure for the data sample, the algorithm shown here uses entropy, the measure used by the ID3 algorithm [35]. In the context of this Thesis, the "XGBoost" [36] library is used to set up the BDTs.

4.4. Maximum Likelihood Estimation

To extract the TDCPV parameters S_{CP} and C_{CP}, a maximum likelihood estimation (MLE) is performed [37]. Given a dataset $\vec{y} = \{y_1, y_2, ..., y_k\}$, one can assume that this dataset is a random sample drawn from an underlying Probability Distribution Function (PDF) with a set of unknown parameters $\vec{\theta} = \{\theta_1, \theta_2, ..., \theta_n\}$. The goal of a MLE is to find the set of parameters $\vec{\theta}$ within the parameter space which best model the observed data.
For this, a likelihood function $L(\vec{\theta})$ is defined as

$$L(\vec{\theta}) = f(\vec{y}, \vec{\theta}), \tag{4.7}$$

where $f(\vec{y}, \vec{\theta})$ describes the likelihood to observe the dataset $\vec{y}$ given the parameter set $\vec{\theta}$. In the context of high energy particle physics, each observed event y_k is randomly drawn and independent, for which the likelihood is defined as

$$L(\vec{\theta}) = \prod_{i=1}^{k} f(y_k, \theta). \tag{4.8}$$

The goal of the MLE is now to find the set of parameters which maximises this likelihood function. In practice, often the negative natural logarithm is applied to simplify the product to a sum, thus enabling a faster computation. In the context of this Thesis, the "iminuit" [38] python interface is used to perform the MLE.

5. Event Selection

The $B \to K_S\,\ell^+\ell^-$ decay channel suffers from very low statistics. The $B \to K_S\,\ell^+\ell^-$ decay channel has a branching fraction of $\mathbf{BR}(\Upsilon(4S) \to K_\mathrm{S}\ell^+\ell^-) \approx 1.1 \cdot 10^{-9}$ [4]. Together with a cross-section of $\sigma(e^+e^- \to \Upsilon(4S)) = 1.1 \cdot 10^{-9}\,b$ and an available luminosity of $L = 1197\,\mathrm{fb}^{-1}$, the available signal events are approximately 280 for the ee and $\mu\mu$ decay channel combined, assuming a signal efficiency of the event selection of 100%. In practise, the number of available signal events after the event selection will be significantly lower due to signal rejection during the event selection. For comparison, the branching fraction of the $B \to J/\psi K_\mathrm{S}$ decay is $\mathbf{BR}(\Upsilon(4S) \to J/\psi K_\mathrm{S}) \approx 4 \cdot 10^{-4}$ and thus five orders of magnitudes higher compared to the branching ratio of $B \to K_S\,\ell^+\ell^-$. Therefore, in order to make the $B \to K_S\,\ell^+\ell^-$ decay channel accessible, a sophisticated event selection is needed.

5.1. Data Sets

The sensitivity study is performed using simulated events. For Belle, one million signal events are generated for each of the ee and $\mu\mu$ decay channels. The events are generated using the EvtGen simulation package [39]. The detector response is simulated using GEANT3 [40]. For background, simulations of basic SM processes, called "generic MC samples", corresponding to four times the total Belle luminosity of $711\,\mathrm{fb}^{-1}$, recorded at the $\Upsilon(4S)$ resonance, are used. The Belle dataset is converted to the Belle II data format for further analysis using B2BII [41].

For Belle II, approximately two million signal events per decay channel are used. For background, the used dataset corresponds to four times the Belle II luminosity of $486\,\mathrm{fb}^{-1}$, recorded at the $\Upsilon(4S)$ resonance during Run 1 and Run 2 of Belle II. Run 1 of Belle II lasted from 2019 to 2022 during which approximately $357\,\mathrm{fb}^{-1}$ of data were recorded at the $\Upsilon(4S)$ resonance. After a long shutdown period, Run 2 took place from 2024 to early 2025, approximately $129\,\mathrm{fb}^{-1}$ of data were recorded at the $\Upsilon(4S)$ resonance during Run 2. For the Belle II simulation, the MC16rd campaign is used. The data files are processed using the Belle II analysis software basf2 [42].

5.2. Preselection

For the Belle data, charged particles are required to have a distance to the interaction point of less than 1.0 cm in the transverse plane and less than 4.0 cm in the z-direction to suppress beam background. For Belle II, these requirements are tightened such that charged particles are required to have a distance to the interaction point of less than 0.5 cm in the transverse plan and less than 2.0 cm in the z-direction in order to reflect the higher vertex resolution of Belle II compared to Belle. Charged particles are furthermore required to satisfy $17^{\circ} < \theta < 150^{\circ}$ to ensure that the particles are within the acceptance region of the CDC and they need to have a transverse momentum of $p_{\mathrm{T}} > 0.1$ GeV to suppress background from low-momentum particles. Electron and muon candidates need to satisfy $L_{\mathrm{e}} > 0.9$ or $L_{\mu} > 0.9$, respectively, in order to distinguish them from charged hadrons, where L_{e} and L_{μ} are the electron and muon ID. Furthermore electrons (muons) are required to satisfy $p > 0.5$ (0.8) GeV.

To account for Bremsstrahlung, the momenta of photons which have an energy below 1 GeV and are detected within a cone of 0.05 rad around the electron are added to the momentum of the electron. To distinguish K_{S} candidates from Lambda Baryons and misreconstructions steming from wrongly combinated tracks at Belle, a NN is trained on the K_{S} momentum, the distance along the z axis between the two track helices at their closest approach, the flight length of the K_{S} in the transverse plane, the angle between the K_{S} momentum and the vector joining the IP with the K_{S} decay vertex, the angle between the pion momentum and the laboratory-frame direction in the K_{S} rest frame, the distances-of-closest-approach in the transverse plane between the IP and the two pion helices and the number of hits in the CDC and the presence or absence of hits in the SVD for each pion track.

For Belle II, a BDT is trained on kinematics of the K_{S} and its daughter pions, flight length of the K_{S} and number of hits in the VXD to separate K_{S} candidates from misreconstructions steming from wrongly combinated tracks. Furthermore, the proton ID of the pions, the Lambda mass which is calculated assuming proton mass hypothesis for one of the daughter pions and kinematics of the pions are used for the separation of K_{S} candidates and Lambda particles. Furthermore, K_{S} candidates need to satisfy 0.46871 GeV $<$ M $<$ 0.5081 GeV, corresponding to a 3σ window around the nominal K_{S} mass and $p_{\mathrm{T}} > 0.1$ GeV.

B meson candidates are reconstructed from combinations of oppositely charged lepton pairs and K_{S} meson candidates. To distinguish signal from background events, the beam-constrained mass M_{bc}, defined as $M_{\mathrm{bc}} = \sqrt{(E_{\mathrm{beam}}/c^2)^2 - (p_{\mathrm{B}}/c)^2}$ and the energy difference ΔE between the B meson candidate and the beam energy E_{B} in the CMS are

considered. B meson candidates need to satisfy 5.20 GeV $< M_{\rm bc} <$ 5.29 GeV and -0.10 GeV $< \Delta E <$ 0.25 GeV. For events with multiple B meson candidates, the candidate with the lowest χ^2 value calculated by a vertex fit is retained. The vertex fits for the B meson candidates are performed using the TreeFitter module of basf2 [43].
Many $B \to J/\psi(\to \ell^+\ell^-)K_{\rm S}$ and $B \to \psi(2S)(\to \ell^+\ell^-)K_{\rm S}$ events survive those cuts. In order to suppress those decays, a veto on the invariant mass of the dilepton system is applied. The veto is illustrated in Figure 5.1 for the Belle decay channels, the corresponding plot for Belle II can be found in Appendix A.

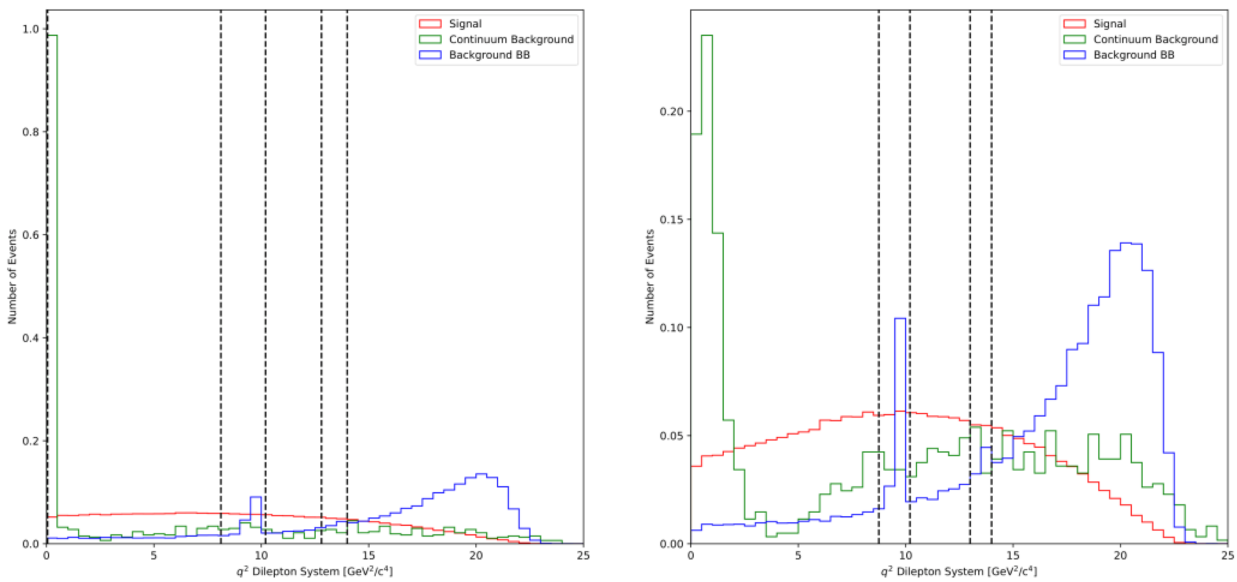

Figure 5.1.: Veto applied on the invariant mass of the dilepton system for the Belle ee decay channel (left) and the Belle $\mu\mu$ decay channel (right). Signal events are shown in red, continuum background events in green and BB related background events are shown in blue. The cuts applied are indicated with black doted lines.

The decay $B \to J/\psi(\to \ell^+\ell^-)K_{\rm S}$ is suppressed by applying a veto on the region $q^2 \in$ [8.12, 10.20] GeV^2 for the Belle ee decay channel, $q^2 \in$ [8.2, 11.0] GeV^2 for the Belle II ee channel and $q^2 \in$ [8.75, 10.20] GeV^2 for the $\mu\mu$ decay channels, where q^2 refers to the squared invariant mass of the dilepton system. Similarly, in order to suppress $B \to \psi(2S)(\to \ell^+\ell^-)K_{\rm S}$ decays, a veto is applied on $q^2 \in$ [12.8, 14.0] GeV^2 for the Belle ee decay channel, $q^2 \in$ [12.8, 15.0] GeV^2 for the Belle II ee channel and $q^2 \in$ [13.0, 14.0] GeV^2 for the $\mu\mu$ decay channels. Additionally, for the ee decay channels a veto on the region $q^2 <$ 0.05 GeV^2 is applied to suppress low energy photo conversions $\gamma^* \to e^+e^-$ and pion decays $\pi^0 \to \gamma e^+e^-$.

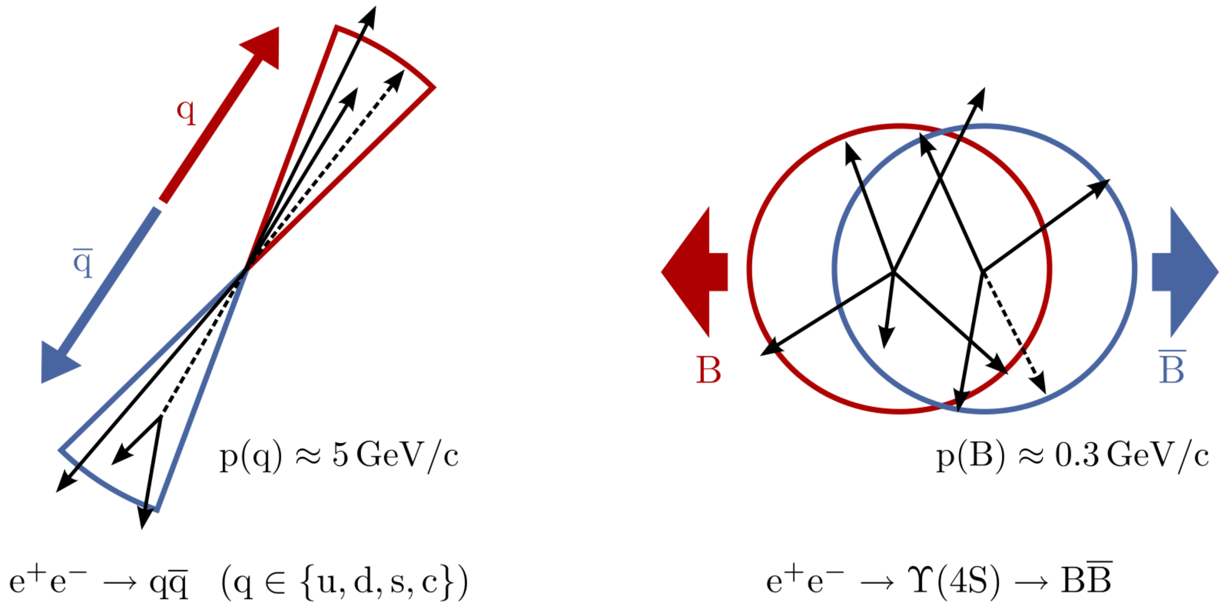

Figure 5.2.: Typical event topologies for continuum events (left) and BB events (right). Continuum events tend to have a back-to-back dijet event structure whereas BB events tend to have a spherically distributed topology.

5.3. MVA Based Event Selection

Background events at Belle and Belle II can be divided into two categories: Background events arising from $e^- e^+ \to q\,\bar{q}$, where $q \in \{u, d, c, s\}$, called "Continuum Events", and background events containing a pair of B mesons, called "BB Background". Given that the B mesons within the BB background are produced almost at rest, BB background events tend to have a spherically shaped event topology. Continuum events in contrast tend to have a back-to-back dijet topology due to their higher amount of energy available to each quark. The typical event topologies are sketched in Figure 5.2.

To suppress the remaining background events, BDTs are used to perform a signal-background separation. The training variables are chosen such that they have a discriminating power between signal and background events. The BDTs are trained using the helicity of the B mesons, the cosine of the angles between the thrust axis of the signal B meson and the z-axis as well as the thrust axis of the Rest-of-Event (ROE), the thrust magnitudes of the signal B meson and the ROE, a set of Fox-Wolfram moments and modified Fox-Wolfram moments used to characterise the event shape topology [44], mass, flight significance and MVA discriminators of the K_S, a set of Cleo-Cones [45], variables characterising the ROE like the multiplicity of charged Kaons and K_S in the ROE or the energy deposited within the ROE, PID variables and the χ^2 values of the dilepton system and B meson candidate vertex fits. A detailed list of the used training variables together with their importance to the signal-background separation can be found in Appendix B. The amount of signal and background events the BDTs are trained on are summarised in Table 5.1. There is a significant imbalance between the amount of signal events and the amount of background events, therefore weights are assigned to the signal and background events to account for this imbalance. In order to avoid overtraining, the dataset is randomly divided into ten different folds. For each decay channel, ten BDTs are set up. BDT i is trained on all folds j and evaluated on fold i with $i \neq j$ and $i, j \in \{1, 2, ..., 10\}$. The BDTs output a single value ranged between 0 and 1, where 0 corresponds to background events and 1 corresponds to signal events. Figure 5.3 and Figure 5.4 show the output distribution and the ROC curve of the BDT training for the Belle ee decay channel, corresponding plots for the other decay channels are shown in Appendix A.

Table 5.1.: Amount of Signal and Background Events the BDTs are trained on.

Channel	Signal Events	Background Events
Belle ee	312,858	22,856
Belle $\mu\mu$	340,211	28,353
Belle II ee	210,186	13,567
Belle II $\mu\mu$	177,552	16,393

The training output for an event is calculated by averaging the training output of all nine BDTs trained on the event. The training and test output distributions agree well with each other, indicating that overtraining is negligible. Furthermore the distributions show that the BDTs have a high power of separating signal and background events, which is in agreement to the ROC curve. In order to remove background events, a cut on the BDT output is applied at 0.6, the value of the cut was chosen such that the signal efficiency of the cut is 90%-95%, dependent on the decay channel, with a background rejection rate of approximately 90%, thus allowing to remove a huge proportion of the background without a significant loss of signal. Table 5.2 summarises the signal efficiencies of the event selection, Table 5.3 summarises the corresponding signal event yields.

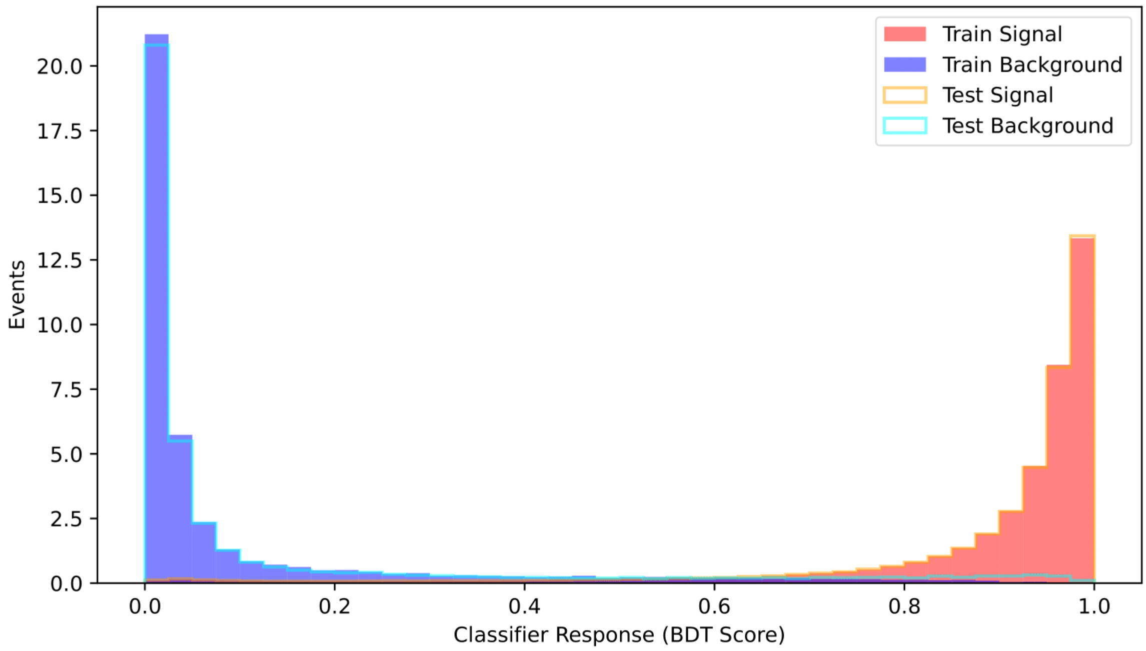

Figure 5.3.: BDT output distribution for the Belle *ee* decay channel. Shown are the distributions of the training and the test scores. The events are normalised such that the area of the histogram corresponds to 1.

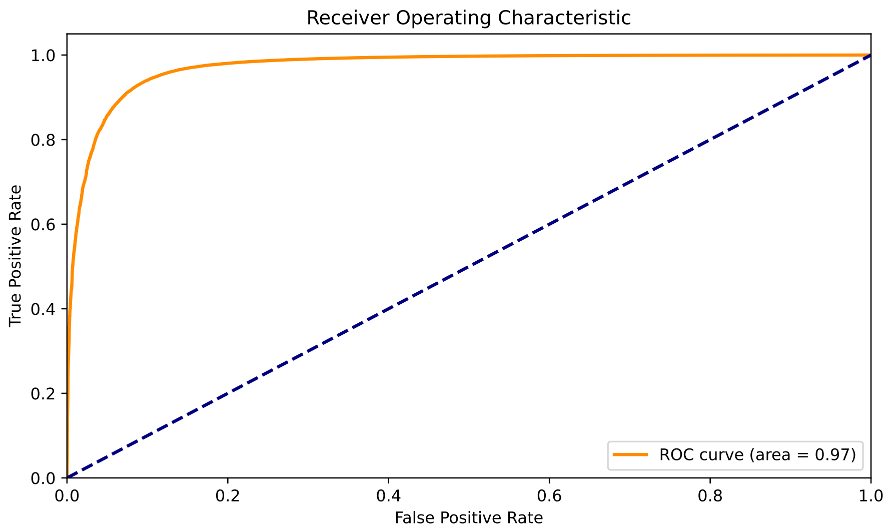

Figure 5.4.: ROC curve for the Belle *ee* decay channel, shown in orange. The AUC score equals 0.97, indicating a high separation power between signal and background events. The dotted blue line shows the hypothetical ROC curve for an AUC score of 0.5.

Table 5.2.: Accumulated Signal Efficiencies of the Event Selection.

Cut	Belle ee	Belle $\mu\mu$	Belle II ee	Belle II $\mu\mu$
Reconstruction	53.32%	50.14%	55.56%	58.47%
Lepton Cuts	47.84%	44.65%	51.50%	35.00%
K_S Cuts	44.29%	41.25%	47.81%	32.45%
M_{bc} and ΔE Cut	38.23%	39.67%	43.04%	31.23%
q^2 Veto	31.29%	34.02%	31.61%	26.71%
BDT Cut	29.13%	31.58%	30.04%	25.32%

Table 5.3.: Signal Events accessible after corresponding Cut.

Cut	Belle ee	Belle $\mu\mu$	Belle II ee	Belle II $\mu\mu$	Belle	Belle II	Combined
Reconstruction	44.59	42.16	31.73	33.40	76.32	75.56	151.88
Lepton Cuts	40.01	37.34	29.43	20.00	77.35	49.43	126.78
K_S	37.04	34.50	27.32	18.55	71.54	45.87	117.41
M_{bc} and ΔE Cut	31.93	33.18	24.60	17.85	65.11	42.45	107.56
q^2 Veto	26.17	28.45	18.06	15.27	54.62	33.33	87.95
BDT Cut	24.36	26.41	17.17	14.47	50.77	31.64	82.41

6. CPV Parameter Extraction

6.1. Set Up of the Fit

After the event selection described in Chapter 5 is applied, a likelihood fit is performed to extract the S_{CP} and C_{CP} parameters. The fit region is defined as 5.23 GeV $< M_{\mathrm{bc}} <$ 5.29 GeV, -0.1 GeV $< \Delta E <$ 0.2 GeV and $|\Delta t| <$ 10 ps. Further cuts are applied to ensure the vertex resolution quality. Those cuts are 0.05 ps $< \Delta t_{\mathrm{err}} <$ 2 ps and $\chi^2 > 0$ for the signal and tag-side B meson vertex fits. The fit is performed over M_{bc}, ΔE, Δt and the BDT output score.
To model the BDT output score distribution with an analytic function, the BDT output is transformed via $O' = \log\left(\frac{O-0.6}{1-O}\right)$, where O and O' are referring to the original and transformed BDT output scores, and 0.6 and 1 are the minium and maximum values of the original BDT output according to the event selection described in Chapter 5. The data is fitted in seven different bins of $|qr|$, where qr refers to the output of the Flavour Tagger introduced in Chapter 4.1. The different bins are defined as $|qr| \in$ (0.0, 0.1), (0.1, 0.25), (0.25, 0.45), (0.45, 0.6), (0.6, 0.725), (0.725, 0.875) and (0.875, 1.0).
The PDF shapes assumed for the signal, BB background and continuum background distributions for the different fit variables are summarised in Table 6.1.

Table 6.1.: Assumed PDF shapes for the signal and background distributions.

Variable	Signal Shape	BB Background Shape	Continuum Background Shape
M_{bc}	Double Crystal Ball [46]	ARGUS [47] and Gaussian	ARGUS
ΔE	Double Crystal Ball	Exponential	Exponential
Transformed BDT Output	Bifurcated Gaussian [48]	Bifurcated Gaussian	Bifurcated Gaussian
Δt	Δt Signal PDF	Exp. + Double Gaussian	Exp. + Double Gaussian

For Δt, the signal PDF is given by the convolution of equation (4.3) with the resolution function (4.4) as described in Chapter 4.2. To determine the PDF shape parameters, 100,000 signal events per decay channel and all available background events are used to perform a likelihood fit which outputs the PDF shape parameters best suitable to model the given distribution. For the PDF shape parameter determination, all fit projections are normalised to unity. The Figures 6.1, 6.2 and 6.3 show the result of the PDF shape determination for M_{bc}, whereas Figures 6.4, 6.5 and 6.6 show the corresponding plots for ΔE. The plots for M_{bc} and ΔE show the results for all $|qr|$ bins combined. The plots indicate a good fit quality. For the BDT output score distribution a $|qr|$ bin dependent disagreement between the MC Data and the fit projection is observed however for the signal PDF shape determination, where the disagreement becomes more distinct for an increasing value of $|qr|$. Figures 6.7, 6.8 and 6.9 show the PDF fit projections for the BDT output score in the $|qr|$ bins 0 and 6 for Belle, the corresponding plots for Belle II can be found in Appendix C. The background PDF fit projections for Δt can be seen in Figure 6.10 and 6.11. The PDF shape parameters for the Δt signal PDF were determined using the methods described in Chapter 4.2.

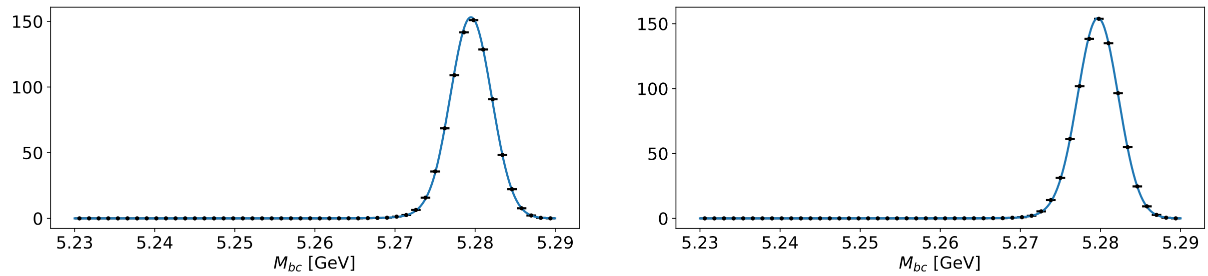

Figure 6.1.: Signal PDF shape for M_{bc} for Belle (left) and Belle II (right). The black dots indicate the MC data, the blue line is the result of the likelihood fit.

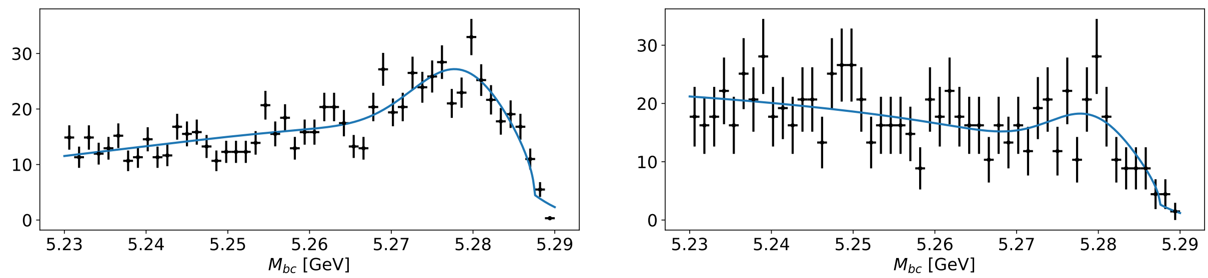

Figure 6.2.: BB background PDF shape for M_{bc} for Belle (left) and Belle II (right). The black dots indicate the MC data, the blue line is the result of the likelihood fit.

6.2. Results

After the PDF shape parameters are fixed, another likelihood fit is performed to extract the signal and background event yields as well as the S_{CP} and C_{CP} parameters. For the

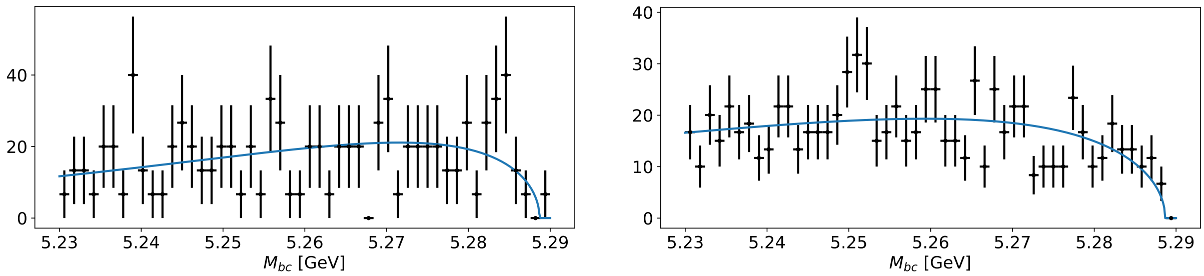

Figure 6.3.: Continuum Background PDF shape for M_{bc} for Belle (left) and Belle II (right). The black dots indicate the MC data, the blue line is the result of the likelihood fit.

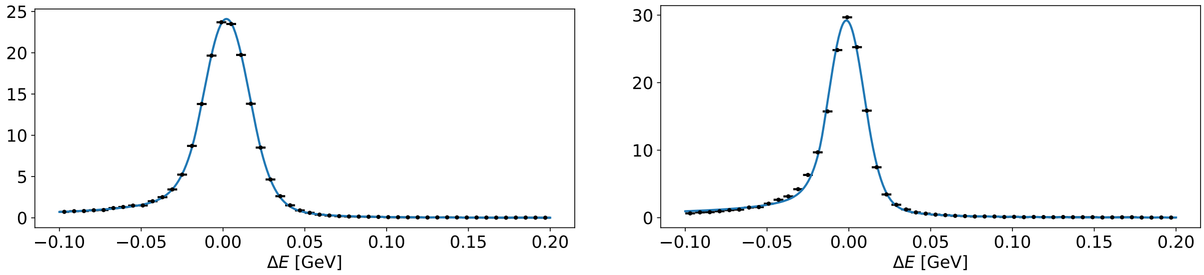

Figure 6.4.: Signal PDF shape for ΔE for Belle (left) and Belle II (right). The black dots indicate the MC data, the blue line is the result of the likelihood fit.

fit, the MC data is sampled to reflect the expected signal and background event yields at Belle and Belle II. The MC event yields available for the fit after the applied event selection and sampling are summarised in Table 6.2.

Table 6.2.: MC Event Yields available for the Fit.

Dataset	Signal	BB Background	Continuum Background
Belle	48	622	27
Belle II	29	136	134
Belle + Belle II	77	758	161

In the case of ΔE and the transformed BDT output, the PDF shape functions are the same for the BB background and the continuum background. The fit therefore has problems to distinguish the continuum and BB background events. To solve this issue, the event yields of the continuum background are fixed to the MC expectation during the fit whereas the signal and BB background event yields are kept free floating. During a complete analysis on real data, this set up can be realised by estimating the continuum background event yields before the likelihood fit using off-resonance data, then fixing the

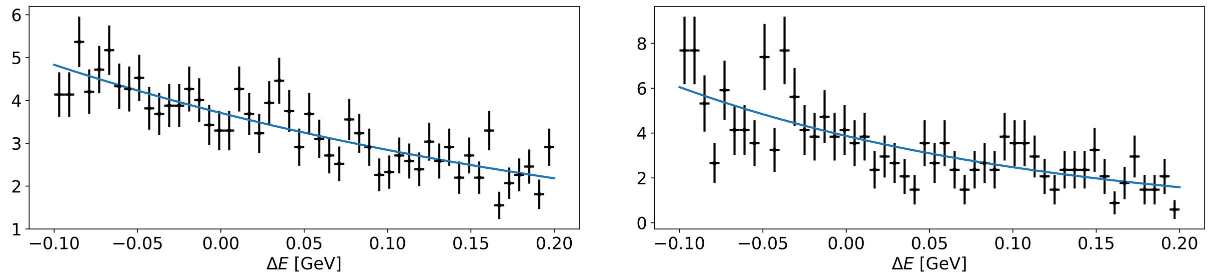

Figure 6.5.: BB background PDF shape for ΔE for Belle (left) and Belle II (right). The black dots indicate the MC data, the blue line is the result of the likelihood fit.

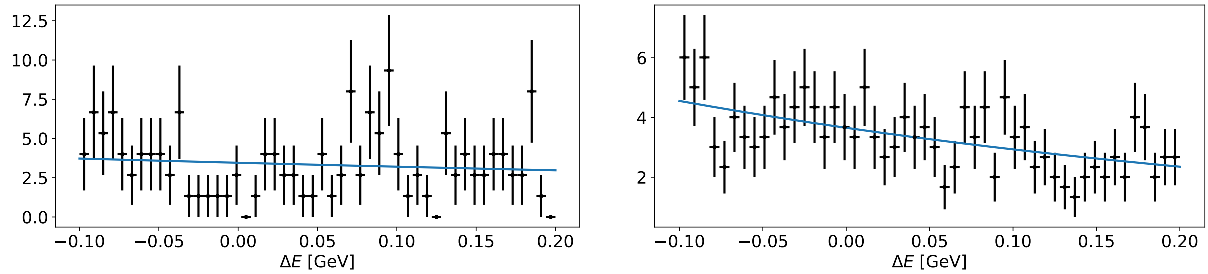

Figure 6.6.: Continuum Background PDF shape for ΔE for Belle (left) and Belle II (right). The black dots indicate the MC data, the blue line is the result of the likelihood fit.

continuum event yield to the estimated value. The fit is performed on the Belle and the Belle II datasets separately to ensure the quality of the fit as well as on the combination of the Belle and Belle II datasets to extract the event yields and the CPV parameters. In this Chapter, the results for the fit on the combined Belle and Belle II datasets are presented. The corresponding plots and tables for the Belle and Belle II datasets separated can be found in Appendix C. To perform the likelihood fit, the PDFs of the different fit variables are assumed to factorise, the logarithm of the likelihoods of the Belle and Belle II datasets are summed up. The likelihood for Belle is given via

$$\begin{aligned} L_{\text{Belle}} &= \text{PDF}_{\text{Belle}}(M_{\text{bc}}) * \text{PDF}_{\text{Belle}}(\Delta E) \\ &\quad * \text{PDF}_{\text{Belle}}(BDT) * \text{PDF}_{\text{Belle}}(\Delta t). \end{aligned} \tag{6.1}$$

Additionally, the event yields $n_{\text{Belle}}(s)$ and $n_{\text{Belle}}(b)$ of the signal and background events are added to the logarithm of the likelihood to take the normalisation of the PDF shapes into account. The corresponding logarithm of the likelihood for Belle II is added to get the full likelihood. Table 6.3 summarises the results of the fit and compares it to the

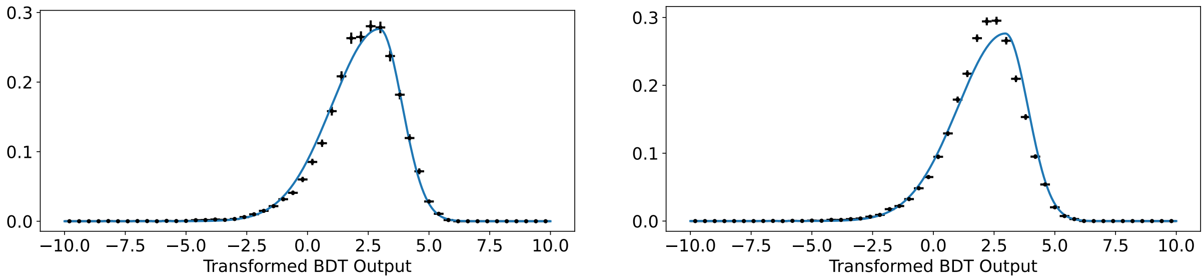

Figure 6.7.: Signal PDF shape for the transformed BDT output for Belle in the $|qr|$ bin 0 (left) and 6 (right). The black dots indicate the MC data, the blue line is the result of the likelihood fit.

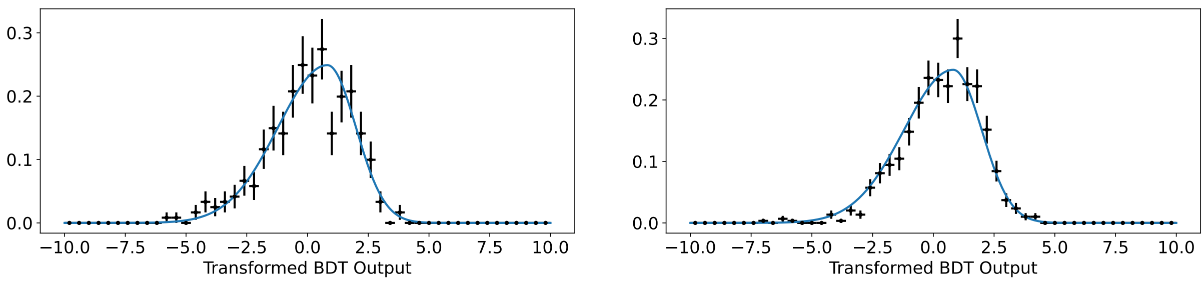

Figure 6.8.: BB background PDF shape for the transformed BDT output for Belle in the $|qr|$ bin 0 (left) and 6 (right). The black dots indicate the MC data, the blue line is the result of the likelihood fit.

expectation from MC. Figures 6.12, 6.13, 6.14 and 6.15 show the fit projections for the different fit variables. Figure 6.16 shows the Δt projection of the fit for the signal region, defined as $M_{\text{bc}} > 5.27$ GeV and $|\Delta E| < 0.05$ GeV.

Table 6.3.: Results of the Fit on the combined Belle and Belle II Dataset.

Variable	Result	Uncertainty	MC Expectation
S_{CP}	-0.28	0.49	0.0
C_{CP}	-0.45	0.40	0.0
Signal Events Belle	51	13	48
Signal Events Belle II	32	9	29
BB Background Events Belle	619	37	622
BB Background Events Belle II	133	22	136
Continuum Background Events Belle	27	0	27
Continuum Background Events Belle II	134	0	134

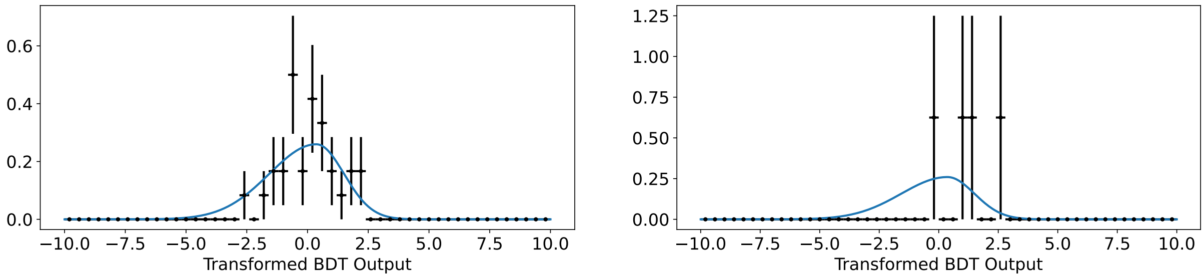

Figure 6.9.: Continuum background PDF shape for the transformed BDT output for Belle in the $|qr|$ bin 0 (left) and 6 (right). The black dots indicate the MC data, the blue line is the result of the likelihood fit.

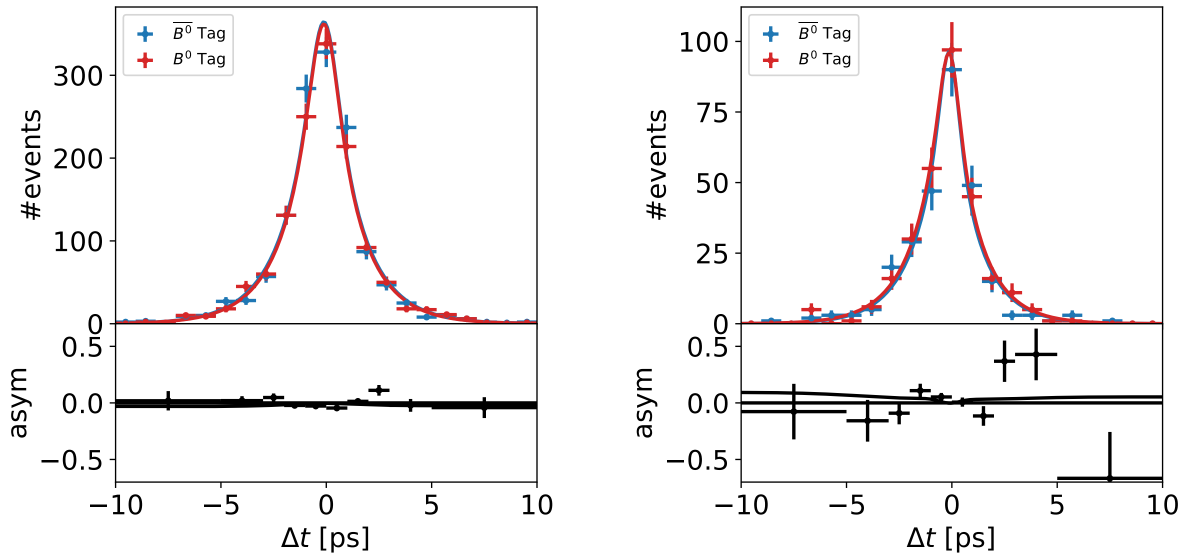

Figure 6.10.: *BB* background PDF shape for Δt for Belle (left) and Belle II (right). The PDF shape parameters are determined using the full background MC sample and a maximum likelihood fit.

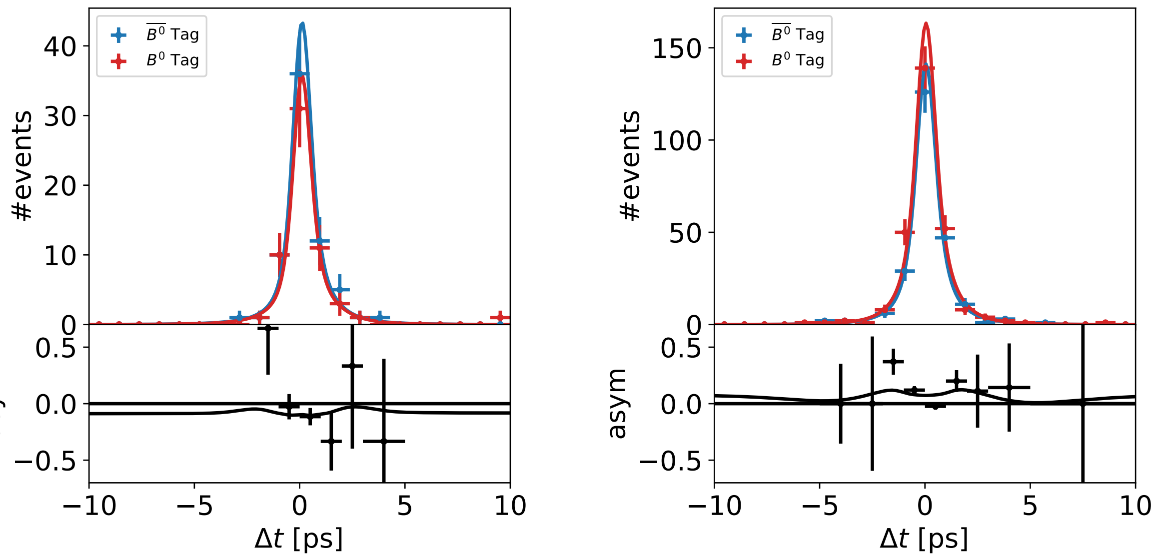

Figure 6.11.: Continuum background PDF shape for Δt for Belle (left) and Belle II (right). The PDF shape parameters are determined using the full background MC sample and a maximum likelihood fit.

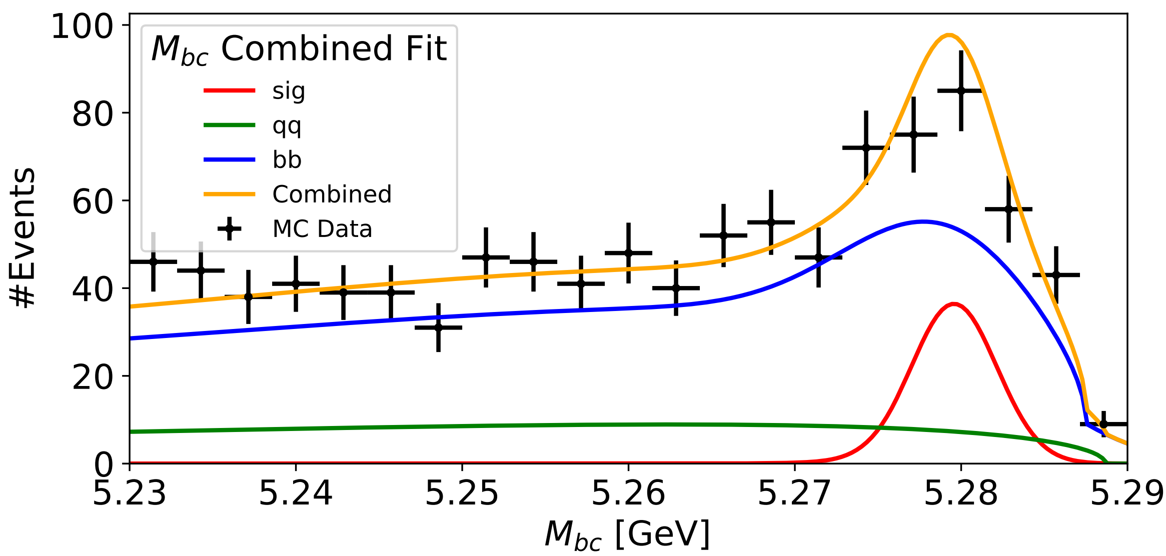

Figure 6.12.: Fit Projection of M_{bc} for the fit on the combined Belle and Belle II dataset. The Signal, BB background and continuum background yields are shown in red, blue and green, respectively. The sum of the event yields is shown in yellow, the MC data is shown in black.

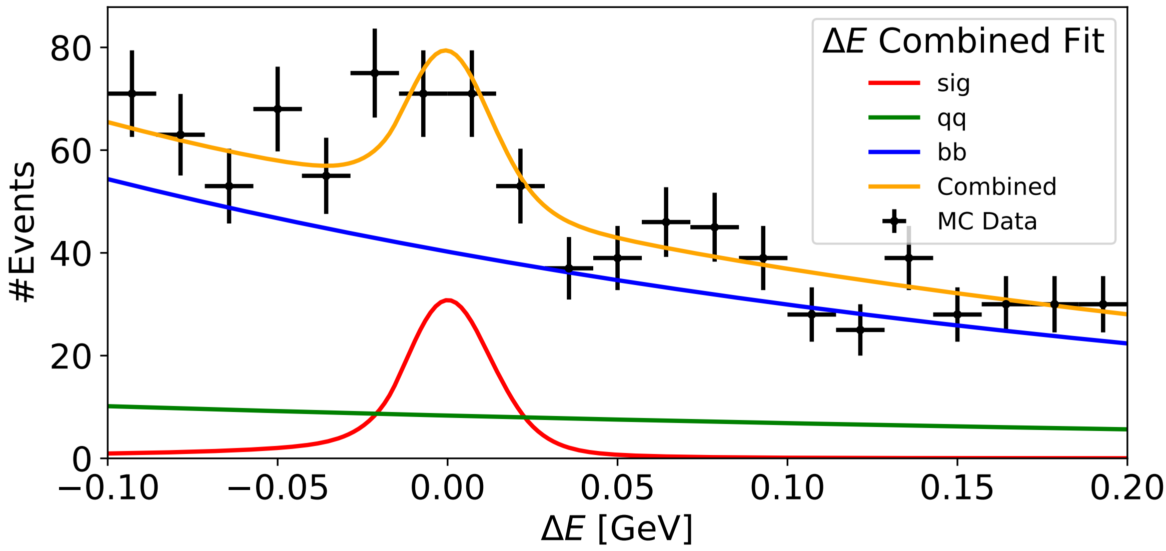

Figure 6.13.: Fit Projection of ΔE for the fit on the combined Belle and Belle II dataset. The Signal, BB background and continuum background yields are shown in red, blue and green, respectively. The sum of the event yields is shown in yellow, the MC data is shown in black.

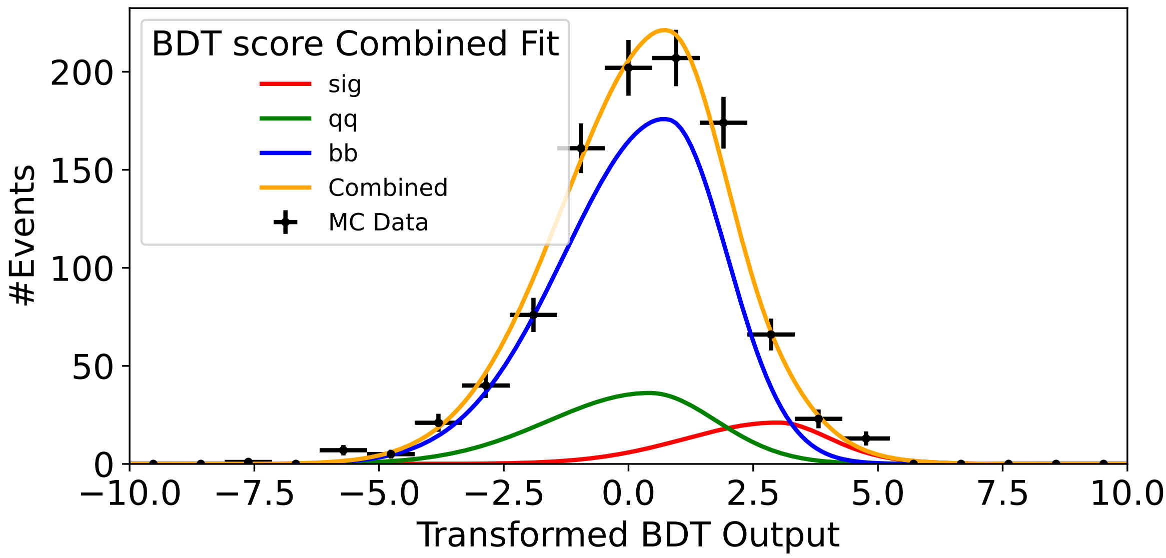

Figure 6.14.: Fit Projection of the transformed BDT output for the fit on the combined Belle and Belle II dataset. The Signal, BB background and continuum background yields are shown in red, blue and green, respectively. The sum of the event yields is shown in yellow, the MC data is shown in black.

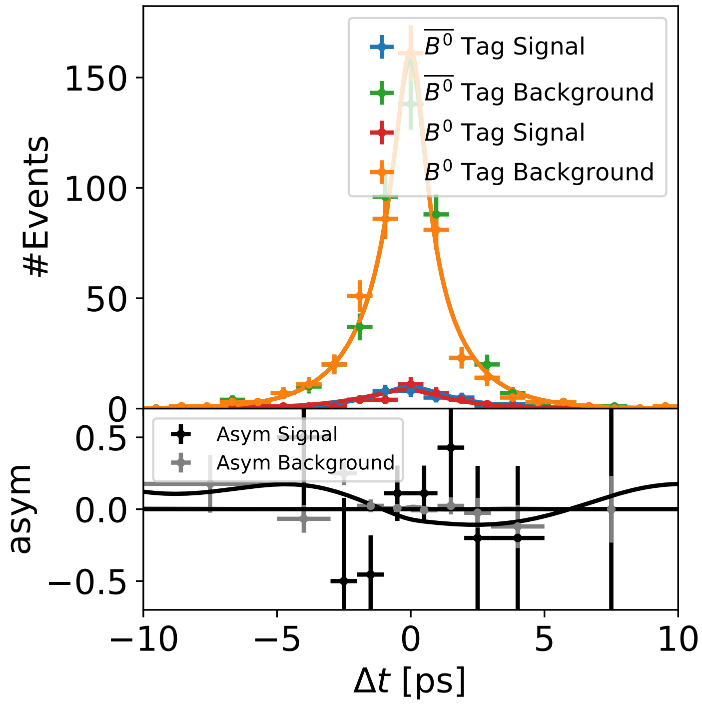

Figure 6.15.: Fit Projection of Δt for the fit on the combined Belle and Belle II dataset, separated into signal and background events for an B^0 and $\overline{B^0}$ tag, respectively.

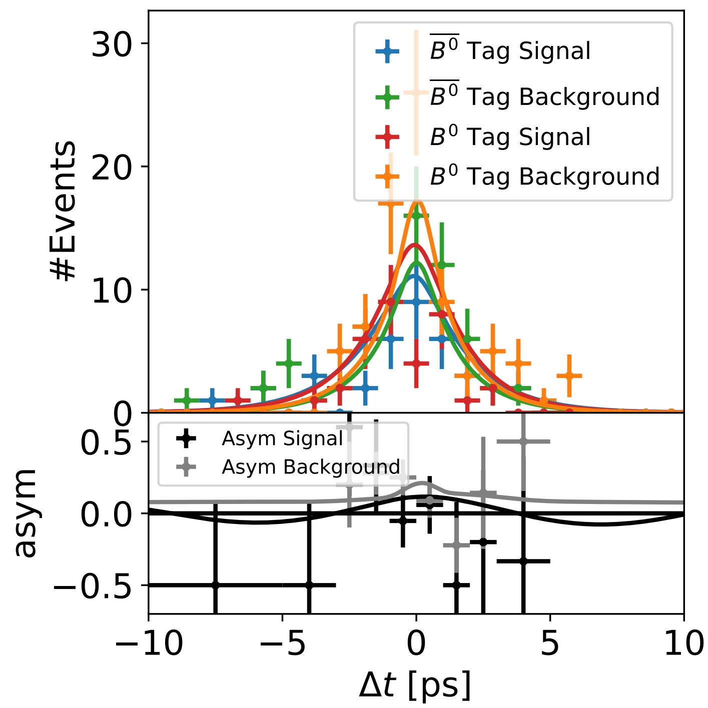

Figure 6.16.: Fit Projection of Δt in the signal region for the fit on the combined Belle and Belle II dataset, separated into signal and background events for an B^0 and $\overline{B^0}$ tag, respectively.

7. Discussion

7.1. Event Selection

In this section, the set up and performance of the event selection, especially regarding the MVA based part of the selection, are discussed.

7.1.1. MVA Based Event Selection

The signal region is considered to be $M_{\mathrm{bc}} > 5.27$ GeV and $|\Delta E| < 0.05$ GeV. In an early stage of the analysis, the idea was to use two BDTs per decay channel: one for the signal vs BB background separation and one for the signal vs continuum background separation. The cuts on the BDT output scores were then optimized simultaneously with respect to the signal significance in the signal region, where the significance is defined as $S = \frac{s}{\sqrt{s+b}}$, where S is referring to the signal significance and s and b are the amounts of signal and background events within the signal region, respectively. This strategy is illustrated in Figure 7.1 for Belle and 7.2 for Belle II.

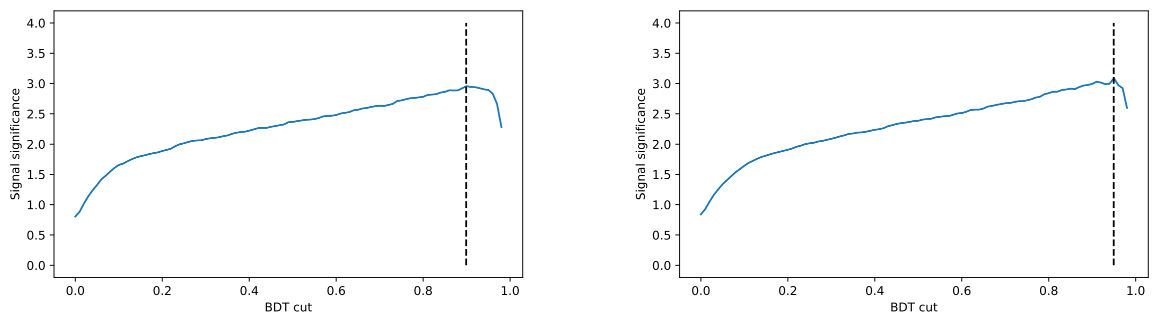

Figure 7.1.: Signal Significance depending on the applied cut on the BDT output for Belle ee (left) and Belle $\mu\mu$ (right), the blue line indicates the signal significance in the signal region, the black dotted line shows the optimal cut with respect to the significance.

However, this approach proved to be not optimal due to a very low signal efficiency of only 50 % for the BDT cut due to the very tight cut according to the signal significance

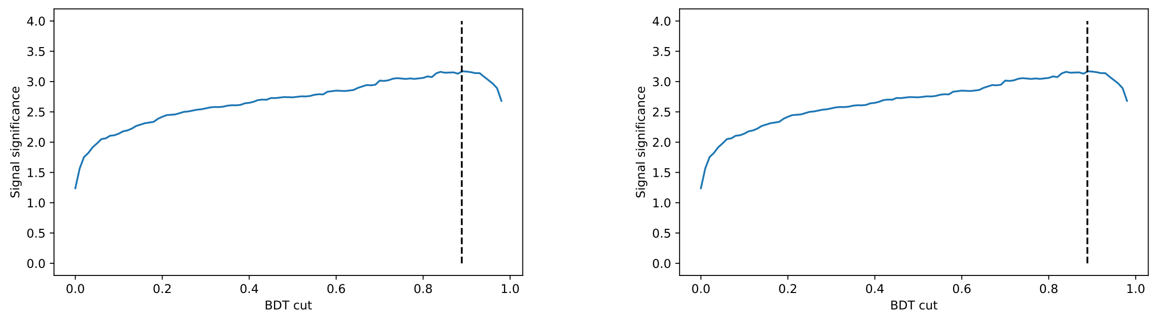

Figure 7.2.: Signal Significance depending on the applied cut on the BDT output for Belle II *ee* (left) and Belle II $\mu\mu$ (right), the blue line indicates the signal significance in the signal region, the black dotted line shows the optimal cut with respect to the significance.

optimization and thus a decreased sensitivity of the fit on S_{CP} and C_{CP}. Thus, to increase the sensitivity of the fit, the strategy to optimize the cut on the BDT score according to the signal significance was abandoned to favour a simple BDT set up with a loose cut on the BDT output which results in a higher signal statistic. Therefore the event selection presented in this Thesis consists of a simple BDT trained for signal vs background separation followed by a loose cut on the output at 0.6, the value of the cut was chosen such that the signal efficiency of the cut is 90%-95%, dependent on the decay channel, with a background rejection rate of approximately 90%, thus allowing to remove a huge proportion of the background without a significant loss of signal.

7.1.2. Peaking Background

Figures 7.3 and 7.4 show the distributions of M_{bc} for Belle and Belle II for $|\Delta E| < 0.05$ GeV and $O' \in (1.0, 8.0)$. For Belle II, the signal PDF shape is clearly distinguishable from the background PDF shape and the signal yield significantly exceeds the background yield around 5.27 GeV. For Belle however, the signal-to-background ratio is noticeably lower compared to Belle II and the BB background exhibits a peak around 5.27 GeV, where the signal peak is located as well. This peak in the BB background is not modeled by the background PDF. Tests on four times the available Luminosity of Belle and Belle II show no bias in the fit however. Figures 7.5 and 7.6 show the distribution of M_{bc} for Belle separately for 0.00 GeV $< \Delta E <$ 0.05 GeV and -0.05 GeV $< \Delta E <$ 0.00 GeV. The distributions indicate that the peak in the BB background occurs for both regions of ΔE.

Figure 7.7 shows the M_{bc} distribution of the BB background for Belle differentiated into different components based on lepton and K_{S} identification. Especially for the Belle $\mu\mu$

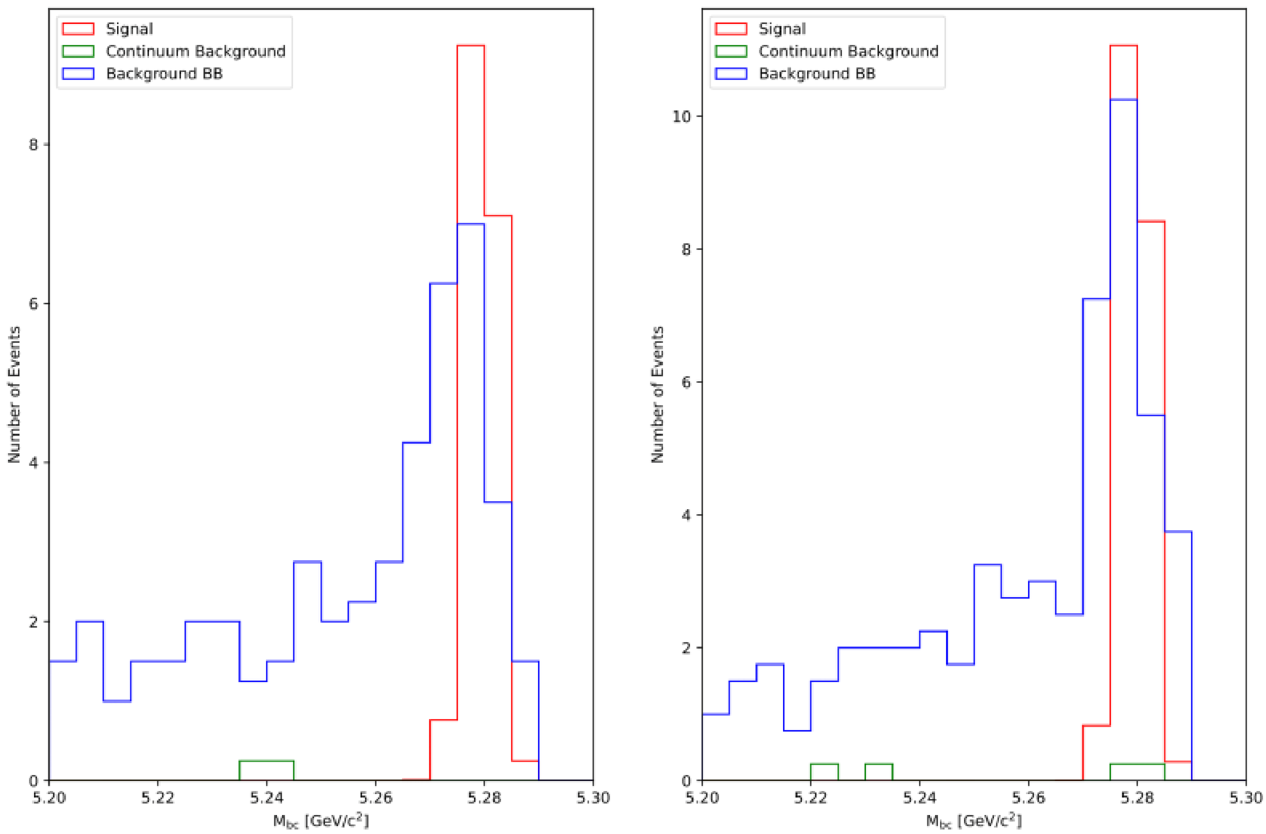

Figure 7.3.: M_{bc} distribution for Belle ee (left) and Belle $\mu\mu$ (right) for $|\Delta E| < 0.05$ GeV and $O' \in (1.0, 8.0)$. The signal is shown in red, the BB and continuum backgrounds are shown in blue and green, respectively.

decay channel the peak within the background mostly consists of events containing fake leptons or fake K_{S}. Fake leptons mostly originate from misIDs on charged pions whereas fake K_{S} stem from a track mismatch during the reconstruction of the K_{S}. Attempts to reduce the contribution of fake lepton and fake K_{S} events by applying cuts on PID variables have proven unsuccessfully due to high signal rejection rates. As a test, a likelihood fit was performed on ΔE, the transformed BDT output score and Δt with a cut requiring $M_{\mathrm{bc}} > 5.27$ GeV to avoid fitting the peaking background. The sensitivity on S_{CP} however decreased from 0.49 to 0.60, proving that the removal of M_{bc} as a fit dimension does not improve the sensitivity on the CPV parameters.

Figures 7.8 and 7.9 show the ΔE distributions for Belle and Belle II for $M_{\mathrm{bc}} > 5.27$ GeV and $O' \in (1.0, 8.0)$. The PDF shapes of signal and background are clearly distinguishable and the signal yield exceeds the background yield around 0.0 GeV.

Figures 7.10 and 7.11 show the distribution of the transformed BDT output score for

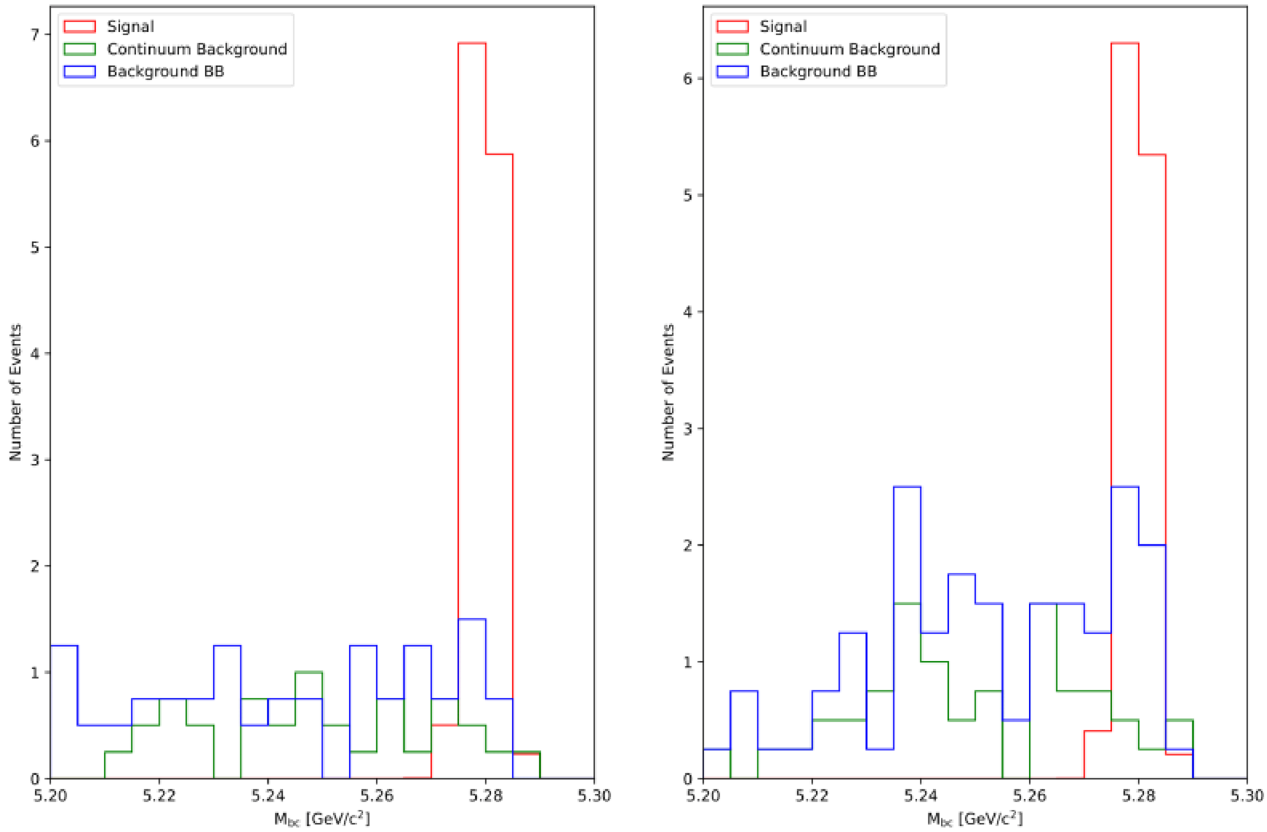

Figure 7.4.: M_{bc} distribution for Belle II ee (left) and Belle II $\mu\mu$ (right) for $|\Delta E| < 0.05$ GeV and $O' \in (1.0, 8.0)$. The signal is shown in red, the BB and continuum backgrounds are shown in blue and green, respectively.

Belle and Belle II. As for ΔE, the PDF shapes of signal and background are clearly distinguishable from each other.

To conclude, the distributions of ΔE and the transformed BDT output score as well as M_{bc} for Belle II look healthy and thus should provide no issue for the likelihood fit. The peak in the BB background for M_{bc} for Belle is not modeled by the fit, but tests on four times the available luminosity do not show any bias. Unfortunately, neither the removal of M_{bc} as a fit dimension nor attempts to reduce the peak in the background without removing to much signal have proven successful with respect to increasing the sensitivity of S_{CP} and C_{CP}.

7.2. Likelihood Fit

Over the course of the analysis, different fit strategies have been tried to study their impact on the sensitivity of the fit. A simple variation of the cut imposed on the BDT output

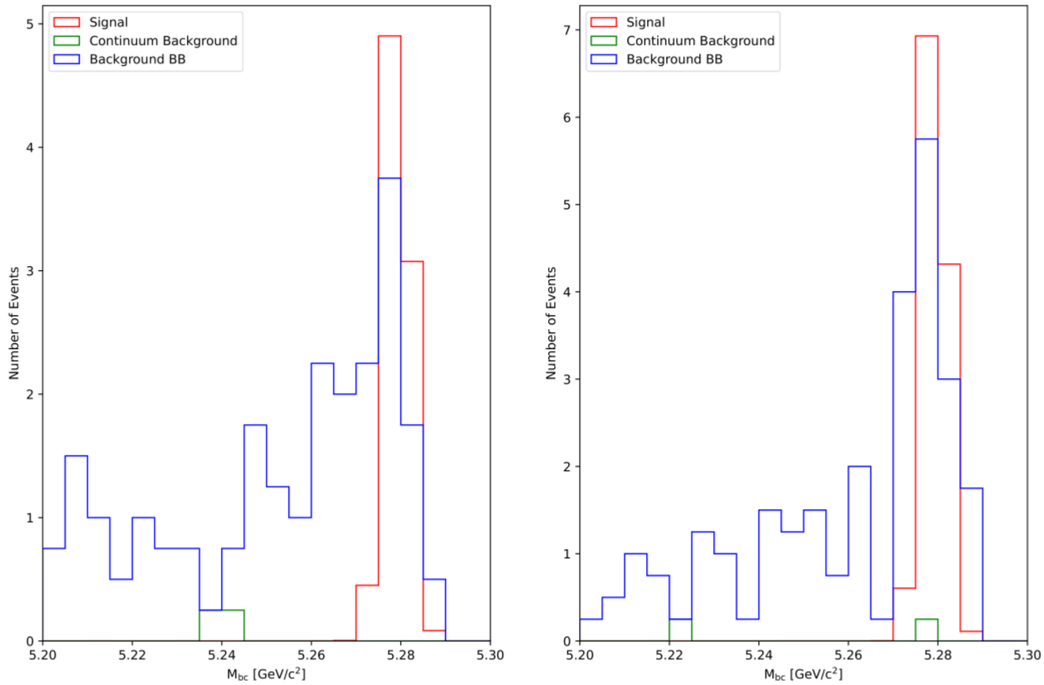

Figure 7.5.: M_{bc} distribution for Belle ee (left) and Belle $\mu\mu$ (right) for 0.00 GeV $< \Delta E <$ 0.05 GeV and $O' \in (1.0, 8.0)$. The signal is shown in red, the BB and continuum backgrounds are shown in blue and green, respectively.

score has proven to not affect the fit result to a noticeable degree. Another approach was to apply a BDT output cut which is dependent on the $|qr|$ bin. This set up can exploit the higher signal-to-background ratio in higher $|qr|$ bins by imposing tighter cuts on the $|qr|$ bins 0-4 while keeping the cuts loose for the $|qr|$ bins 5 and 6. Various combinations of cuts however showed no significant effect on the fit performance.
In Chapter 6.1 it was shown that the fit projections of the PDF shapes agree with the expectation for M_{bc}, ΔE and Δt as well as the background PDFs for the transformed BDT output whereas the projections are off for the signal PDF of the transformed BDT output. While it does not impose a bias to the fit, it could still lead to an increased uncertainty of the fit result. As a test, the BDT output score was removed as a fit dimension, but it led to a decreased sensitivity on S_{CP} and C_{CP}. To solve this issue, another transformation of the BDT output score was tested, where the BDT output is not transformed into a bifurcated Gaussian, but instead the signal is transformed to have a linear distribution while the background is distributed double exponentially. The PDF shapes for the transformed BDT output can be seen in Figure 7.12, 7.13 and 7.14.

While the PDF shapes look healthy, the sensitivity of the fit to extract S_{CP} and C_{CP} has not improved.

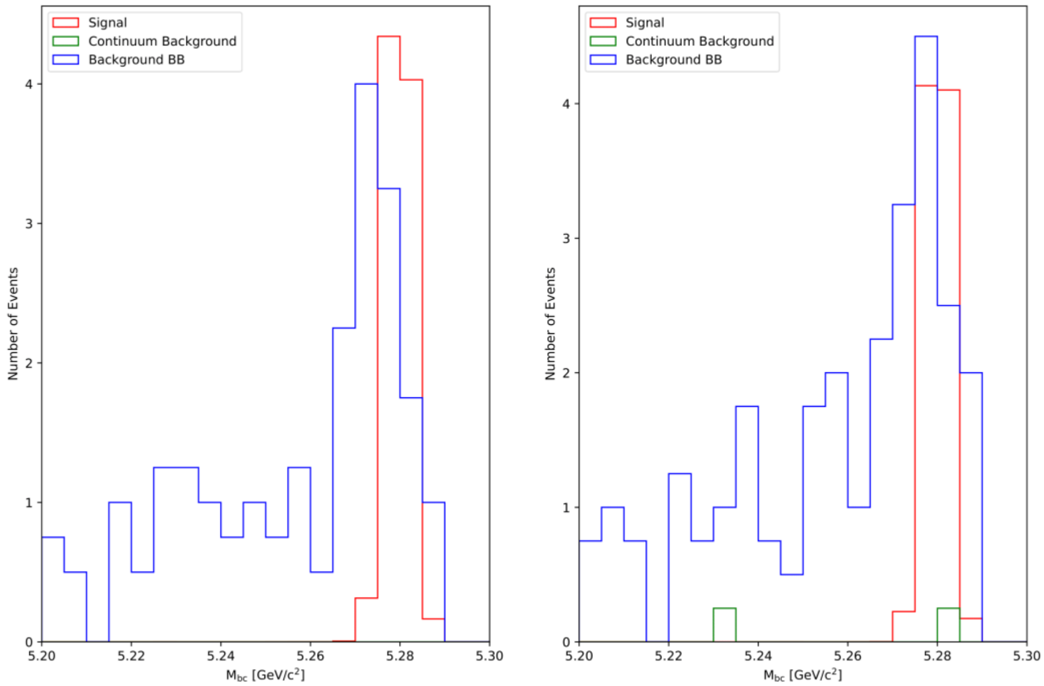

Figure 7.6.: $M_{\rm bc}$ distribution for Belle ee (left) and Belle $\mu\mu$ (right) for -0.05 GeV $< \Delta E <$ 0.00 GeV and $O' \in (1.0, 8.0)$. The signal is shown in red, the BB and continuum backgrounds are shown in blue and green, respectively.

Other approaches were the use of a polynomial PDF shape for the background PDFs of ΔE as well as to allow the continuum background yields to be free floating during the fit. Both attempts however led to a decrease of the fit sensitivity.
To summarise, the fit seems to have either a worse sensitivity or to be not significantly affected by applying different fit strategies. Hence, it seems that with the current amount of data at Belle and Belle II, a precision significantly better than 0.5 on $S_{\rm CP}$ is not achievable independent of the fit strategy. Other sources of systematic uncertainties are assumed to be of the order of $\mathrm{O}(10^{-2})$, thus significantly smaller than the statistical uncertainty. Therefore other systematic uncertainties are not addressed in the context of this Thesis.

7.3. Outlook

The main contribution to the uncertainty on $S_{\rm CP}$ and $C_{\rm CP}$ is the low statistics of the signal decay. Thus, in order to improve the performance of the fit, more data needs to be available. The result in this Thesis is obtained by using the full currently accessible dataset of Belle and Belle II, corresponding to 711 fb^{-1} at Belle and 486 fb^{-1} fb at Belle II.

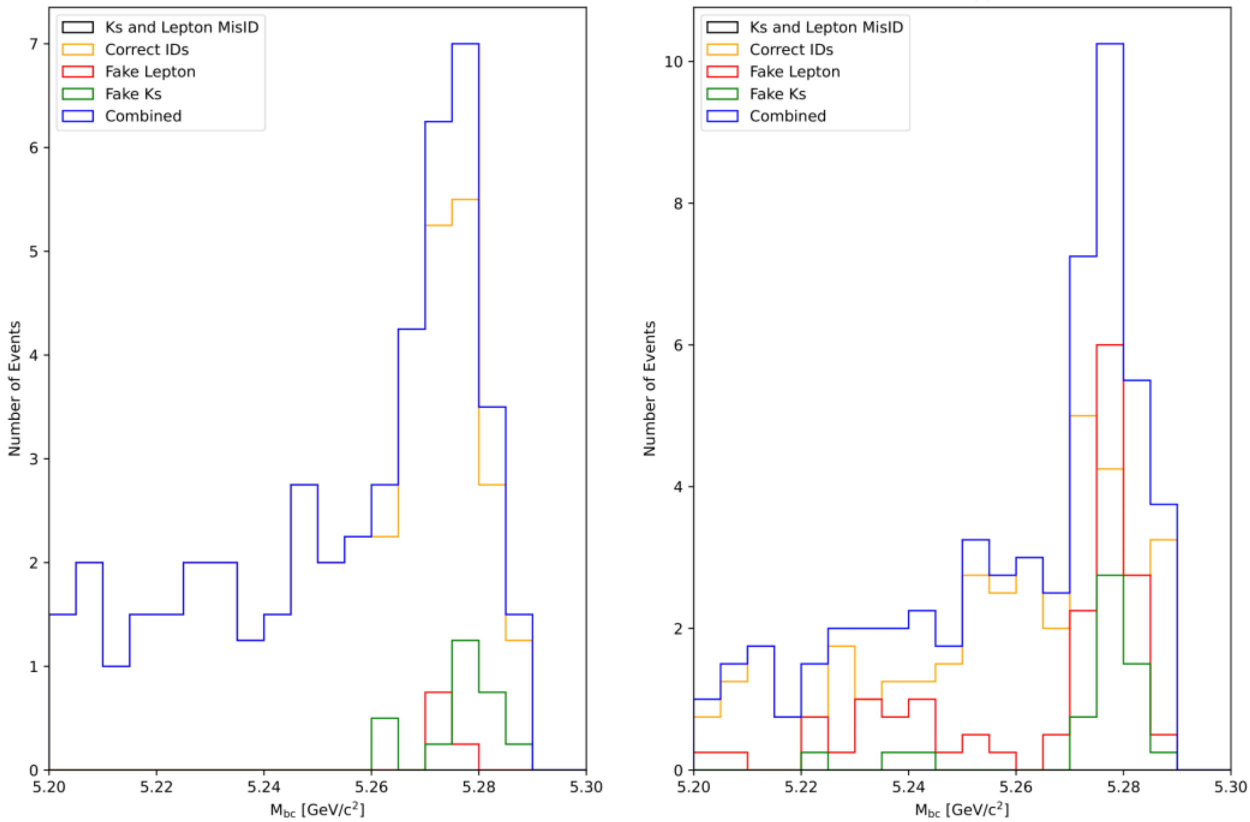

Figure 7.7.: M_{bc} distribution for the BB background for Belle ee (left) and Belle $\mu\mu$ (right) for $|\Delta E| < 0.05$ GeV and $O' \in (1.0, 8.0)$, divided into different components based on Lepton and K_{S} IDs.

As a test, a fit is performed assuming the luminosity of Belle II is four times the current accessible, corresponding to 1944 fb^{-1}. The sensitivity on S_{CP} and C_{CP} was improved to 0.39 and 0.25, respectively. The SM expectation for S_{CP} is 0.7, thus a sensitivity of at least 0.23 is needed to claim evidence for TDCPV in the $B \to K_S \ell\ell$ decay channel. To achieve this sensitivity, a dataset of approximately four and a half times the size of the current accessible is needed given that the sensitivity scales with $\sim \frac{1}{\sqrt{L}}$ with L being the available Luminosity.

To search for deviations from the SM however, a higher sensitivity is needed, which will be difficult to archive for Belle II alone. Thus, in order to nail down the Wilson Coefficient C_9 and search for deviations from the SM, a combination of measurements of various decay channels involving $b \to s\ell^+\ell^-$ transitions needs to be performed by LHCb and Belle II.

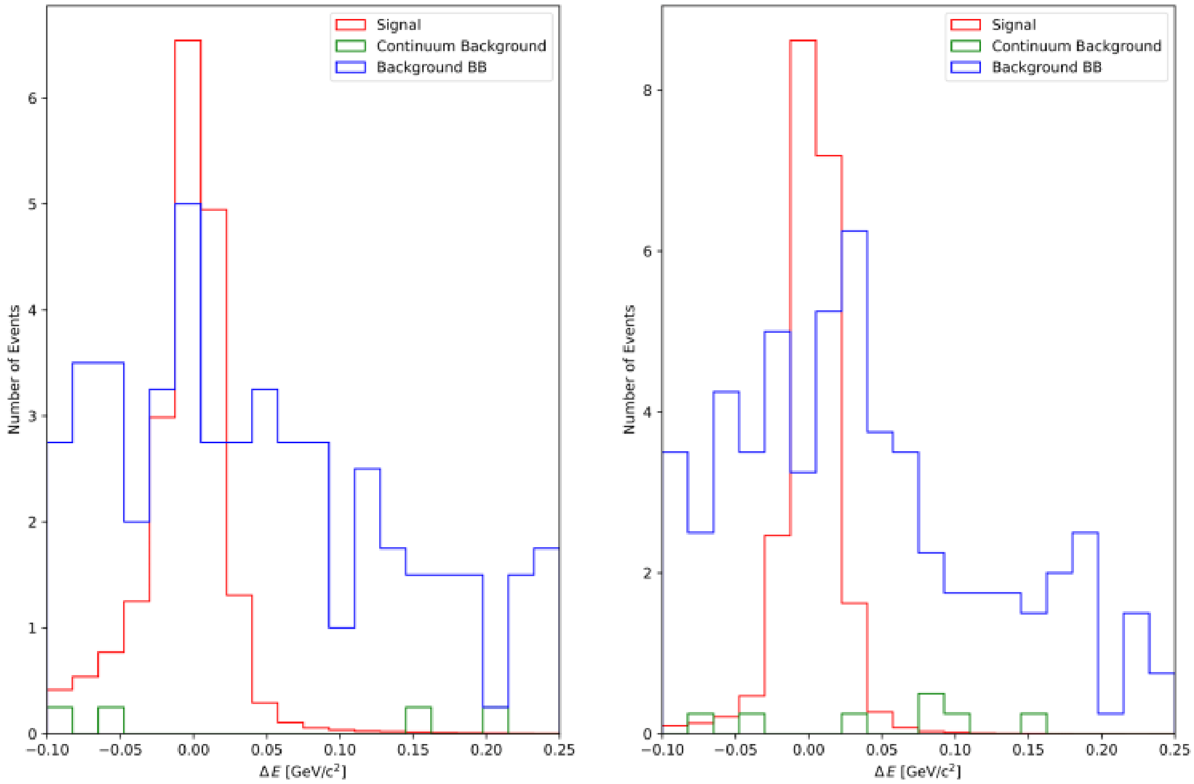

Figure 7.8.: ΔE distribution for Belle ee (left) and Belle $\mu\mu$ (right) for $M_{\mathrm{bc}} > 5.27$ GeV and $O' \in (1.0, 8.0)$. The signal is shown in red, the BB and continuum backgrounds are shown in blue and green, respectively.

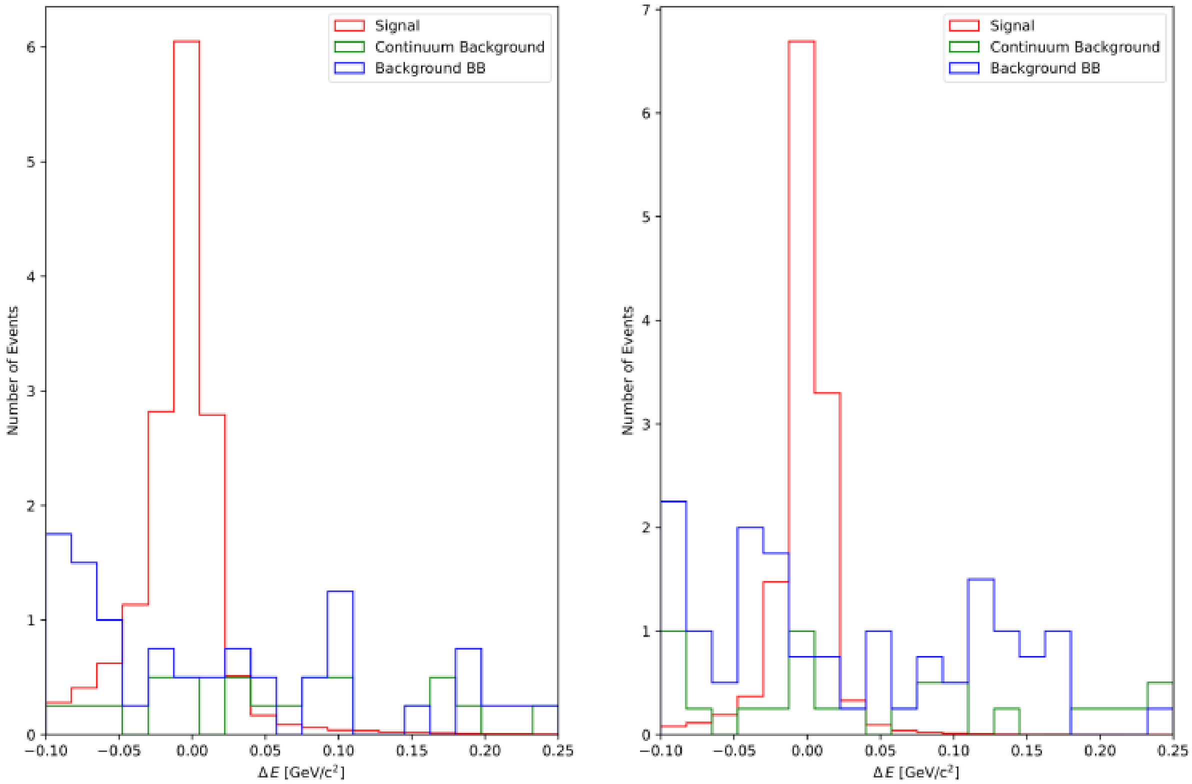

Figure 7.9.: ΔE distribution for Belle II ee (left) and Belle II $\mu\mu$ (right) for $M_{\text{bc}} > 5.27$ GeV and $O' \in (1.0, 8.0)$. The signal is shown in red, the BB and continuum backgrounds are shown in blue and green, respectively.

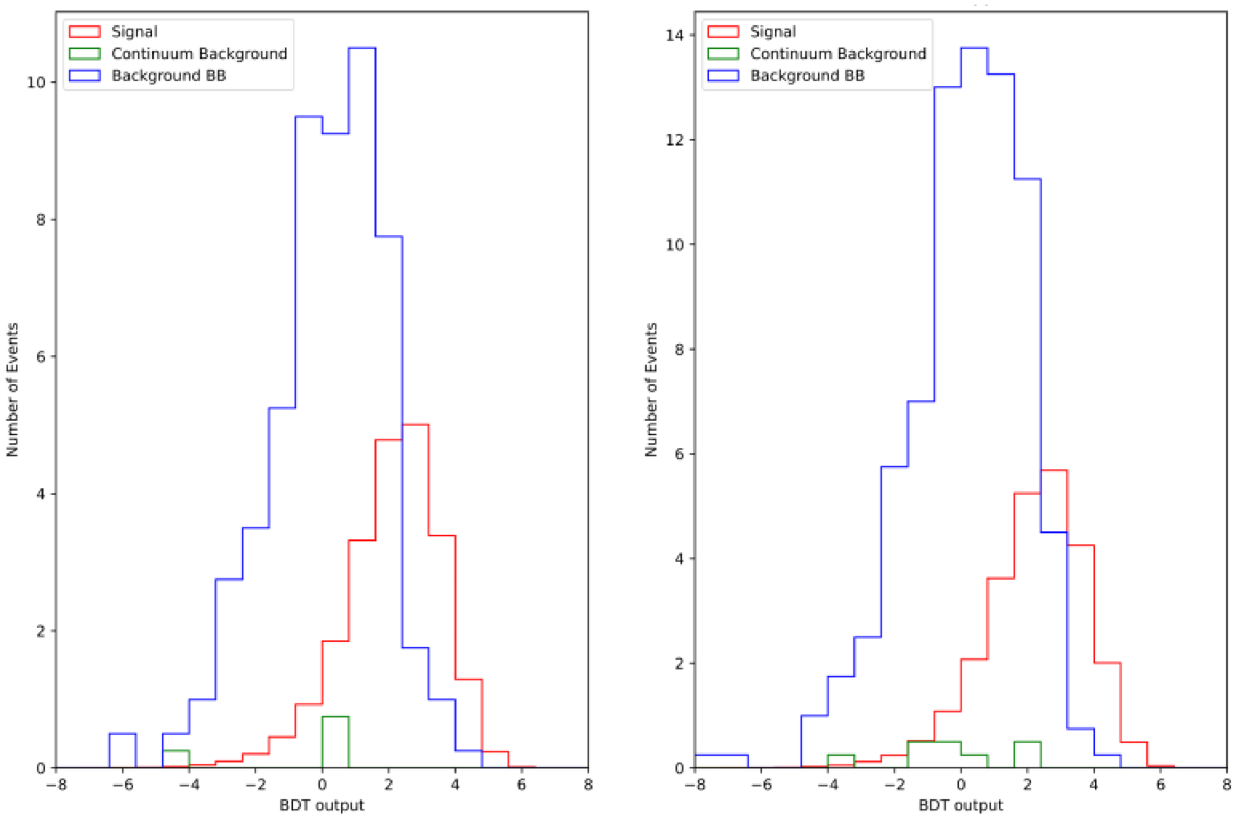

Figure 7.10.: Transformed BDT output score distribution for Belle ee (left) and Belle $\mu\mu$ (right) for $M_{\rm bc} > 5.27$ GeV and $|\Delta E| < 0.05$ GeV. The signal is shown in red, the BB and continuum backgrounds are shown in blue and green, respectively.

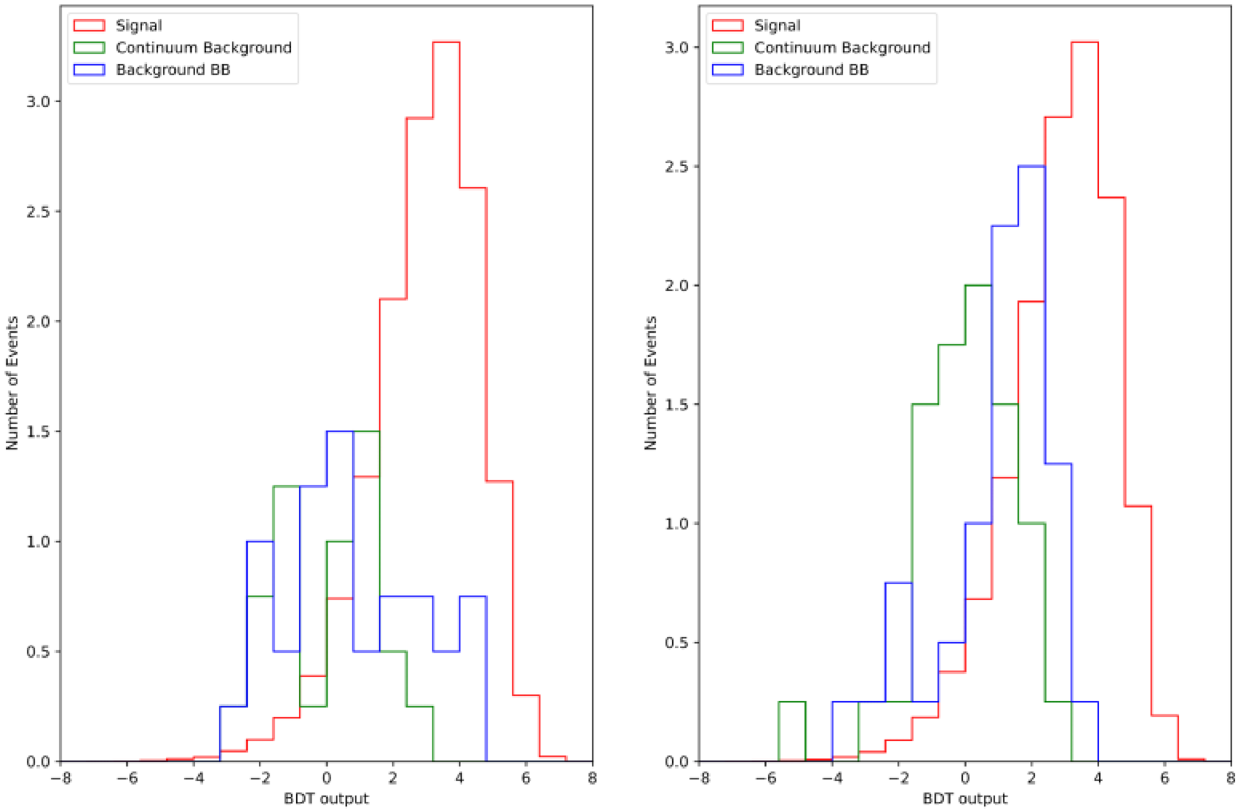

Figure 7.11.: Transformed BDT output score distribution for Belle II ee (left) and Belle II $\mu\mu$ (right) for $M_{\rm bc} > 5.27$ GeV and $|\Delta E| < 0.05$ GeV. The signal is shown in red, the BB and continuum backgrounds are shown in blue and green, respectively.

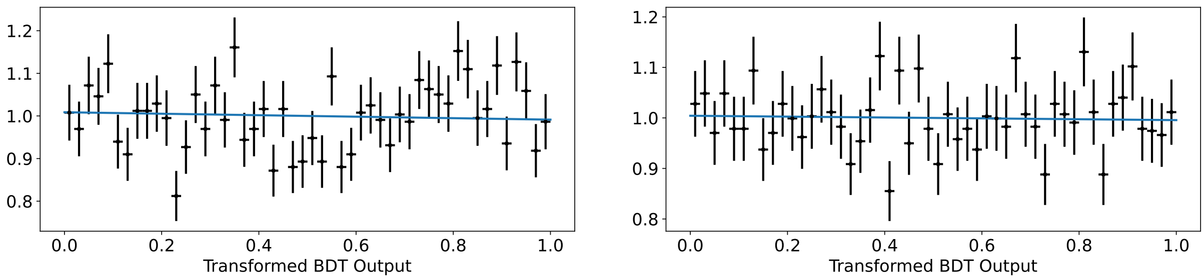

Figure 7.12.: Signal PDF shape for the transformed BDT output for Belle (left) and Belle II (right) in the $|qr|$ bin 0. The black dots indicate the MC data, the blue line is the result of the likelihood fit.

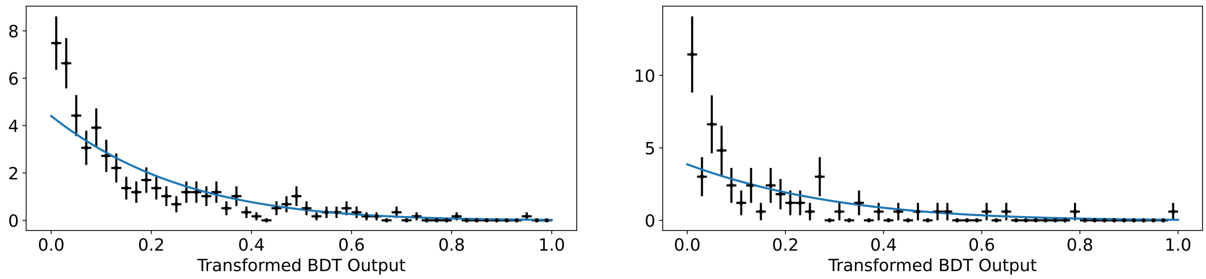

Figure 7.13.: BB Background PDF shape for the transformed BDT output for Belle (left) and Belle II (right) in the $|qr|$ bin 0. The black dots indicate the MC data, the blue line is the result of the likelihood fit.

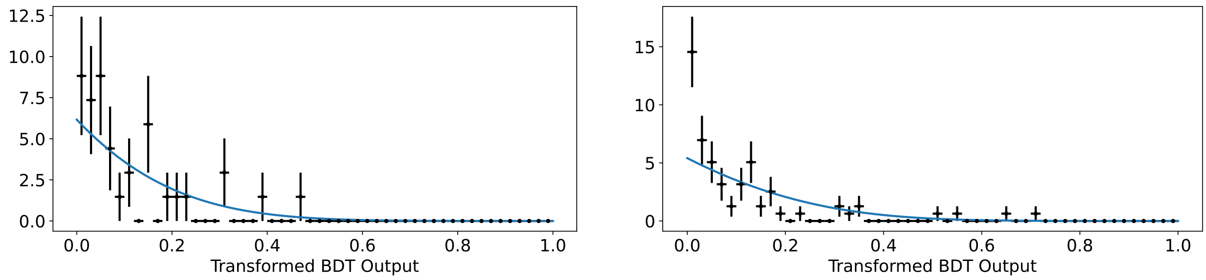

Figure 7.14.: Continuum Background PDF shape for the transformed BDT output for Belle (left) and Belle II (right) in the $|qr|$ bin 0. The black dots indicate the MC data, the blue line is the result of the likelihood fit.

Bibliography

[1] A.D. Sakharov. “Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe”. In: *Pisma Zh.Eksp.Teor.Fiz.* ***5****, 32-35* (1967).

[2] Makoko Kobayashi and Toshihide Maskawa. “CP-Violation in the Renormalizable Theory of Weak Interaction”. In: *Progress of Theoretical Physics, Volume* ***49****, Issue 2* (1973).

[3] Sebastien Descotes-Genona, Martin Novoa-Bruneta and K. Keri Vos. “The time-dependent angular analysis of Bd KS , a new benchmark for new physics”. In: *TUM-HEP-1275/20* (2021).

[4] The Particle Data Group. “The Review of Particle Physics (2024)”. In: *Phys. Rev. D* ***110****, 030001* (2024).

[5] A. Salam. “Weak and Electromagnetic Interactions”. In: *Conf. Proc. C* ***680519****, 367* (1968).

[6] S. L. Glashow. “Partial-symmetries of weak interactions”. In: *Nucl. Phys.* ***22(4)****, 579* (1961).

[7] S. Weinberg. “Conceptual Foundations of the Unified Theory of Weak and Electromagnetic Interactions”. In: *Rev. Mod. Phys.* ***52****, 515-523* (1980).

[8] S. Weinberg. “A Model of Leptons”. In: *Phys. Rev. Lett.* ***19****, 1264* (1967).

[9] Peter W. Higgs. “Broken Symmetries and the Masses of Gauge Bosons”. In: *Phys. Rev. Lett.* ***13****, 508-509* (1964).

[10] Wikipedia. The Free Encyclopedia. *Wikipedia: Standard Model, Online [Accessed 9th of August 2025].* 2025.

[11] Chien-Shiung Wu. “Experimental Test of Parity Conservation in Beta Decay*”. In: *Phys. Rev.* ***105****, 1413* (1957).

[12] Nicola Cabibbo. “Unitary Symmetry and Leptonic Decays”. In: *Phys. Rev. Lett.* ***10****, 531* (1963).

[13] J. Charles et al. “Current status of the Standard Model CKM fit and constraints on $\Delta F = 2$ New Physics”. In: *Phys. Rev. D* ***91****, 073007* (2015).

[14] J.H. Christenson et al. "Evidence for the 2π decay of the $K_2^0 Meson$". In: *Phys Rev. Lett.* ***13*** (1964).

[15] J.R. Batley et al. "A precision measurement of direct CP violation in the decay of neutral kaons into two pions". In: *Physics Letters B* ***544****, 97-112* (2002).

[16] KTeV-Collaboration. "Measurements of Direct CP Violation, CPT Symmetry, and Other Parameters in the Neutral Kaon System". In: *Phys. Rev. D* ***67****, 012005* (2002).

[17] Belle-Collaboration. "Observation of Large CP Violation in the Neutral B Meson System". In: *Phys. Rev. Lett.* ***87****, 091802* (2001).

[18] BaBar-Collaboration. "Measurement of the CP-Violating Asymmetry Amplitude sin2β". In: *Phys. Rev. Lett.* ***89****, 201802* (2002).

[19] Douglas Clowe et al. "A Direct Empirical Proof of the Existence of Dark Matter". In: *ApJ* ***648****, L109* (2006).

[20] A.V. Gladyshev and D.I. Kazakov. "Supersymmetry and LHC". In: *Phys. Atom. Nuclei* ***70****, 1553* (2007).

[21] T. Hurth, F. Mahmoudi, and S. Neshatpour. "B anomalies in the post-$R_{K^{(*)}} era$". In: *Phys. Rev. D* ***108****, 115037* (2023).

[22] SuperKEKB Collaboration. "SuperKEKB Collider". In: *Nucl. Instrum. Meth. A* ***907****, 188-199* (2018).

[23] Yukiyoshi Ohnishi et al. "Accelerator design at SuperKEKB". In: *PTEP 2013* ***03A011*** (2013).

[24] KEK. "KEKB B-Factory Design Report". In: *KEK Report No.* ***95-7*** (1995).

[25] Belle-II Collaboration. "Belle II Technical Design Report". In: *KEK-REPORT-* ***2010-1*** (2010).

[26] Belle Collaboration. "The Belle Detector". In: *Nucl. Instrum. Meth. A* ***479****, 117-232* (2002).

[27] Dmitry Matvienko. "The Belle II experiment: status and physics program". In: *EPJ Web of Conferences* ***191****, 02010* (2018).

[28] T. Hara, T. Kuhr, and Y. Ushiroda. "Belle II Coordinate System and Guideline of Belle II Numbering Scheme". In: *https://indico.mpp.mpg.de/event/2308/contributions/4092/attachments/3414/3799/Belle2NumberingScheme.pdf, Online [Accessed 20th of October 2025]* (2011).

[29] F. Abudinén et al. "B-flavor tagging at Belle II". In: *Eur. Phys. J. C **82**, 283* (2022).

[30] F. Abudinén et al. "Measurement of the B^0 lifetime and flavor-oscillation frequency using hadronic decays reconstructed in 2019-2021 Belle II data". In: *Phys. Rev. D **107**, L091102* (2023).

[31] I. Olkin and A.R. Sampson. "Multivariate Analysis: Overview". In: *International Encyclopedia of the Social Behavioral Sciences, 10240-10247* (2001).

[32] J. Ross Quinlan. *Learning Efficient Classification Procedures and Their Application to Chess End Games, 463 - 482.* 1. Morgan Kaufmann, 1983.

[33] Wikipedia. The Free Encyclopedia. *Wikipedia: Decision tree learning, Online [Accessed 20th of October 2025].* 2025.

[34] Bryon P. Roe et al. "Boosted decision trees, an alternative to artificial neural networks". In: *Nucl. Instrum. Meth. A **543** 2-3, 577-584* (2005).

[35] J.R. Quinlan. "Induction of Decision Trees". In: *Mach. Learn. **1**, 1, 81-106* (1986).

[36] Tianqi Chen and Carlos Guestrin. "XGBoost: A Scalable Tree Boosting System". In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785-794* (2016).

[37] Richard J. Rossi. *Mathematical Statistics: An Introduction to Likelihood Based Inference.* 1. John Wiley & Sons, Inc, 2018.

[38] F. James and M. Roos. "Minuit - a system for function minimization and analysis of the parameter errors and correlations". In: *Computer Physics Communications **10**, 343-367* (1975).

[39] D.J. Lange. "The EvtGen particle decay simulation package". In: *Nucl. Instrum. Meth. A **462**, 152-155* (2001).

[40] R. Brun et al. "GEANT3". In: *CERN Report No. **DD/EE/84-1*** (1984).

[41] Moritz Gelb et al. "B2BII - Data conversion from Belle to Belle II". In: *Computing and Software for Big Science **1*** (2018).

[42] T. Kuhr et al. "The Belle II Core Software". In: *Comput. Softw. Big Sci. **3**, no.1* (2019).

[43] J. F. Krohn et al. "Global Decay Chain Vertex Fitting at B-Factories". In: *Nucl. Instrum. Meth. A **976**, 164269* (2020).

[44] Geoffrey C. Fox and Stephen Wolfram. "Observables for the Analysis of Event Shapes in e^+e^- Annihilation and Other Processes". In: *Phys. Rev. Lett. **41**, 1581* (1987).

[45] D.M. Asner et al. "Search for exclusive charmless hadronic B decays". In: *Phys. Rev. D* ***53****, 1039* (1996).

[46] Tomasz Skwarnicki. "A study of the radiative CASCADE transitions between the Upsilon-Prime and Upsilon resonances". In: *DESY Internal Report No. DESY F31-86-02* (1986).

[47] H. Albrecht et al. "Search for hadronic $b \to u$ decays". In: *Physics Letters B, Volume* ***241****, Issue 2* (1990).

[48] J.F. Gibbons and S. Mylroie. "Estimation of impurity profiles in ion-implanted amorphous targets using joined half-Gaussian distributions". In: *Appl. Phys. Lett.* ***22****, 568-569* (1973).

A. Additional Plots Event Selection

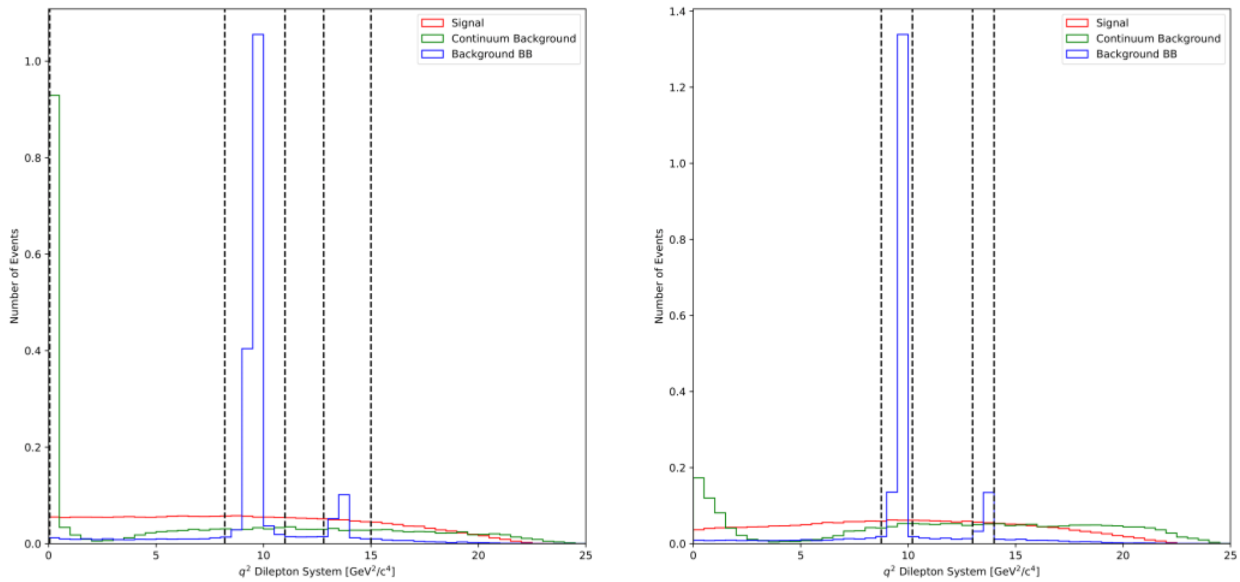

Figure A.1.: Veto applied on the invariant mass of the dilepton system for the Belle II ee decay channel (left) and the Belle II $\mu\mu$ decay channel (right). Signal events are shown in red, continuum background events in green and BB related background events are shown in blue. The cuts applied are indicated with black doted lines.

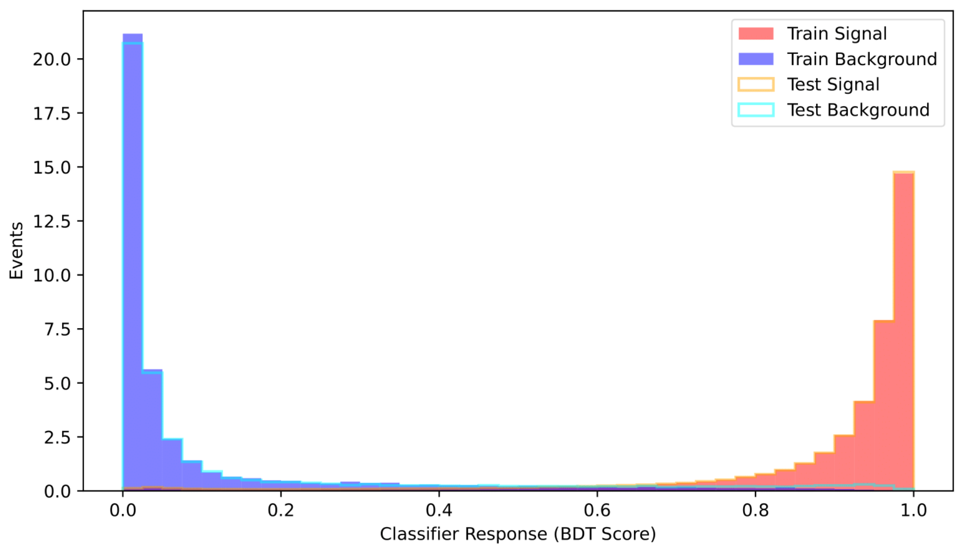

Figure A.2.: BDT output distribution for the Belle $\mu\mu$ decay channel. Shown are the distributions of the training and the test scores. The events are normalised such that the area of the histogram corresponds to 1.

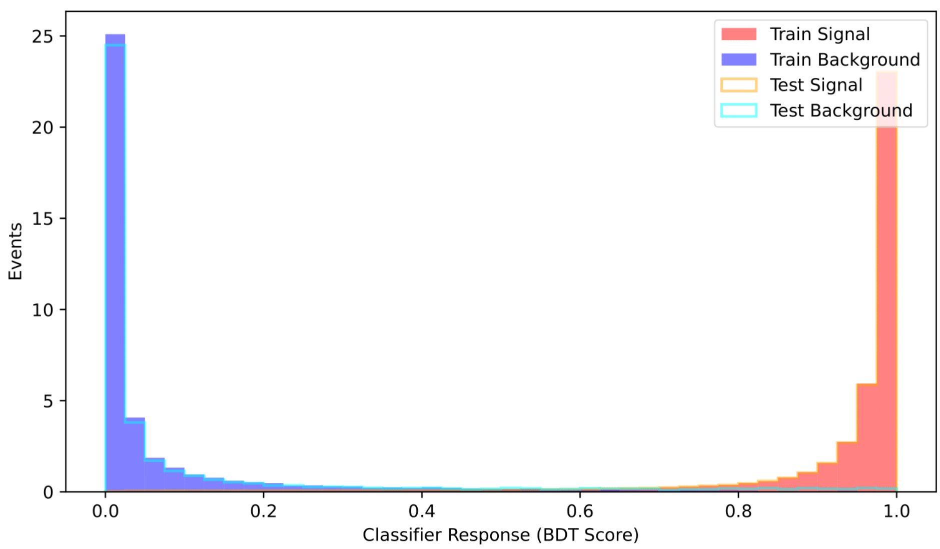

Figure A.3.: BDT output distribution for the Belle II ee decay channel. Shown are the distributions of the training and the test scores. The events are normalised such that the area of the histogram corresponds to 1.

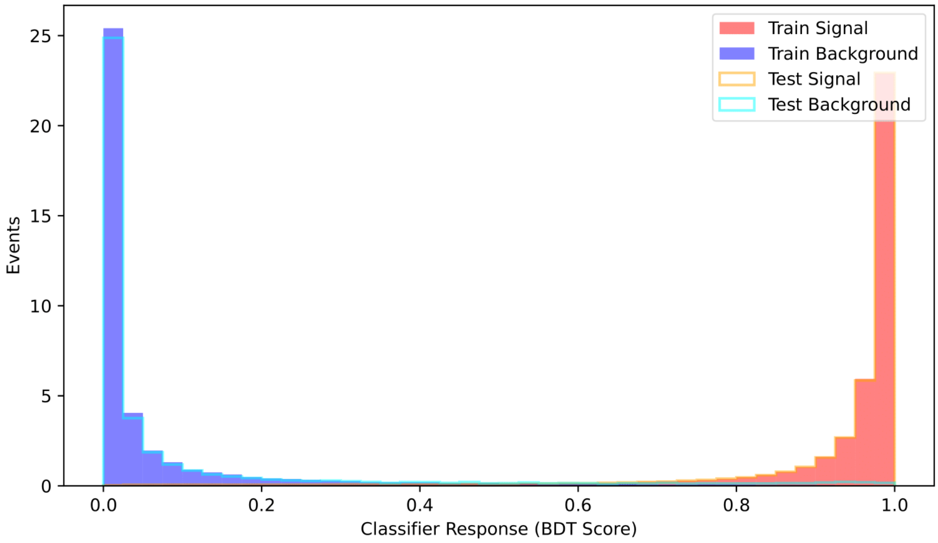

Figure A.4.: BDT output distribution for the Belle II $\mu\mu$ decay channel. Shown are the distributions of the training and the test scores. The events are normalised such that the area of the histogram corresponds to 1.

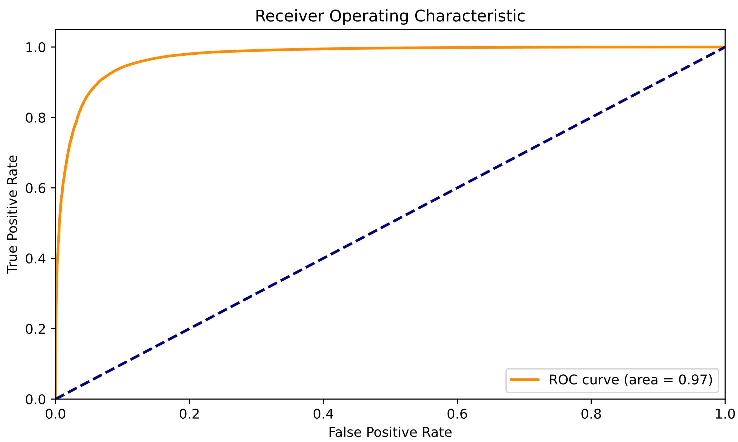

Figure A.5.: ROC curve for the Belle $\mu\mu$ decay channel, shown in orange. The AUC score equals 0.97 indicating a high separation power between signal and background events. The dotted blue line shows the hypothetical ROC curve for an AUC score of 0.5.

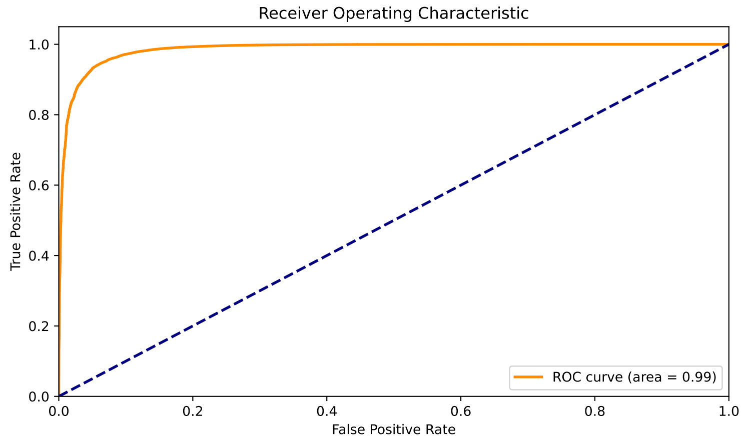

Figure A.6.: ROC curve for the Belle II *ee* decay channel, shown in orange. The AUC score equals 0.99 indicating a high separation power between signal and background events. The dotted blue line shows the hypothetical ROC curve for an AUC score of 0.5.

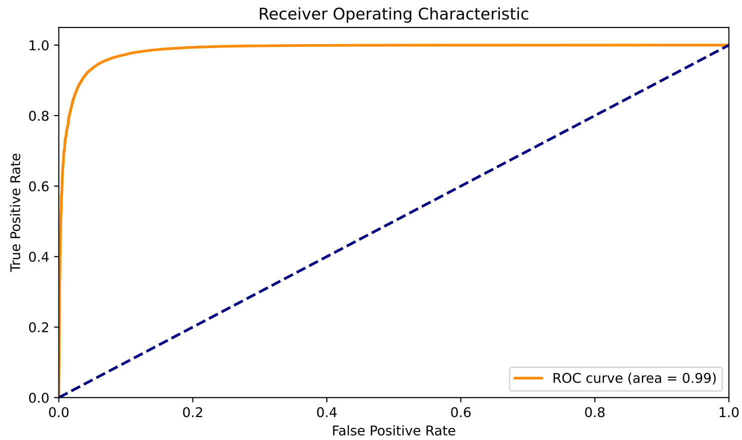

Figure A.7.: ROC curve for the Belle II $\mu\mu$ decay channel, shown in orange. The AUC score equals 0.99 indicating a high separation power between signal and background events. The dotted blue line shows the hypothetical ROC curve for an AUC score of 0.5.

B. Input Variables for BDT Training

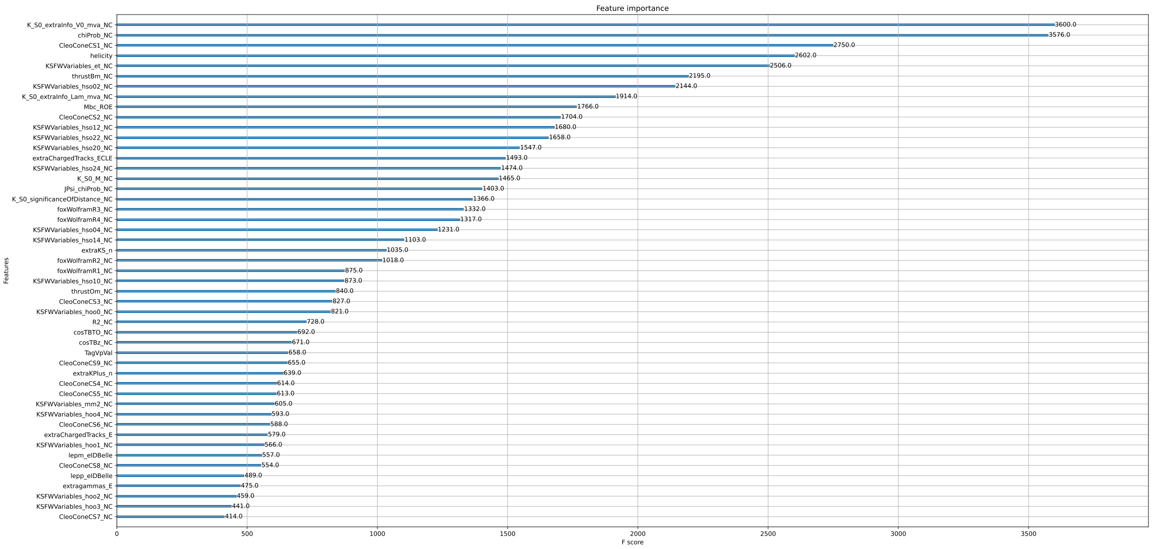

Figure B.1.: List of input variables used for the BDT training for Belle *ee* decay channel. The input variables are ordered according to their importance for the signal background separation with the F score indicating their relevance.

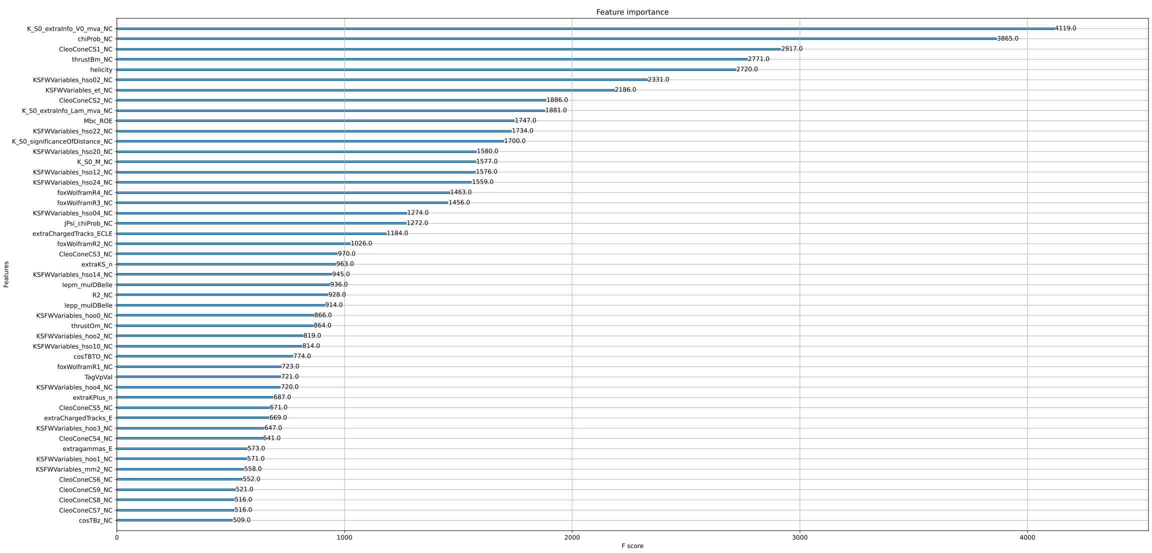

Figure B.2.: List of input variables used for the BDT training for Belle $\mu\mu$ decay channel. The input variables are ordered according to their importance for the signal background separation with the F score indicating their relevance.

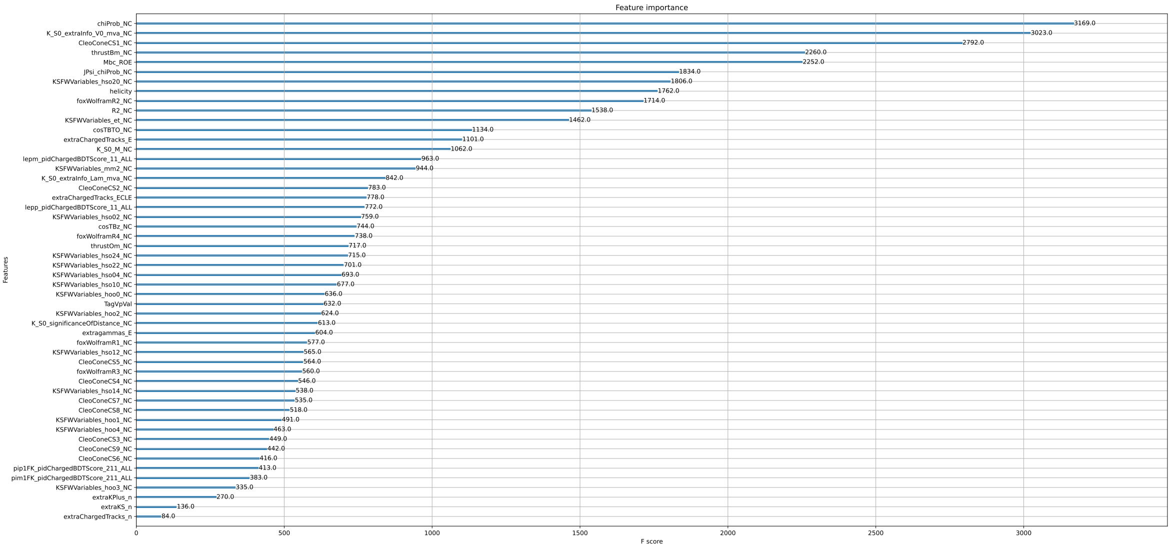

Figure B.3.: List of input variables used for the BDT training for Belle II ee decay channel. The input variables are ordered according to their importance for the signal background separation with the F score indicating their relevance.

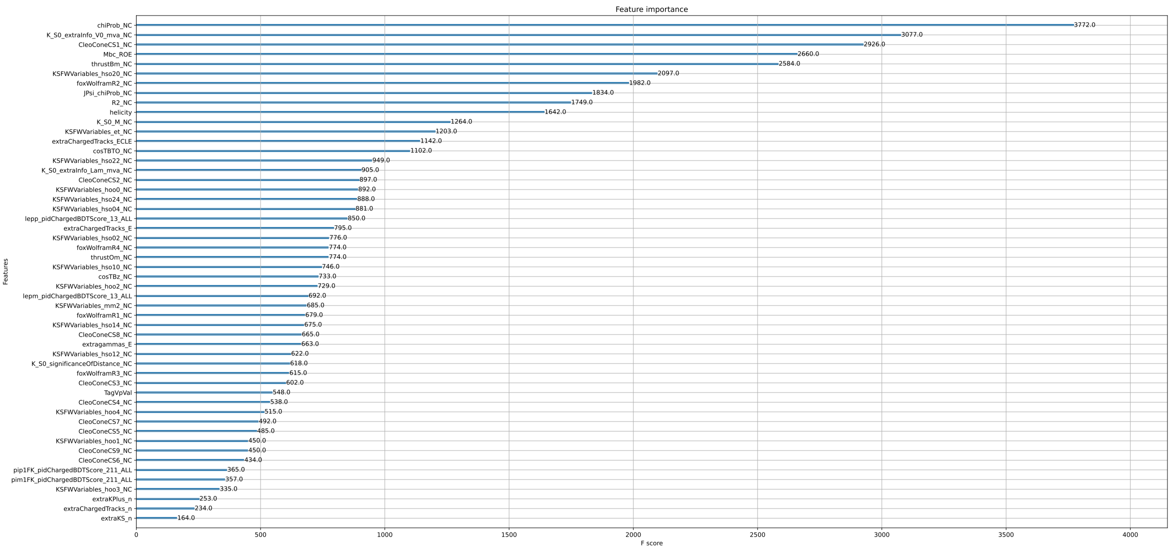

Figure B.4.: List of input variables used for the BDT training for Belle II $\mu\mu$ decay channel. The input variables are ordered according to their importance for the signal background separation with the F score indicating their relevance.

C. Additional Plots Likelihood Fit

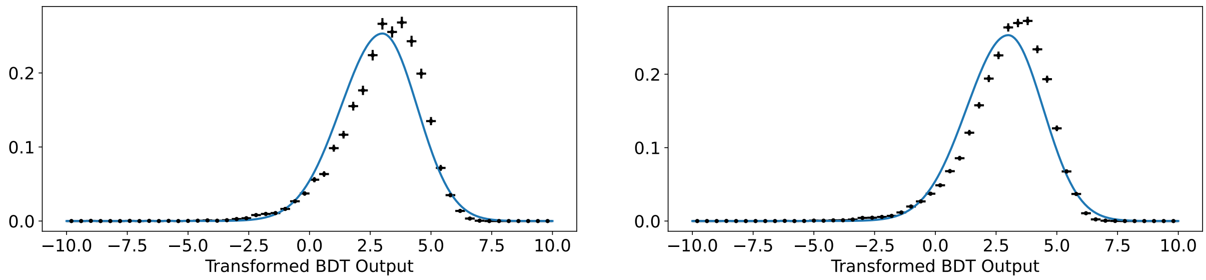

Figure C.1.: Signal PDF shape for the transformed BDT output for Belle II in the $|qr|$ bin 0 (left) and 6 (right). The black dots indicate the MC data, the blue line is the result of the likelihood fit.

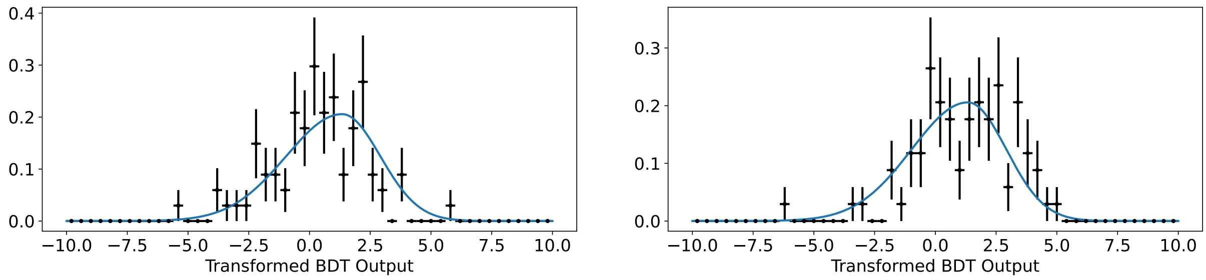

Figure C.2.: BB background PDF shape for the transformed BDT output for Belle II in the $|qr|$ bin 0 (left) and 6 (right). The black dots indicate the MC data, the blue line is the result of the likelihood fit.

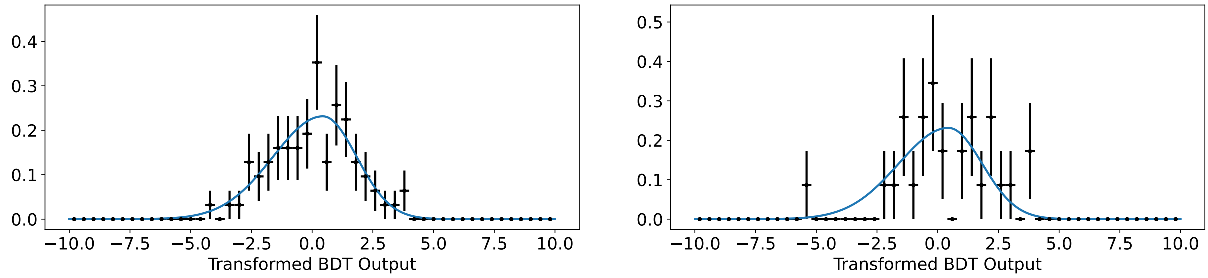

Figure C.3.: Continuum background PDF shape for the transformed BDT output for Belle II in the $|qr|$ bin 0 (left) and 6 (right). The black dots indicate the MC data, the blue line is the result of the likelihood fit.

Table C.1.: Results of the Fit on the Belle Dataset.

Variable	Result	Uncertainty	MC Expectation
S_{CP}	-0.03	0.72	0.0
C_{CP}	-0.48	0.49	0.0
Signal Events	51	13	48
BB Background Events	619	37	622
Continuum Background Events	27	0	27

Table C.2.: Results of the Fit on the Belle II Dataset.

Variable	Result	Uncertainty	MC Expectation
S_{CP}	-0.59	1.19	0.0
C_{CP}	-0.52	0.71	0.0
Signal Events	32	9	29
BB Background Events	133	22	136
Continuum Background Events	134	0	134

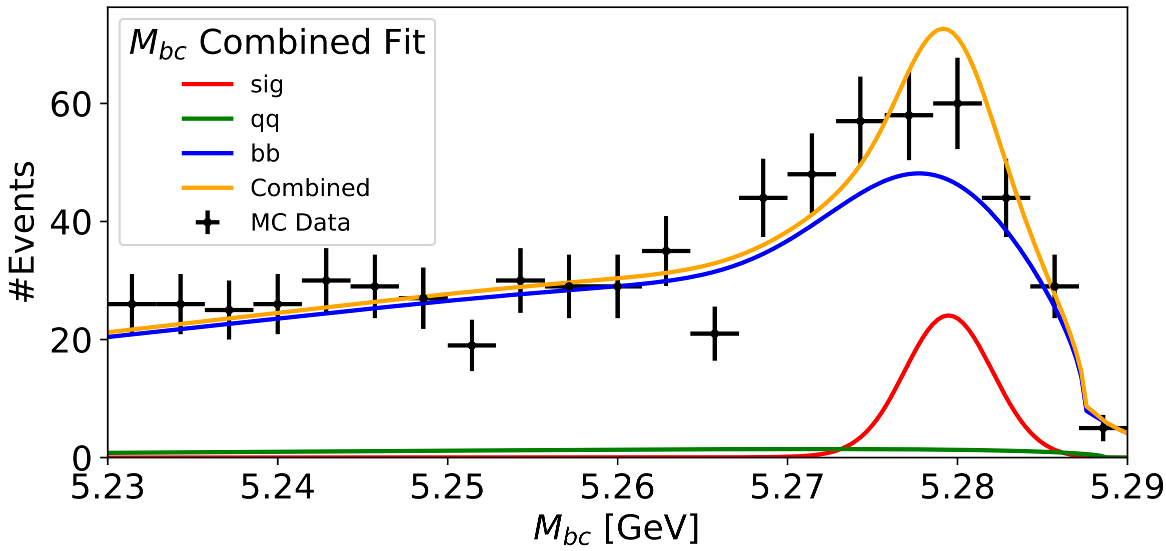

Figure C.4.: Fit Projection of M_{bc} for the fit on the Belle dataset. The Signal, BB background and continuum background yields are shown in red, blue and green, respectively. The sum of the event yields is shown in yellow, the MC data is shown in black.

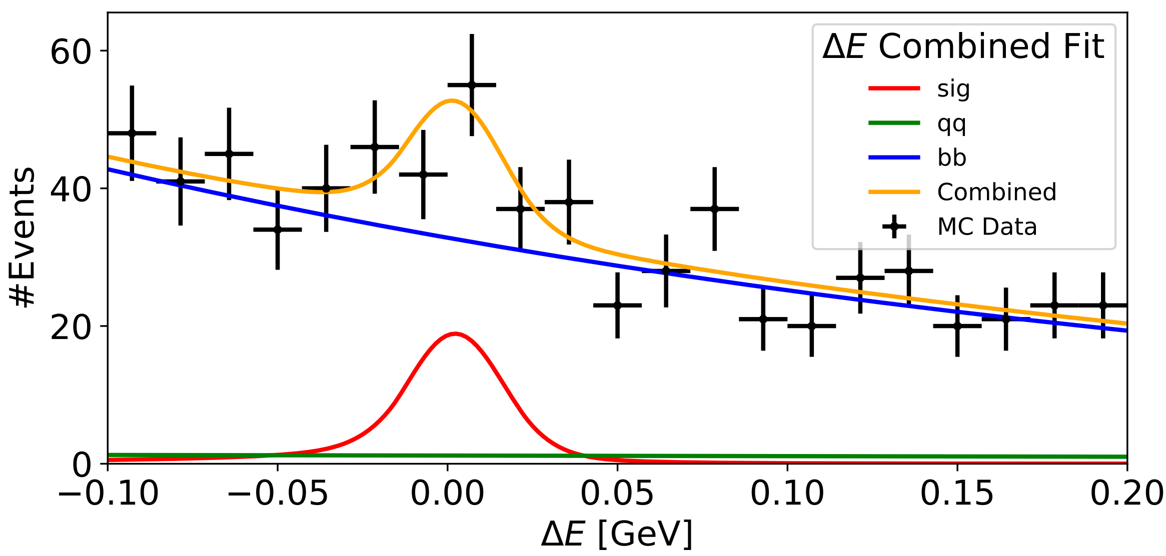

Figure C.5.: Fit Projection of ΔE for the fit on the Belle dataset. The Signal, BB background and continuum background yields are shown in red, blue and green, respectively. The sum of the event yields is shown in yellow, the MC data is shown in black.

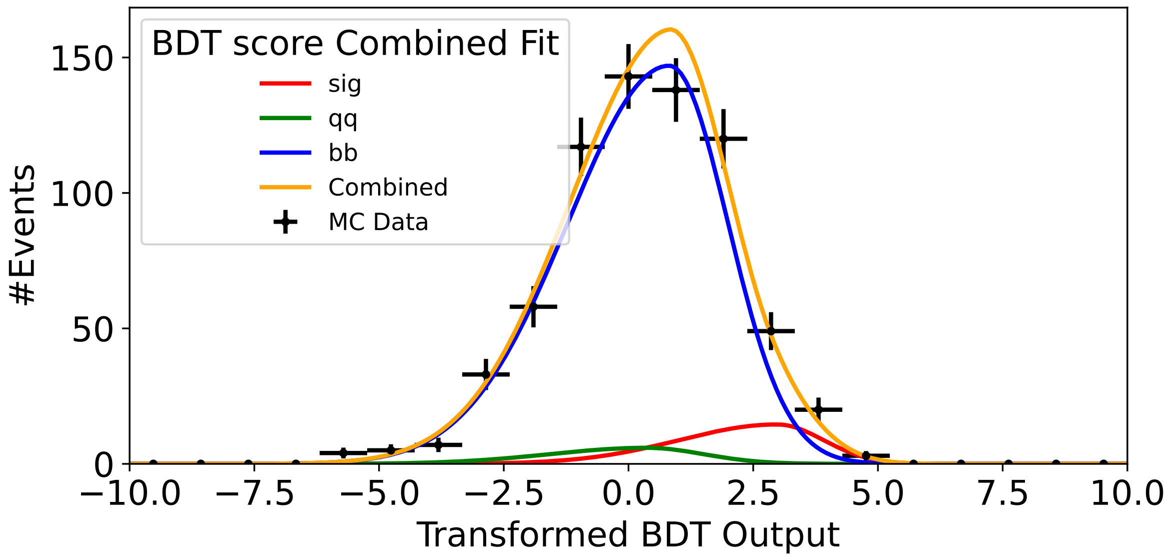

Figure C.6.: Fit Projection of the transformed BDT output for the fit on the Belle dataset. The Signal, BB background and continuum background yields are shown in red, blue and green, respectively. The sum of the event yields is shown in yellow, the MC data is shown in black.

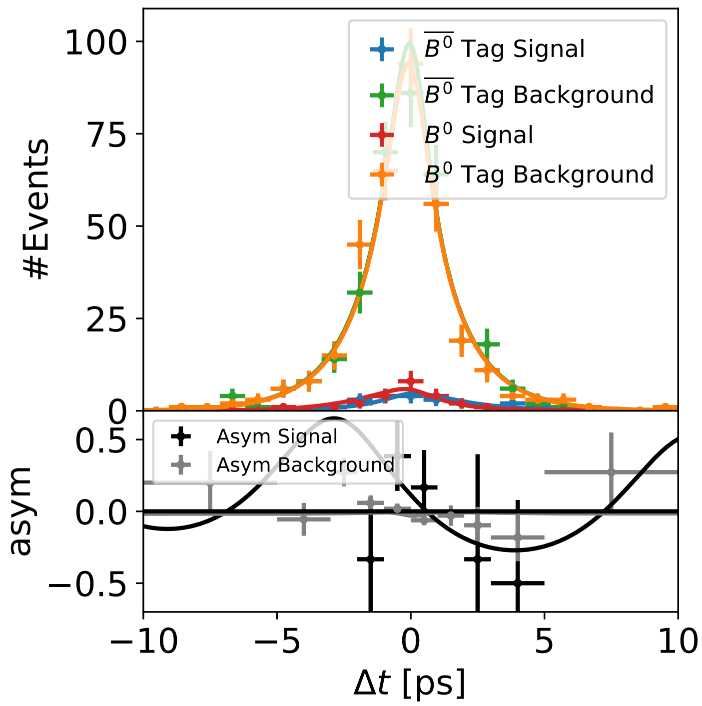

Figure C.7.: Fit Projection of Δt for the fit on the Belle dataset, separated into signal and background events for an B^0 and $\overline{B^0}$ tag, respectively.

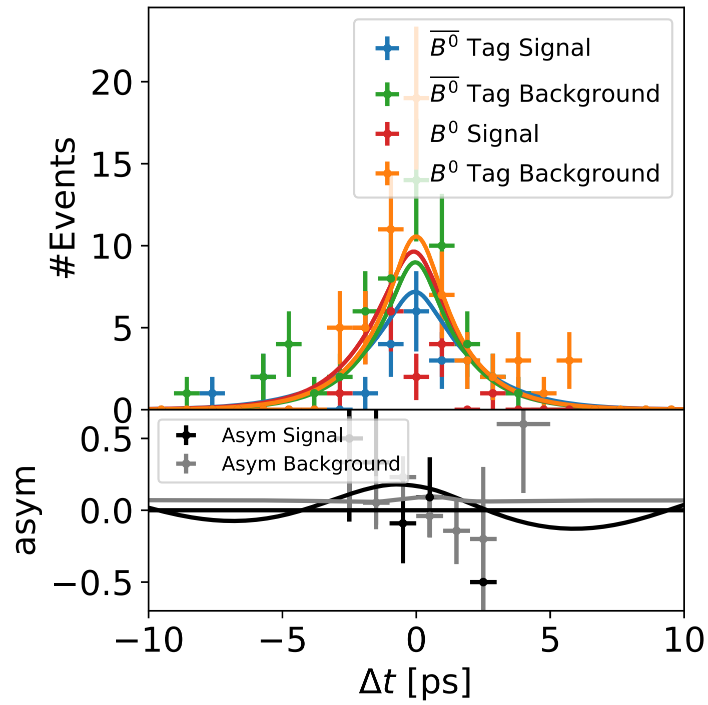

Figure C.8.: Fit Projection of Δt in the signal region for the fit on the Belle dataset, separated into signal and background events for an B^0 and $\overline{B^0}$ tag, respectively.

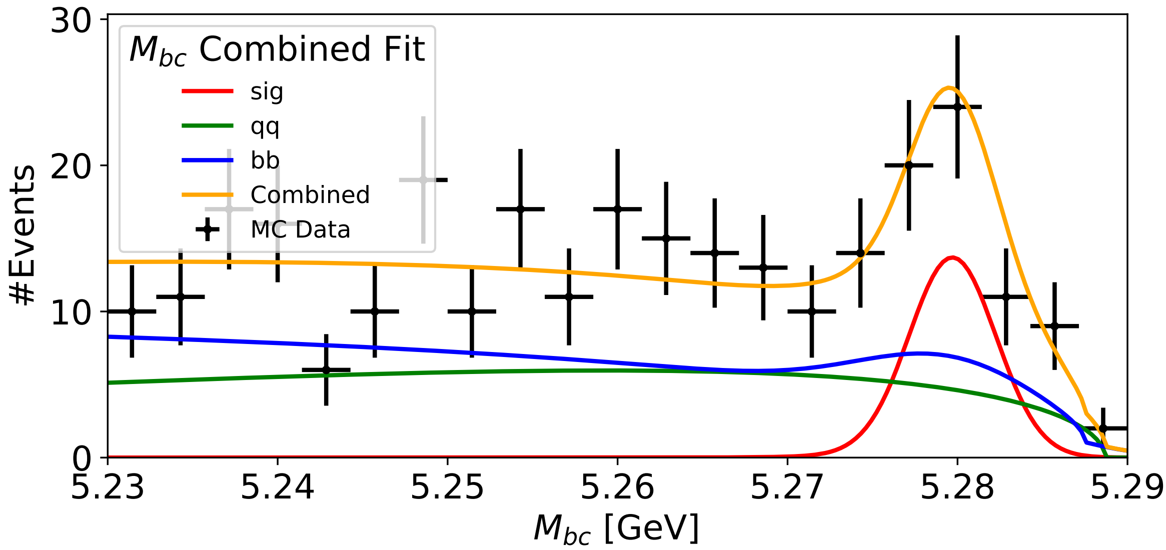

Figure C.9.: Fit Projection of M_{bc} for the fit on the Belle II dataset. The Signal, BB background and continuum background yields are shown in red, blue and green, respectively. The sum of the event yields is shown in yellow, the MC data is shown in black.

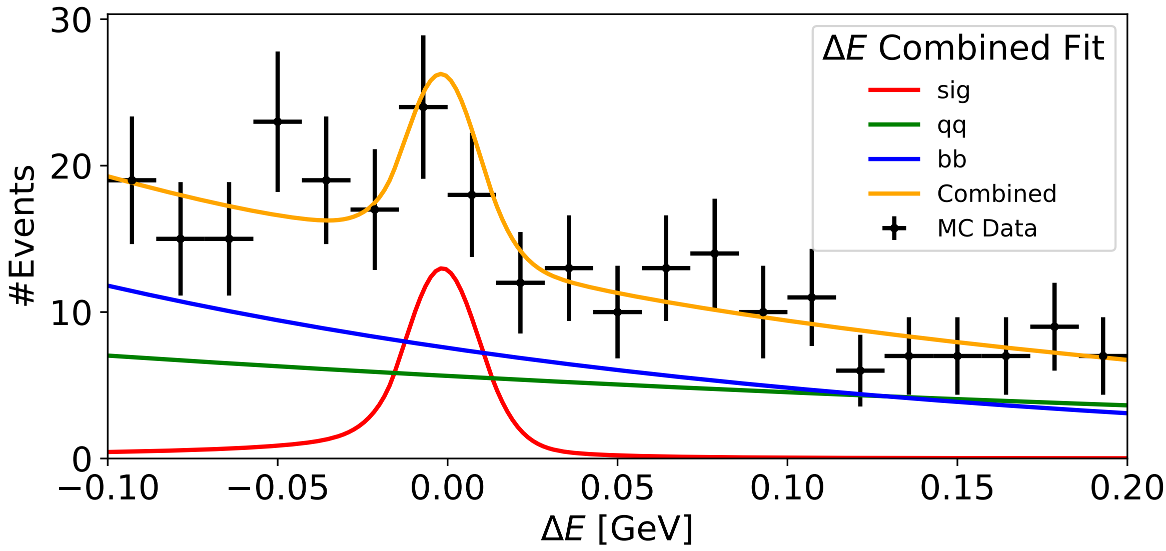

Figure C.10.: Fit Projection of ΔE for the fit on the Belle II dataset. The Signal, BB background and continuum background yields are shown in red, blue and green, respectively. The sum of the event yields is shown in yellow, the MC data is shown in black.

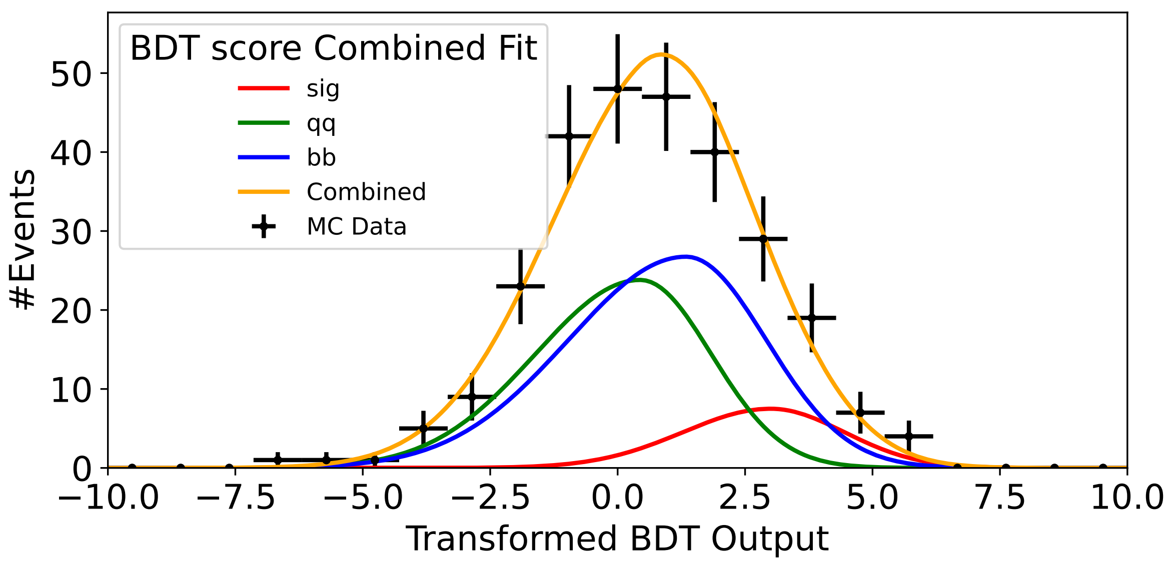

Figure C.11.: Fit Projection of the transformed BDT output for the fit on the Belle II dataset. The Signal, BB background and continuum background yields are shown in red, blue and green, respectively. The sum of the event yields is shown in yellow, the MC data is shown in black.

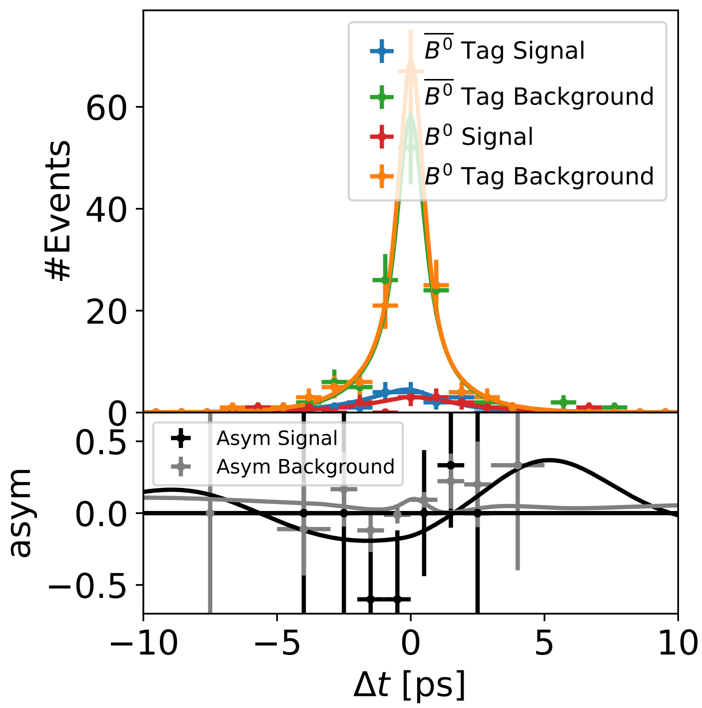

Figure C.12.: Fit Projection of Δt for the fit on the Belle II dataset, separated into signal and background events for an B^0 and $\overline{B^0}$ tag, respectively.

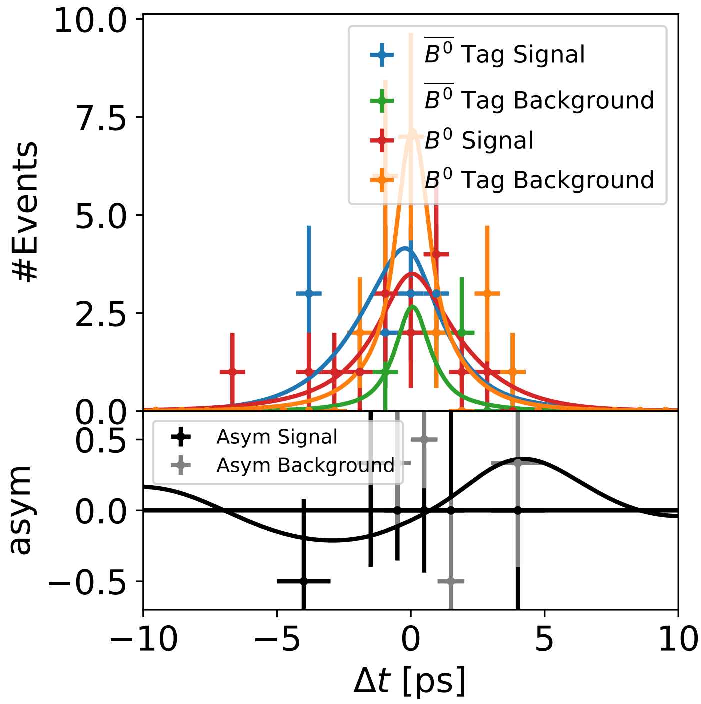

Figure C.13.: Fit Projection of Δt in the signal region for the fit on the Belle II dataset, separated into signal and background events for an B^0 and $\overline{B^0}$ tag, respectively.

Danksagung

First of all, I want to thank Ariane for the opportunity to write my Master Thesis in her working group and for the supervision of my thesis. Also I want to say thank you to the whole groups of Ariane and Thibaud for their help and advice during the weekly meetings.
Furthermore, I want to thank my family and friends for their continuous support throughout the duration of my studies and during my Thesis. I know that I can always count on you, regardless of my problems.
And lastly, of course, a very big thank you to Thibaud for his inexhaustible support and advice during my whole Thesis! No matter the issue I was facing nor the time, I always knew that I could come to you for advice.

Erklärung nach §17(9) der Prüfungsordnung für den Bachelor-Studiengang Physik und den Master-Studiengang Physik an der Universität Göttingen: Hiermit erkläre ich, dass ich diese Abschlussarbeit selbständig verfasst habe, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe und alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten Schriften entnommen wurden, als solche kenntlich gemacht habe.

Darüberhinaus erkläre ich, dass diese Abschlussarbeit nicht, auch nicht auszugsweise, im Rahmen einer nichtbestandenen Prüfung an dieser oder einer anderen Hochschule eingereicht wurde.

Göttingen, den 8. Mai 2026

(Dennis Rodermund)