
MESSUNG DER \mathcal{T} -VERLETZUNG
MIT DEM **Belle** EXPERIMENT

MEASUREMENT OF THE \mathcal{T} -VIOLATION
WITH THE **Belle** EXPERIMENT

Master's Thesis

prepared by

Fabian Wilk

from Kassel

at the II. Institute of Physics

Thesis Period: 15th April 2014 – 14th November 2014

Thesis Number: II.Physik-UniGö-MSc-2014/08

First Referee: Prof. Dr. Ariane Frey

Second Referee: Prof. Dr. Arnulf Quadt

Abstract

A first measurement of direct \mathcal{T} -violation with the Belle experiment is presented. The input data that are used amount to about 619 million $B\bar{B}$ pairs produced in $\Upsilon(4S)$ decays measured by the Belle experiment from e^+e^- collisions. The analysis technique is based on reconstructing the decays of entangled neutral B mesons into states of a definite flavour (B^0 or \bar{B}^0) and \mathcal{CP} (B_+ or B_- which are approximated by $J/\Psi K_L$ and $J/\Psi K_S$ respectively) eigenvalue. These data then enable to form comparisons between pairs of \mathcal{T} -conjugate transitions, yielding a total of four such pairs. Since the centre-of-mass system of the experiment is boosted, the decay time difference between two tags can be measured which then allows to compare the reconstructed decay probabilities for the comparison pairs as a function of the time difference. A method for selecting a highly efficient and pure data sample for usage by this analysis scheme is presented and its performance is evaluated. Finally, an attempt at performing a fit to the signal distributions, with the aim of extracting the \mathcal{T} -violation asymmetry parameters, is presented and its outcome and shortcomings are discussed.

Zusammenfassung

Eine erste Messung von direkter \mathcal{T} -Verletzung mit dem Belle Experiment wird vorgestellt. Die verwendeten Daten entsprechen ca. 619 Millionen $B\bar{B}$ Paaren, welche in Zerfällen von $\Upsilon(4S)$ Teilchen — die in e^+e^- Kollisionen erzeugt worden sind — entstehen. Die Analysemethode basiert auf der Rekonstruktion von Zerfällen von verschränkten, neutralen B Mesonen in Zustände mit einem eindeutigen Flavour (B^0 oder \bar{B}^0) oder \mathcal{CP} (B_+ oder B_- , welche durch $J/\Psi K_L$ bzw. $J/\Psi K_S$ approximiert werden) Eigenwert. Die so gewonnenen Daten erlauben es, Vergleichspaare von Prozessen zu erstellen, welche durch eine \mathcal{T} -Transformation verbunden sind; es können insgesamt vier solche Paare gebildet werden. Aufgrund des geboosteten Schwerpunktsystems des verwendeten Experiments, kann die Zeitdifferenz zwischen den beiden rekonstruierten Zerfällen bestimmt werden. Dies erlaubt es die rekonstruierten Zerfallswahrscheinlichkeitsverteilungen der Vergleichspaare als Funktion der Zeitdifferenz zu vergleichen. Eine Methode zur hocheffizienten Selektion von einem sehr reinen Datensatz, welcher für diese Analyse verwendet werden kann, wird vorgestellt und ihr Leistungsverhalten wird evaluiert. Schlussendlich wird der Versuch eines Fits zur Extraktion der \mathcal{T} -Verletzungsasymmetrieparameter an die Signalverteilungen präsentiert.

TABLE OF CONTENTS

	Abstract	i
	Table of Contents	iii
I	Introduction	1
II	Theoretical Background	3
	1 The Standard Model of Particle Physics	3
	1.1 Mathematical Formalism	4
	1.2 Bosons and the Fundamental Interactions	4
	1.3 The CKM Mechanism	6
	2 Symmetries	7
	2.1 Discrete Symmetries	7
	2.2 \mathcal{CP} Violation	8
	3 The Time Symmetry	9
	3.1 Types of Time Asymmetry	9
	3.2 Time Reversal Violation	10
III	Methodology	11
	1 From \mathcal{CP} to \mathcal{T}	11
	1.1 Direct \mathcal{T} -Violation	11
	1.2 \mathcal{T} -Violation in Mixing	12
	1.3 \mathcal{T} -Violation in Interference	12
	2 Redefining Transition	12
	2.1 Quantum Entanglement	13
	2.2 Measuring \mathcal{T} -Conjugate Transitions	13
	2.3 Tagging the B Mesons	15
	2.4 Quantifying the \mathcal{T} -Violation	16
IV	Experimental Apparatus	19
	1 The KEKB Accelerator Facility	19
	1.1 KEKB as B -Factory	19
	1.2 Data and Luminosity	20
	1.3 Asymmetric Beam Energies	21
	2 The Belle Detector	21
	2.1 Detector Subsystems	22
	3 Data Samples	25
	3.1 MONTE-CARLO Data	25
	3.2 Experimental Data	26

V	The Event Reconstruction	27
1	Recipe	27
2	Reconstructing the J/Ψ	27
2.1	Electron Reconstruction	28
2.2	Muon Reconstruction	28
2.3	Forming and Selecting Dilepton Candidates	29
3	Reconstructing the K_S	29
3.1	The V^0 Particle	29
3.2	Identifying the K_S	30
4	Reconstructing the K_L	30
4.1	Estimating the Four-Momentum	31
5	Reconstructing the B_{CP}	32
5.1	Obtaining CP -even Candidates	32
5.2	Obtaining CP -odd Candidates	33
5.3	Selecting the Final Candidate	33
5.4	The \mathcal{V}_{CP} Vertex	34
6	Reconstructing the B_{Flv}	35
6.1	Flavour Tagging	35
6.2	Estimating the \mathcal{V}_{Flv} Vertex	36
7	Refining the Data Set	38
7.1	Multivariate Analysis Techniques	38
7.2	Boosted Decision Trees	40
7.3	Application for this Analysis	40
8	Continuum Suppression	46
VI	Measuring the \mathcal{T}-Violation	49
1	Estimating the Time Resolution	49
2	Evaluating the \mathcal{T} -Violation	50
2.1	Implementation	51
2.2	Application	51
2.3	Results	51
VII	Conclusion	57
1	Synopsis of the Achievements	57
2	Discussion of the Shortcomings	57
3	Prospects and Outlook	58
	Appendices	61
A	References	61
B	Comparison Pairs	67
C	Asymmetry Parameters	67
D	List of Cuts	68
E	MVA Parameters	69
F	MVA Input Variables	70
F.1	List of Variables	70
F.2	Distributions for minus MVA	71
F.3	Distributions for plus_KLM MVA	72
F.4	Distributions for plus_KLM_ECL MVA	73
	Acknowledgements	75
	Declaration of Conduct	77

I Introduction

Generally, our understanding of the world of elementary particles is very good. The Standard Model of Elementary Particle Physics provides a theory that describes the fundamental interactions of elementary particles with high precision and excellent overall consistency. Its parameters may be numerous, but the accuracy achieved by dedicated experiments for each of those parameters is extremely high. Furthermore, the overall consistency of the model — which can be tested by overconstraining the model parameters using independent measurements — is satisfactory: no statistically relevant deviations have been found thus far.

The so called B -factory experiments Belle and $BABAR$ have been instrumental in the endeavour to prove the practical relevance, that is the real-world applicability, of the theory. First and foremost, they have performed the first conclusive measurement of the violation of the combination \mathcal{CP} of the fundamental symmetries \mathcal{C} and \mathcal{P} in the system of neutral B mesons [1–4]. Following this highly anticipated discovery, the B -physics duo performed numerous precision measurements of \mathcal{CP} -violation and B -physics, including the first ever measurement of the \mathcal{CP} -violating complex phase δ . Furthermore, the two experiments conducted many analyses studying (heavy) flavour physics and the electroweak model of the Standard Model. Many of their analyses were enabled by the unique properties of the neutral and charged B mesons and the production mechanism used at the B -factories. The combination of all those measurements has played a pivotal role in the validation of the predictions and the verification of the consistency of the Standard Model.

Studying the physical operations under which a theory remains invariant is without doubt a task of considerable importance for the testing of a theory. After all, if one can observe an asymmetry in reality this should be modelled correctly by theory. Any deviation between the two is an inadequacy of the accepted theory which cannot be approved. Moreover, the discovery of such inadequacies has more than once sparked scientific progress.

One of the stepping stones used by numerous theoretical models — among them the Standard Model — is the so called \mathcal{CPT} -theorem. This theorem itself can be derived from assumptions that are even more fundamental and somewhat axiomatic to the nature of space and time. It states that the combination of the fundamental space-time symmetries \mathcal{C} , \mathcal{P} , and \mathcal{T} must not be violated, i.e. theory must be invariant under the combined transformations. In the context of the measured \mathcal{CP} -violation, however, this necessitates that the symmetry \mathcal{T} must be violated too: only by combining the asymmetries of \mathcal{CP} and \mathcal{T} , the invariance under \mathcal{CPT} can be facilitated.

Fundamentally, the non-invariance of the \mathcal{T} -symmetry is not a problem for the theory: it can be accommodated in a straightforward manner within the Standard Model using the same mechanism that introduces \mathcal{CP} -violation. However, providing a conclusive and indisputable measurement is significantly more delicate. Over the years, several measurements have been performed, in particle physics as well as other disciplines of physics, but none of them could claim to be direct and truly devoid of doubt. In particle physics these measurements suffer from the \mathcal{CPT} -theorem which, as a side effect, makes it difficult to measure an observable only related to \mathcal{T} . The entanglement of \mathcal{T} and \mathcal{CP} given the \mathcal{CPT} -theorem interferes with an untainted and independent measurement.

Finally, in 2012 the $BABAR$ collaboration published the results of an analysis proving that the fundamental symmetry \mathcal{T} is in fact directly violated by the weak interaction [5]. Not only did they use a method

which can claim to be truly direct, but they also discovered the \mathcal{T} -violation with a significance that makes the result indisputable. Once again, their measurement was facilitated in particular by two things: the fact that they study B mesons and the specific production mechanism employed.

Due to the fact that the two experiments Belle and $BABAR$ are akin to each other, there is the unique opportunity to validate the discovery using an independent experiment. Following the scientific method, the claim of a discovery is only truly credible after a second, independent measurement has replicated the results which makes such a measurement no less important.

This thesis reports the results of a first attempt at doing just that. That is to say, replicating the analysis performed by $BABAR$ using the data measured by the Belle experiment. It should be seen as an investigative study, hence it is by no means complete and many aspects of the analysis have not been fully validated. In fact, several important topics such as the evaluation of statistical and systematic uncertainties have not been considered at all. The lion's share of the analysis work is based upon the reconstruction scheme deployed by previous analyses performed by Belle. Only the particular technique used for measuring an observable related to \mathcal{T} is borrowed from the analysis leading to the discovery by $BABAR$.

In Chapter II the theoretical background of the analysis — namely the Standard Model and the concept of symmetries — is briefly reviewed hereby placing the analysis into a broader context. Following that, in Chapter III the technique that was developed and employed by $BABAR$ is discussed. This should also provide additional insight into the possible concrete manifestations of the \mathcal{T} -violation. Afterwards the experimental setup, i.e. the KEKB accelerator and the Belle detector, are summarised in Chapter IV. In the course of this discussion, the different detector components and their purpose within the detector as a scientific apparatus are presented as well as the data samples that are used by this analysis. Following that, in Chapter V the event reconstruction is discussed. This chapter amounts to the largest portion of the thesis which is not surprise considering its share in the total effort and importance for the analysis. Finally, in Chapter VI the actual \mathcal{T} -violation measurement is presented and the results are discussed. A summary of the achievements of this analysis, the remaining problems, and a discussion of the prospects are given in Chapter VII which also contains the concluding remarks.

Remark: Throughout this thesis the system of natural units (where $\hbar = c = k_B = 1$) is adopted unless stated otherwise. Electric charges are implicitly given in multiples of the elementary charge e . Furthermore, all natural constants and the values of particle properties are taken from [6].

II Theoretical Background

In this chapter, the theoretical background of the analysis will be presented. Section 1 places the analysis into a broader context by briefly discussing the Standard Model of Particle Physics which is the theory of the fundamental particles and how they interact. In Section 2, a brief overview over the concept of symmetries, different types of symmetries, and their relevance for physics will be given. Finally, in Section 3, the time (a)symmetry — the essence of the analysis — will be discussed.

1 The Standard Model of Particle Physics

The Standard Model of Particle Physics (SM) [6–9] is a refined theory which has been extended and improved over the course of the last century. It represents the current knowledge of the world of fundamental particles and the forces acting between them. This, of course, implies that it is by no means a full-fledged or complete theory: it has several known deficiencies and shortcomings. However, history provides reason to trust in its ability to evolve, adapt to new discoveries, and accommodate new, extending theories.

The SM distinguishes between elementary and composite particles: the former are point-like, meaning they have no intrinsic substructure, whereas the latter are compound objects. Elementary particles are further categorised by their spin-eigenvalue: **Fermions** (which are spin- $1/2$ particles) are the fundamental matter-particles and **Bosons** (which have integer spin) are the force-carriers that are required for particle interactions. Figure 1 depicts a graphical representation of the fundamental particles in the SM.

To this date, there are twelve known fermions (each paired with its anti-particle) and five fundamentally different bosons (the gluon, two weakly interacting massive bosons, the photon, and the Higgs boson) which are elementary.

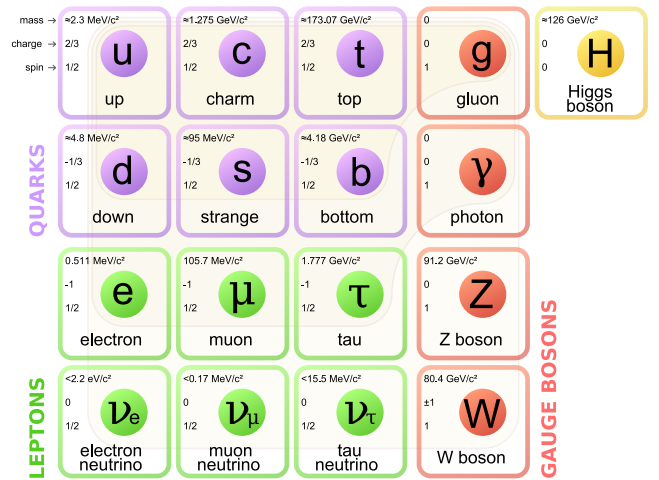


Fig. 1: Graphical representation of the elementary particles in the Standard Model [10].

The family of elementary fermions can be further categorised into the **Quarks** and **Leptons**, marked in purple respectively green in Figure 1. There are six different types — called *flavours* — of quarks and leptons, which are grouped in three generations (depicted as columns in Figure 1). All particles of a given flavour are equivalent regarding the remaining (non-flavour) quantum numbers.¹ Fermions are sometimes called *matter particles*, however, only the particles of the first generation are observed as being constituents of stable matter. The members of generations two and three are unstable themselves and can be created only for a short moment in high energy collisions.

Quarks are grouped further as *up-type* quarks (u , c , t) which have electric charge $Q = +2/3$ and *down-type* quarks (d , s , b) with $Q = -1/3$. Similarly, one can group leptons into the charged leptons

¹ *Flavour* can refer to both, a particle's species and the collection of *internal* quantum numbers (e.g. lepton/baryon number, strangeness, etc.) which are defining for a particle's species.

(e , μ , τ) and the electrically neutral neutrinos (ν_e , ν_μ , ν_τ). For each² fermion f there exists an anti-fermion \bar{f} which is completely equivalent except for its internal quantum numbers³ which have the *flipped* value.

1.1 Mathematical Formalism

The theoretical background of the SM comes from Quantum Field Theory (QFT) where particles are described by fields. Using this concept, a given model is fully described by its LAGRANGE density \mathcal{L} . Free particles are introduced into the model by directly adding their respective field equation to \mathcal{L} . Usually, the fields corresponding to an interaction are introduced by requiring that \mathcal{L} — which already contains the field(s) of the elementary free particle(s) — shall be invariant under local gauge transformation [9]. The modifications that must be made to ensure this invariance give rise to one or multiple new fields (the so called gauge fields).

Quantisation of a gauge field yields a so called gauge boson, which is the mediator particle associated with the fundamental interaction that was introduced to \mathcal{L} in the first place. At the same time this mechanism gives rise to an intrinsic quantum number which can be interpreted (in analogy to the electric charge) as the *charge* of the corresponding interaction. The electromagnetic interaction, for example, is introduced by requiring invariance under $U(1)_Q$ and gives rise to the gauge field A_μ and the quantum number Q which can be associated with the electric charge.

Mathematically the SM is a quantum gauge theory with the internal symmetries of the composite group

$$SU(3)_C \otimes SU(2)_I \otimes U(1)_Y .$$

This group product is composed of the $SU(3)_C$ subgroup of the QCD — that is the quantum theory of the **strong force**, cf. Section 1.2.1 — and the $SU(2)_I \otimes U(1)_Y$ which is the symmetry group of the **electroweak interaction**, cf. Section 1.2.2.

1.2 Bosons and the Fundamental Interactions

The Standard Model covers three fundamental interactions known as the strong, weak, and electromagnetic forces. The gravitational force, although apparently being fundamental, is not described by the standard model, which is one of its shortcomings. Finding and testing a mathematical formalism for describing quantum gravity remains a top priority for theoretical particle physics. However, for all ordinary experimental matters, gravity can be neglected without further discussion since it is several orders of magnitude weaker than the other interactions. Table 1 compares the coupling strengths of the fundamental forces.

Force	Relative Strength
Strong	~ 1
Electromagnetic	$\sim 10^{-3}$
Weak	$\sim 10^{-8}$
Gravity	$\sim 10^{-37}$

Tab. 1: Relative coupling strengths of the four fundamental forces as exerted between two fundamental particles at a distance of 1 fm [7].

Each fundamental force is mediated by at least one exchange particle: these particles are bosons (i.e. they have integer spins) and are marked in red in Figure 1. However, due to conservation of different quantum numbers only certain types of interactions (transformations between particles) are allowed. Figure 2 shows all possible fundamental interaction vertices in the Standard Model.

In addition to the bosons which mediate the fundamental forces, there is an additional bosonic member of the SM: the HIGGS boson. Invention of the HIGGS mechanism⁴ [11–13] solved the conceptual problem that, although massive mediator bosons had been measured, the SM could not provide a mathematical formalism that describes how they acquired mass in the first place.

By means of spontaneous symmetry breaking, the HIGGS mechanism gives rise to the masses of the mediators of the weak interaction and all fermions (except for that of the neutrinos which is, however,

² Whether the neutrino is a DIRAC or MAJORANA particle — that is to say whether it has an anti-particle or is its own anti-particle — is still a topic open for discussion: no conclusive evidence for either theory has been found thus far.

³ Those are the flavour quantum numbers and others such as electric charge Q , weak isospin, etc.

⁴ More aptly named ABEGHHK'tH mechanism by P. HIGGS (for ANDERSON, BROUT, ENGLERT, GURALNIK, HAGEN, HIGGS, KIBBLE and 't HOOFT) due to the variety of scientists who co-invented the mathematical mechanism first in the context of superconductivity and later particle physics.

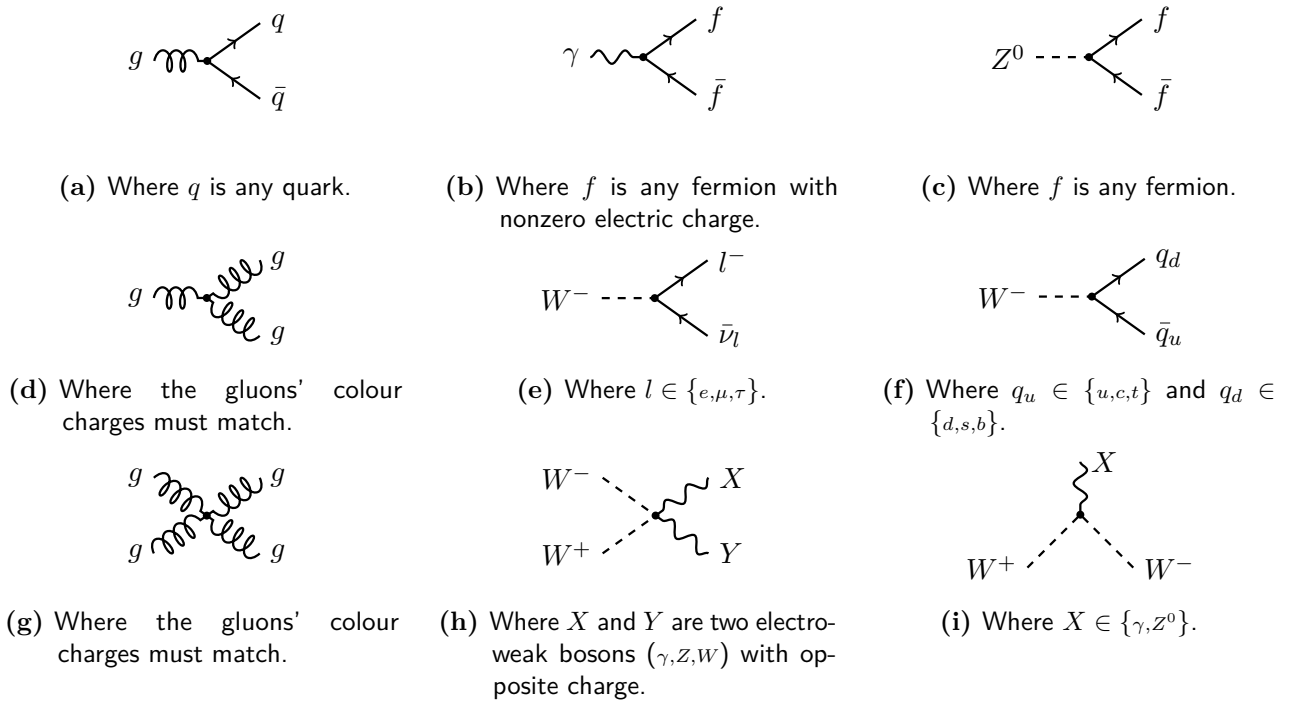


Fig. 2: FEYNMAN diagrams of all fundamental interaction vertices possible in the Standard Model. Charge conjugates — if they exist — are implied [7, 8].

a design choice and not a limitation). The search for the HIGGS boson had been going on for several decades until it was finally discovered in 2012 by the ATLAS and CMS collaborations [14, 15]. Since the HIGGS mechanism has no relevance to this analysis, it will not be discussed any further.

1.2.1 The Strong Interaction

The dynamics of the strong force are described by Quantum Chromodynamics (QCD). Its mediator — the gluon g — has no mass and none of the charges required to interact with any of the other forces.

Various experiments have determined that the charge of the strong interaction C is in fact three-valued. Appropriately the charge is called *colour charge* and its valid states are labeled *red*, *green*, and *blue* (each with its anti-colour counterpart). Since the gluon itself carries one quantum of colour and one of anti-colour charge, there are nine possible ways to form a gluon. However, one of them is a colour singlet (zero-valued total colour charge) and thus cannot partake in the strong interaction,⁵ hence, there is a total of eight known gluons. The only known particles other than the gluon carrying colour charge are the quarks.

Because the strong mediator carries colour too, it can couple to other gluons: self interaction, including so called three- and four-vertices (where three or four gluons couple to each other), is possible (cf. Figures 2d and 2g). This property of the gluon gives rise to a unique property of the strong force: the *colour confinement*.

As two colour-connected quarks are moved apart, the gluon field forms a narrow tube between the two quarks, due to their self-attraction, rather than spreading out uniformly [7]. Since the energy stored within the colour field increases nearly linearly, at some point it becomes energetically more favourable for the field to produce a new quark-antiquark pair. This is because afterwards the distance between two colour connected states is reduced to roughly half of the previous value. Consequentially, free quarks are suppressed and have not been observed so far. Furthermore, a bound quark state (a so called hadron) can exist only if it is a colour singlet, i.e. the sum of the colour charges of its constituents must be zero.

The strong force can bind quarks to form compound states. So far only states of a quark and an anti-quark — called **Mesons** — as well as three-quark states — called **Baryons** — have been observed,

⁵ Albeit it would be able to interact with other colour singlet states, it is usually assumed that this colour singlet gluon does not exist.

however, there is no theoretical reason for this limitation. Mesons are generally unstable and short-lived, whereas baryons can be very long-lived (e.g. the neutron) or even be considered stable (e.g. the proton).

1.2.2 The Electroweak Interaction

In the 1960s GLASHOW, SALAM, and WEINBERG proposed the model of *electroweak unification* (also called GSW theory) which fully describes electromagnetic and weak interactions of particles [16–18]. Their model solved the problem that introduction of the weak interaction into the LAGRANGE density of the SM by using the $SU(2)_I$ symmetry group results in a field without correspondence to any known particle. The GSW model proposes that instead of introducing the electromagnetic and weak interactions separately, they should be included using a common symmetry group, hence the $SU(2)_I \times U(1)_Y$ subgroup of the SM was born. Its group generators are the weak isospin I and the hypercharge $Y := 2 \cdot (Q - T_3)$, where Q is the electric charge and T_3 the third component of the weak isospin of a given state.

The $SU(2)_I$ symmetry directly produces two new bosons, the W^\pm bosons which interact with particles that have a nonzero value of T_3 . By mixing the third component of $SU(2)_I$ and the $U(1)_Y$ symmetry, two new fields are produced which correspond to the photon γ and the Z^0 boson. Nearly twenty years after the theoretical work by GLASHOW, SALAM, and WEINBERG, the UA1 and UA2 collaborations discovered the three proposed particles [19–22]. Over the course of the following twenty years several experiments performed precision measurements of numerous Standard Model parameters confirming the predictions of the GSW model [23, 24].

It can be found that the $SU(2)$ component of the newly introduced coupling has a V-A-structure.⁶ A pure V-A-coupling — which the W^\pm bosons have — has a rather peculiar property: it interacts exclusively with left-handed particles and right-handed anti-particles. In the theory left-handed particles form doublets of up- and down-type fermion pairs whereas right-handed particles form singlets, except for the right-handed neutrino which does not exist in the SM:

$$\begin{pmatrix} u \\ d \end{pmatrix}_L \quad \begin{pmatrix} \nu_l \\ l \end{pmatrix}_L \quad u_R \quad d_R \quad l_R.$$

The charged-current interaction of the electroweak model is unique in the sense that it is not flavour conserving in the quark sector. A W^+ boson can couple to any pair of up-type quark q_u and down-type antiquark \bar{q}_d (charge conjugates implied) because the weak and flavour eigenstates are not equal in the CKM mechanism, cf. Figure 2f and Section 1.3.

1.3 The CKM Mechanism

The **Cabibbo-Kobayashi-Maskawa (CKM) mechanism** is the theory of the flavour-changing charged-current weak interactions in the quark sector. Initially, the theory predicted that the coupling of the W bosons should be flavour ignorant, that is to say its strength should be equal for all flavours. However, various experiments showed discrepancies in the measured coupling strengths: universality of the weak interaction was challenged. In 1963 CABIBBO first proposed a solution to the problem [25] which was later generalised to three quark generations by KOBAYASHI and MASKAWA in 1973 [26]. The newly theorised particles — the bottom and top quarks — were discovered in 1977 [27] and 1995 [28, 29] respectively.

The central idea of the CKM mechanism is that the flavour eigenstates of the weak interaction $|q\rangle$ are not equal to the mass eigenstates $|q'\rangle$. By multiplying a unitary, complex transformation matrix one can be transformed into the other. Conceptually this is a basis transformation rotating the states from one space to the other. By convention the effect is absorbed in the down-type quarks i.e. $|q'_u\rangle \equiv |q_u\rangle$ (charge conjugates implied) therefore

$$\begin{pmatrix} |d'\rangle \\ |s'\rangle \\ |b'\rangle \end{pmatrix} = \underbrace{\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}}_{V_{\text{CKM}}} \cdot \begin{pmatrix} |d\rangle \\ |s\rangle \\ |b\rangle \end{pmatrix}, \quad (1)$$

⁶ Which stands for *Vector-Axialvector Structure/Coupling*.

where V_{CKM} is the **CABIBBO-KOBAYASHI-MASKAWA Matrix**. It can be shown that the CKM matrix has four irreducible degrees of freedom: three mixing angles $\theta_{12}, \theta_{23}, \theta_{13}$ and one complex phase δ . There exist different parametrisations of the CKM matrix, of which the **WOLFENSTEIN** parametrisation [30] is the most natural one since it directly reflects the magnitude of the different matrix elements.

Any charged-current vertex that couples two quarks i, j introduces a proportionality to V_{ij} (for the \mathcal{CP} -conjugate of that vertex the complex conjugate V_{ij}^* is introduced) to the matrix element used for calculating the transition probability. Hence, assuming that V_{CKM} is non-trivial (i.e. not the unit matrix), the *mismatch* of the quantum states of propagation and interaction results in varying coupling strengths for the different quarks. Equation (2) shows current best-measured values [6] of the CKM matrix elements.

$$\begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} = \begin{pmatrix} \sim 0.97 & \sim 0.22 & \sim 0.003 \\ \sim 0.22 & \sim 0.97 & \sim 0.04 \\ \sim 0.008 & \sim 0.04 & \sim 0.99 \end{pmatrix} \quad (2)$$

Numerous analyses have shown that the off-diagonal matrix elements — corresponding to flavour-changing couplings — are much smaller than the on-diagonal elements which match the in-generation couplings,⁷ cf. Equation (2).

2 Symmetries

Physics defines a symmetry of a system as a transformation that leaves certain physical properties of that system unchanged. These quantities are then invariances of the symmetry. There are two families of transformations, those which are *continuous* and those which are *discrete*:

- A **continuous** transformation maps from a single input to a continuous and essentially infinite set of output values. The *translational symmetry* comes to mind as an example because it claims that physics remains the same after applying the transformation $T_a(x) : x \mapsto x + a, \forall x, a \in \mathbb{R}$. A more figurative example is the rotation of a sphere around any axis through its centre.
- A **discrete** transformation maps from a single input to a finite set of output values (often just one). Discrete transformations involve some kind of *swapping* between values of a finite set of values. A prime example is the *mirror symmetry* (i.e. reflection on a plane) which claims that physics shall be invariant under $T(x) : x \mapsto -x, \forall x \in \mathbb{R}$. Rotation of a regular polygon around its centre is an example that is more intuitively accessible.

2.1 Discrete Symmetries

There are three elementary discrete transformations and therefore three possible basic discrete symmetries: [9, 31]

- **Parity Inversion \mathcal{P}** reverses the handedness of space, i.e. $\mathcal{P}(\vec{x}) : \vec{x} \mapsto -\vec{x}$. Furthermore, if applied on a field of given chirality it transforms it into its chiral counterpart

$$\mathcal{P}\psi_L(\vec{x}, t) = \gamma^0\psi_R(-\vec{x}, t) \quad \text{and} \quad \mathcal{P}\psi_R(\vec{x}, t) = \gamma^0\psi_L(-\vec{x}, t),$$

where γ^0 is the 0th **DIRAC** matrix [9]. Given that the charged-current weak interaction is asymmetric w.r.t. the chirality, it becomes obvious that it must violate \mathcal{P} .

- **Charge Conjugation \mathcal{C}** interchanges particles with their corresponding anti-particles. As a result the signs of all electric charges are inverted, i.e. $\mathcal{C}(Q) : Q \mapsto -Q$; furthermore all internal quantum numbers are swapped.

It can be shown that the presence of both vector and axialvector structure in the weak interaction requires violation of \mathcal{C} by the weak interaction since it acts differently on the two substructures.

- **Time Reversal \mathcal{T}** mirrors the time component of a state, i.e. $\mathcal{T}(t) : t \mapsto -t$. Hence, spatial positions are left unchanged but the signs of linear and angular momenta are modified: $\vec{p} \mapsto -\vec{p}$ [31, 32].

These elementary transformations can be combined, forming the \mathcal{CP} -, \mathcal{CT} -, \mathcal{PT} -, and \mathcal{CPT} -transformations, by successive application of the component transformations. In the mid 1950s it was shown that

⁷ Refer to [6] for a review of measurements of the CKM matrix elements.

any *canonical* quantum field theory must exhibit invariance under the CPT -symmetry [31, 33, 34]; this is called **CPT theorem**. In this context, a canonical QFT is defined by being local, LORENTZ covariant and having a lowest energy state (*ground state*).⁸ Even in the most extreme models theorists develop for particle physics, the CPT theorem is usually imposed.

One approach to testing the CPT theorem is to compare the mass of a particle with that of its anti-particle since CPT -invariance requires the masses to be equal. Due to its experimental availability and small decay width, the top quark is a prime candidate for such a measurement; recent measurements have found no indication for CPT -violation [36, 37].

2.2 CP Violation

For a thorough discussion of the CP -violation refer to [38], for a review of the CP -violation experiments that have been performed refer to [6].

Until the mid 1950s the invariance of the fundamental symmetries C , P , and T was accepted scientific consensus. However, in 1956 YANG and LEE pointed out the potential for violation of the P symmetry by the weak interaction [39] and suggested experiments which should provide a conclusive answer. Shortly thereafter it was indeed shown that both C and P are violated by the weak interaction [40, 41]. Early on this was considered a fundamental problem and for some time it was assumed that the combination CP would be a proper symmetry. Eventually this assumption was refuted in measurements of kaon decays in 1964 [42]. However, for some years to come there was no proper theoretical explanation for the measured effect.

Recall that the CKM mechanism introduced an irreducible complex phase δ into the CKM matrix, cf. Equation (1). Hence, given that δ is nonzero, it is required that $V_{ij} \neq V_{ij}^*$ and thus the transition probability of a charged-current process must be different from its CP -conjugate counterpart. By solving the coupling-strength problem, KOBAYASHI and MASKAWA introduced CP -violation into the Standard Model. To this date, the CKM mechanism remains the only experimentally conclusive source of CP -violation in the Standard Model. Initially it was thought that the CP -violation of the SM could explain the apparent matter-antimatter asymmetry of the universe, however, it was found that this is not the case, cf. Section 2.2.2.

2.2.1 Types of CP -Violation

In the SM the CP -violation can become apparent in three different ways:

- **Direct CP -Violation** is observed when CP is violated in a decay. For example, this type is observed if

$$\Gamma(M^0 \rightarrow f) \neq \Gamma(\bar{M}^0 \rightarrow \bar{f}),$$

where M^0 is a neutral meson, f is a final state, and the two processes are related by CP . This type is the only source of CP -violation in the decay of charged mesons. It was first observed conclusively by the NA48 and $KTeV$ experiments [43, 44]. Since then it has been observed in various decay channels of kaons and B mesons [6].

- **CP -Violation in Mixing** can be observed through the mixing of neutral states. Figure 3 shows one of the FEYNMAN box-diagrams that describe this process. Given that the probability for a transition from particle to anti-particle is not exactly equal to the probability for the other direction, CP is violated:

$$\Gamma(M^0 \rightarrow \bar{M}^0) \neq \Gamma(\bar{M}^0 \rightarrow M^0).$$

This is the first type of CP -violation that was observed, it was first measured by CRONIN, FITCH, *et al.* [42].

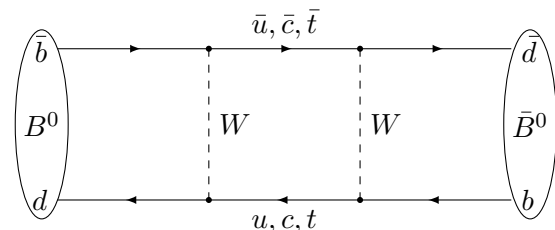


Fig. 3: A possible FEYNMAN diagram of the $B^0 \bar{B}^0$ oscillation.

⁸ In 2002 it was proven that the inverse is partially true as well: it was shown that presence of CPT -violation would imply violation of the LORENTZ symmetry [35].

- **\mathcal{CP} -Violation in Interference** arises from the interfering matrix elements for decays with and without mixing. Figure 4 shows an illustration of the physical processes and interfering terms partaking in this process. Formally, this type of \mathcal{CP} -violation requires that

$$\Gamma(M^0(\rightsquigarrow\bar{M}^0) \rightarrow f)(t) \neq \Gamma(\bar{M}^0(\rightsquigarrow M^0) \rightarrow \bar{f})(t),$$

where $(\rightsquigarrow\bar{M}^0)$ denotes a possible oscillation of the meson state.

This type has been observed in various decay channels of B mesons [6].

2.2.2 Measurement of the \mathcal{CP} -Violation

Since it was first observed, many experiments have performed a \mathcal{CP} -violation measurement. However, it was only in 2001 that the Belle and BABAR experiments published their measurements of the complex phase δ , showing that it is indeed nonzero [2, 4]. The results established \mathcal{CP} -violation, which so far had only been observed in kaon decays [42–44], in the B meson system. In 2008 — following the discovery of the top quark — the theoretical work by KOBAYASHI and MASKAWA was finally honoured with a NOBEL price.

It should be noted that the CKM mechanism with the measured value of δ is not sufficiently strong \mathcal{CP} -violating to explain the observed matter-antimatter discrepancy. Hence, this topic is still under ongoing research in both particle physics and cosmology [6] and is one of the major remaining deficiencies of the SM.

3 The Time Symmetry

In a classical context, \mathcal{T} -invariance can be observed in the fact that the equations of motion for a particle going from I to F along some path C also allow the reversed motion. That is to say, it is permitted for the same particle to follow the path C in reversed direction, i.e. from F to I . In a quantum mechanical description the required transformation is introduced as an anti-unitary operator $U_{\mathcal{T}}$ [31]. This operator transforms a field $|\psi\rangle$ to its complex conjugate, hence it effectively exchanges incoming and outgoing states [31, 45]:

$$\langle U_{\mathcal{T}}\phi|U_{\mathcal{T}}\psi\rangle = \langle\psi|\phi\rangle.$$

Following this formalism, time reversal is a proper symmetry if

$$U_{\mathcal{T}}|\phi\rangle_{\text{out}} = |\tilde{\phi}\rangle_{\text{in}} \quad \text{and} \quad U_{\mathcal{T}}|\psi\rangle_{\text{in}} = |\tilde{\psi}\rangle_{\text{out}}, \quad (3)$$

for some state $|\psi\rangle$ ($|\phi\rangle$) which relates to $|\tilde{\psi}\rangle$ ($|\tilde{\phi}\rangle$) by reversing all momentum directions \vec{p} ; the subscript in (out) denotes an incoming (outgoing) state. This can be deduced from imposing invariance of the scattering matrix \mathcal{S} :

$$\mathcal{S}_{\phi\psi} := {}_{\text{out}}\langle\phi|\psi\rangle_{\text{in}} = {}_{\text{in}}\langle U_{\mathcal{T}}\psi|U_{\mathcal{T}}\phi\rangle_{\text{out}} = {}_{\text{out}}\langle\tilde{\psi}|\tilde{\phi}\rangle_{\text{in}} =: \mathcal{S}_{\tilde{\psi}\tilde{\phi}}. \quad (4)$$

3.1 Types of Time Asymmetry

There are two types of observable time asymmetries, however, only one is related to \mathcal{T} and thus follows the formalism introduced above. In this section the difference between those two types is investigated.

Proper \mathcal{T} -violation — aptly named **genuine time reversal violation** — requires that a process is non-invariant under simultaneous exchange of the sign of the time t (hence reversing the sign of linear and angular momenta) and swapping of in and out states [31, 32]. Consider Equation (3) as an example for proper transformation under \mathcal{T} .

Other processes which have an apparent asymmetry w.r.t. t but do not comply with the above requirements are non-genuine time asymmetries. Such asymmetries are not a result of any asymmetry

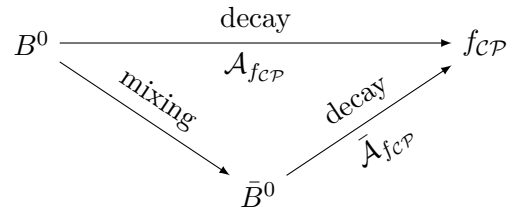


Fig. 4: Interference triangle for the *Golden Channel* (where $f_{\mathcal{CP}} = J/\Psi K_S$) of the interference \mathcal{CP} -violation measurements.

in the fundamental laws of motion but rather produced by the specific set of initial conditions [32, 46]. Examples are the **Universe t -asymmetry** and the so-called **macroscopic arrow of time** [46, 47].

The universe t -asymmetry is a direct result of the second law of thermodynamics given the set of initial conditions specific to our universe. Essentially, this phenomenon is the fact that there exists an obvious time asymmetry in the universe observable through its expansion and the formation of stellar and planetary bodies. However, this is perfectly compatible with a theory that is \mathcal{T} -symmetric since it is a result of the specific initial conditions.

The second example is a macroscopic effect: in a complex isolated system one can observe a time asymmetry w.r.t. the amount of order in that system. This can be observed, for example, in MAXWELL's equations with material absorption or friction in NEWTONIAN mechanics. However, the underlying microscopic processes (i.e. at the level of elementary particles) are still symmetric w.r.t. t [32, 46].

It should be noted that even in particle physics there exist effects which are odd under exchange $t \leftrightarrow -t$, that are, however, not necessarily \mathcal{T} -violating. Once again compliance to both requirements for genuine \mathcal{T} -violation is required. Particle decays are an example of a clearly time-asymmetric process in particle physics, however, the irreversibility of an ordinary decay is not related to \mathcal{T} but rather a result of improper preparation of states [46, 47].

3.2 Time Reversal Violation

Following the formalism of Equations (3) and (4) some further calculation shows that, unlike the charge conjugation, the time reversal transformation acts on vector and axialvector currents in the same manner [31]. Hence, the bare weak charged-current gauge interactions are invariant under \mathcal{T} .

However, it can be shown that the mismatch between weak interaction states and mass eigenstates (which already introduced \mathcal{CP} -violation to the SM, cf. Section 2.2) gives rise to \mathcal{T} -violation in the charged-current weak interaction [31]. This is, of course, not surprising considering that \mathcal{CP} and \mathcal{T} are related in the theory by the \mathcal{CPT} -theorem, cf. Section 2.1. Furthermore, this equivalence predicts that there should exist one type of \mathcal{T} -violation per type of \mathcal{CP} -violation. These pairs of effects should then cancel each other out, therefore ensuring \mathcal{CPT} -invariance.

Following the previous discussion, it can be expected that \mathcal{T} -violation manifests phenomenologically similar to \mathcal{CP} -violation. Furthermore, a model dependent, indirect \mathcal{T} -violation measurement can be performed by directly measuring the violation of the \mathcal{CP} -symmetry and the invariance under \mathcal{CPT} using two independent experiments and analyses. Unless the fundamental concept of symmetries is flawed, the presence of those two effects must enforce that \mathcal{T} is violated. Various indirect measurements have been performed in the neutral kaon [48–51] and B meson [52–54] systems and the overall picture points towards a non-invariance under \mathcal{T} . However, none of these measurements can claim to be truly direct and independent.

Measuring direct \mathcal{T} -violation poses a formidable challenge: in order for the measurement to be untainted and direct one must rely on purely \mathcal{T} -related observables. This ensures absence of contamination from \mathcal{CP} - or \mathcal{CPT} -violation. Obviously, a direct measurement must have direct access to a violated observable. The generic processes to be measured can be formulated as

$$\mathbf{I} : \text{in} \rightarrow \text{out} \quad \text{and} \quad \mathbf{II} : \text{out} \rightarrow \text{in} ,$$

where in and out are the initial and final states of a generic charged-current weak-interaction process which cannot be related by \mathcal{CP} - or \mathcal{CPT} -transformation. The symmetry \mathcal{T} is then violated if $P(\mathbf{I}) \neq P(\mathbf{II})$ because of the relation $\mathbf{I} \xleftrightarrow{\mathcal{T}} \mathbf{II}$.

Although measuring those two transitions seems like a simple enough task, the actual implementation is much less straightforward. The method employed by the analysis presented in this thesis depends on two specific event characteristics which are typical for so called B -factories, namely an entangled neutral B meson pair and a boosted centre-of-mass system. In the following chapter, the prospects for measuring \mathcal{T} -violation and the method that is used by this thesis are discussed.

III Methodology

Following the theoretical discussion, this chapter examines the prospects of measuring the \mathcal{T} -violation if there is any. In order to fully appreciate the actual solution, in Section 1 the different types of \mathcal{CP} -violation are discussed again in the hope of finding an approach that can also be used to measure \mathcal{T} . This should also provide some insight into previous approaches and attempts. Following that, in Section 2 the actual method that is used by this analysis is discussed.

1 From \mathcal{CP} to \mathcal{T}

Recall that, given that \mathcal{CPT} is invariant, each effect of \mathcal{CP} -violation should be accompanied by an effect of \mathcal{T} -violation of the same type. In the following section the different types of \mathcal{T} -violation are discussed in order to examine the prospects of performing a measurement tailored after existing \mathcal{CP} -violation studies. Refer to Section 2.2 in Chapter II for a discussion of the \mathcal{CP} -violation.

1.1 Direct \mathcal{T} -Violation

The most obvious candidate for realising the generic process presented in Section 3.2 of Chapter II is to invert a well-known, well-measured process. Unfortunately however, this is rather difficult in particle physics. Let us consider an example: a possible choice for a measurement is a generic decay and its inverse fusion process, i.e.

$$\mathbf{I} : A \rightarrow B + C \quad \text{and} \quad \mathbf{II} : B + C \rightarrow A,$$

or more explicit:

$$B^0 \xrightleftharpoons[\mathbf{II}]{\mathbf{I}} K^+ + \pi^-.$$

To prepare and measure process **I** is relatively simple, however, measuring the inverse fusion process **II** is not viable: the strong interactions in the inverse direction (**II**) completely swamp the feeble weak fusion process. The cross section of the weak fusion process is multiple orders of magnitude smaller than that of the strong interaction because it is highly suppressed.¹ Apart from that, there remains the problem that preparation of the input for **II** is extremely difficult.

Figure 5 shows the different decay channels that can be used to measure the strength of direct \mathcal{T} - or \mathcal{CP} -violation. For a \mathcal{CP} -violation measurement one can easily measure R_1 and R_2 whereas the rates R'_1 and R'_2 — which are required for a \mathcal{T} -violation measurement — cannot be measured with ease.

This example illustrates why directly measuring \mathcal{T} -violation is somewhat more difficult than the corresponding \mathcal{C} -, \mathcal{P} -, and \mathcal{CP} -violation measurements: the weak production mechanism is highly suppressed and the strong interaction is much more prominent, furthermore the preparation of in and out states is rather difficult.

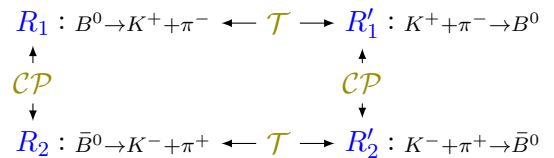


Fig. 5: Decay rates (marked in blue) for measuring \mathcal{T} - and \mathcal{CP} -violation. Arrows show processes which are connected by either transformation (marked in olive).

¹ For comparison: $\text{Br}(B^0 \rightarrow K^+ + \pi^-) \sim 10^{-5}$ [47]

It is clear that one could also choose to measure the \mathcal{T} -violation using a leptonic weak process instead of a hadronic one, for example by measuring the rates of

$$\mu \xrightleftharpoons[\text{II}]{\text{I}} e + \nu_\mu + \bar{\nu}_e.$$

However, even though there is no strong interaction involved in this process it is equally unfit for a measurement: fusion of the neutrinos and the charged lepton can be considered experimentally impossible. Furthermore, this probes the charged-current weak interaction for lepton vertices, whereas the previous method investigates \mathcal{T} -violation for a quark vertex.

1.2 \mathcal{T} -Violation in Mixing

In measurements of neutral meson oscillation, \mathcal{T} -violation is implied if $P(B^0 \rightarrow \bar{B}^0) \neq P(\bar{B}^0 \rightarrow B^0)$, refer to Figure 3 for a FEYNMAN diagram of such a process. However, this difference implies \mathcal{CP} -violation as well because \mathcal{T} and \mathcal{CP} are experimentally identical for a $\mathcal{CP}\mathcal{T}$ -even process such as B^0 — \bar{B}^0 mixing [47].

This measurement has been performed for K^0 — \bar{K}^0 mixing by the CPLEAR collaboration [50] and yielded \mathcal{T} non-invariance. Unfortunately, the measurement cannot be considered an independent \mathcal{T} -violation measurement because of its relation to \mathcal{CP} . The main goal of the CPLEAR experiment was to measure the \mathcal{CP} -violation and as a result thereof the flavour mixing asymmetry measured by CPLEAR is both \mathcal{CP} - and \mathcal{T} -violating.

Furthermore, the approach requires a non-vanishing decay width difference $\Delta\Gamma_K$ between the two neutral meson states. This dependence on $\Delta\Gamma_K$ has been cause of some controversy [6, 55, 56] regarding the interpretation of the observable. Given that $\Delta\Gamma \approx 0$, the proposed \mathcal{T} - and $\mathcal{CP}\mathcal{T}$ -odd observables will vanish themselves [47, 57] and therefore this technique should not be used with B^0 — \bar{B}^0 mixing where $\Delta\Gamma_{B_d}$ is negligible.

Neutrino physics can provide processes for measuring this type of \mathcal{T} -invariance as well: measuring that $P(\nu_e \rightarrow \nu_\mu) \neq P(\nu_\mu \rightarrow \nu_e)$ would imply \mathcal{T} -violation. However, measuring such processes poses an even more cumbersome task and once again probes the weak interaction for a lepton vertex.

1.3 \mathcal{T} -Violation in Interference

It is expected that the \mathcal{T} -violation in interference should be the experimentally most significant effect since the corresponding \mathcal{CP} -violation has the largest effect as well. The associated \mathcal{CP} -violation manifests in the decay rates for B^0 and \bar{B}^0 to the same \mathcal{CP} -eigenstate $f_{\mathcal{CP}}$ ² (following the notation used by this analysis $f_{\mathcal{CP}} \equiv B_{\mathcal{CP}}$).

The measurement technique performed for studies measuring the \mathcal{CP} -violation in interference is constructed such that there is no exchange of in and out states involved. Hence, an apparent asymmetry in a similarly formed \mathcal{T} study does not allow a direct \mathcal{T} -invariance measurement [3, 47]. Direct observation of this class would again require overcoming the irreversibility of the $B \rightarrow J/\Psi K^0$ system.

Refer to Figure 4 for an illustration of the general physics process that is exploited by this approach when measuring \mathcal{CP} -violation.

2 Redefining Transition

None of the approaches discussed in the previous section have been found to be promising. However, with a minor modification to the method presented in Section 1.1 a **direct \mathcal{T} -violation** measurement becomes feasible.

The experimental technique used by this analysis is limited to use at B -factories such as Belle and BABAR, where it was used for the first time [5]. It was initially presented in [58] and further discussed and refined in [46, 47, 57, 59].

The fundamental experimental problem when measuring \mathcal{T} -violation directly by comparing decay rates is that it is impractical or even outright impossible to perform the inverse fusion reaction.

² Typical final states are e.g.: for \mathcal{CP} -odd: $c\bar{c}K_S \equiv J/\Psi K_S$, $\Psi(2S)K_S$, $\chi_{c1}K_S$ and for \mathcal{CP} -even: $J/\Psi K_L$.

However, this problem vanishes if the out state is not chosen as a physical particle state but rather as an **eigenstate of the \mathcal{CP} operator**. Instead of measuring the decay of the physical particle B^0 to its daughter particles, the transition from the *flavour eigenstate* $|B^0\rangle$ to the *\mathcal{CP} eigenstate* $|B_{\mathcal{CP}}^0\rangle$ is measured.³ In the following section the experimental technique and the requirements enabling this approach are discussed.

2.1 Quantum Entanglement

The crucial requirement needed to implement the method used by *BABAR* is an event feature unique to B factories: the quantum entanglement of the neutral meson state. At B factories such as *Belle* and *BABAR*, electrons and positrons are collided at an energy that is just above to the invariant mass of the $\Upsilon(4S)$ resonance. The $\Upsilon(4S)$ is a resonant state of the bottomonium ($b\bar{b}$) which decays almost instantaneously to a pair of B mesons.⁴ The decay to a system of two neutral B_d^0 mesons occurs with a branching ratio of almost 50% and thus allows for large statistics [6].

Since the $\Upsilon(4S)$ has $J^{PC} = 1^{--}$, the binary state of the neutral B mesons is produced such that the two constituents are entangled: although the \mathcal{C} eigenvalue of the whole system is invariant, that of the individual mesons is not; this is called **EINSTEIN-PODOLSKI-ROSEN (EPR) entanglement** [60]. As a result of the EPR entanglement, the actual individual state of each constituent is not well defined prior to a measurement process (such as a decay). Having said this, it should be noted that the constituent states must be orthogonal at all times⁵ until their entanglement vanishes.

Furthermore, once one of these constituents is measured (i.e. it has decayed) the other may oscillate freely.⁶ The coherent state can be parametrised in different ways, typically this is chosen such that the orthogonal states correspond to future decay products. The parametrisations of the coherent state $|\psi\rangle$ that will be used for measuring the \mathcal{T} -violation are:

- Projection by flavour eigenstates: $B^0 \Leftrightarrow \bar{B}^0$

$$|\psi\rangle = \frac{1}{\sqrt{2}}[B^0(t_1)\bar{B}^0(t_2) - \bar{B}^0(t_1)B^0(t_2)]$$

- Projection by \mathcal{CP} eigenstates: $B_+ \Leftrightarrow B_-$

$$|\psi\rangle = \frac{1}{\sqrt{2}}[B_+(t_1)B_-(t_2) - B_-(t_1)B_+(t_2)]$$

The time values t_1 and t_2 are *labels* which denote the neutral B meson states B_1 and B_2 by means of their decay time. Furthermore, the actual meson states at those times are projected from their decay products |1> respectively |2>. The time values allow to define the **actual proper time difference** $\Delta\tau := t_2 - t_1 > 0$ that passed between the two decays [47]. Due to the orthogonality of the coherent state, a decay at t_1 yields a measurement of the state of both constituents at that time.

The states B_+ and B_- correspond to the even respectively odd \mathcal{CP} eigenstate of the neutral B_d meson; note the difference in notation to the charged mesons B^+ and B^- .

2.2 Measuring \mathcal{T} -Conjugate Transitions

In order to measure \mathcal{T} -invariance one must measure the decay rates of two \mathcal{T} -conjugate transitions. The observation of one such transition can be split into three well defined steps:

- 1st The decay of one of the entangled neutral B mesons — denoted by B_1 — is observed at some time t_1 . It decays to a definite flavour (or a definite \mathcal{CP}) final state and the final state particles *project* its pre-decay state (i.e. they allow to infer that state). This decay prepares the state of the other meson — denoted by B_2 — into its orthogonal state, i.e. $|B_1(t_1)\rangle \perp |B_2(t_1)\rangle$, hereby *tagging* it.

³ With $\mathcal{CP}|B_{\mathcal{CP}}^0\rangle = \xi_{\mathcal{CP}}|B_{\mathcal{CP}}^0\rangle$ we define $\mathcal{CP}|B_+\rangle = +|B_+\rangle$ and $\mathcal{CP}|B_-\rangle = -|B_-\rangle$.

⁴ With a net charge of zero, i.e. the $\Upsilon(4S)$ decays (almost) exclusively to either $B^0\bar{B}^0$ or B^+B^- .

⁵ This results from **BOSE** statistics for a system with $\mathcal{C} = -$, which requires the constituents to be in an antisymmetric (\equiv orthogonal) state [47].

⁶ The orthogonality of the components vanishes once the coherent state collapses, i.e. for $t > t_1$ in the formalism given below.

2nd The coherent state has collapsed and thus the remaining meson B_2 may propagate freely.

3rd At some time $t_2 > t_1$ the meson B_2 decays to a definite \mathcal{CP} (or a definite flavour) final state. Once again, the final state particles project a certain pre-decay state.

The steps of this method are illustrated in Figure 6 for two \mathcal{T} -conjugate transitions; the pair of those two transitions forms a *conjugate pair*. For the example given below, the two conjugate transitions are

$$\bar{B}^0 \rightarrow B_- \quad \text{and} \quad B_- \rightarrow \bar{B}^0, \quad (5)$$

however, note that the actual final states used to measure this transition (e.g. J/Ψ , $K_{S,L}$, l^\pm , etc.) are not the physical states of the transitions.

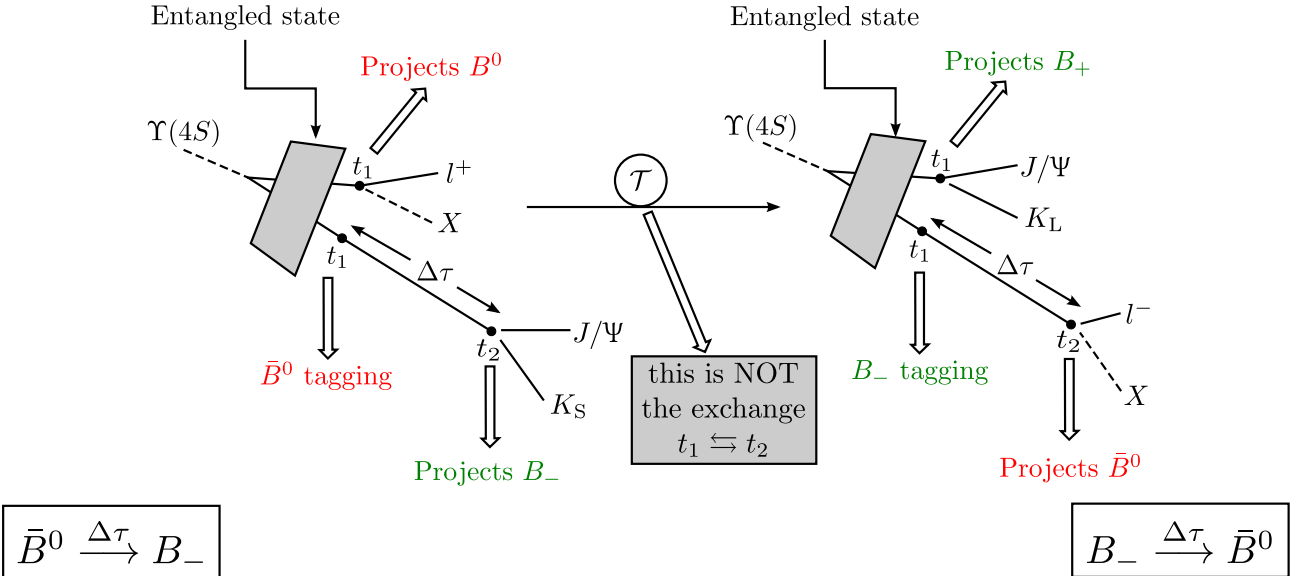


Fig. 6: Illustration showing the reconstruction method deployed for measuring two \mathcal{T} -conjugate transitions [47]. The action *projects* denotes the determination of a B meson's state from its daughter particles, the action *tagging* denotes that the B meson state has been inferred from the decay of its entangled partner.

It is important to note that the two transitions cannot be connected by simply exchanging $t_1 \leftrightarrow t_2$ but are connected only by genuine \mathcal{T} -transformation, i.e. the exchange of in and out states as well as the reversal of the direction of motion.

A measurement of \mathcal{CP} and flavour final state with their corresponding decay times yields a *time ordered final state* ($|1\rangle, |2\rangle$), i.e. the transitions shown in Figure 6 are defined by the states

$$\begin{array}{ll} \text{left hand side} & \text{right hand side} \\ (l^+ X, J/\Psi K_S) & (J/\Psi K_L, l^- X). \end{array}$$

There are three more conjugate transitions pairs, because there are a total of eight different time-ordered transitions; all four pairs of conjugate transitions are listed in Table 2.

Reference		\mathcal{T} -conjugate	
Transition	Final state	Transition	Final state
$\bar{B}^0 \rightarrow B_-$	$(l^+ X, J/\Psi K_S)$	$B_- \rightarrow \bar{B}^0$	$(J/\Psi K_L, l^- X)$
$B_+ \rightarrow B^0$	$(J/\Psi K_S, l^+ X)$	$B^0 \rightarrow B_+$	$(l^- X, J/\Psi K_L)$
$\bar{B}^0 \rightarrow B_+$	$(l^+ X, J/\Psi K_L)$	$B_+ \rightarrow \bar{B}^0$	$(J/\Psi K_S, l^- X)$
$B_- \rightarrow B^0$	$(J/\Psi K_L, l^+ X)$	$B^0 \rightarrow B_-$	$(l^- X, J/\Psi K_S)$

Tab. 2: Comparison pairs for \mathcal{T} -violation measurement [47].

Similarly, one can define pairs of conjugate transitions, which can only be connected by the \mathcal{CP} - and \mathcal{CPT} -transformations. These transitions can be used to measure the \mathcal{CP} - or \mathcal{CPT} -invariance respectively. Measuring the (already well measured) \mathcal{CP} -violation and the expected \mathcal{CPT} -invariance [36, 37] with the

method described here allows to cross-check the analysis procedure. The complete list of transition pairs which can be used for measuring \mathcal{T} - and \mathcal{CP} -violation and $\mathcal{CP}\mathcal{T}$ -invariance are included in Appendix B.

Initially, measuring the decay times t_1 and t_2 may not seem straightforward. However, a measurement of the decay times of the B mesons is fairly simple with the Belle experiment due to its boosted centre-of-mass frame, cf. Section 1.3 in Chapter IV.

2.3 Tagging the B Mesons

In addition to the *flavour tag* that is performed by standard \mathcal{CP} -violation measurements, the \mathcal{T} -violation measurement requires a \mathcal{CP} tag. Below the general scheme for obtaining both tags and the corresponding projection rules are discussed.

2.3.1 \mathcal{CP} Eigenstates

Tagging of the \mathcal{CP} eigenvalue $\xi_{\mathcal{CP}}$ of a neutral B meson \mathcal{CP} eigenstate, labeled as $B_{\mathcal{CP}}$, can be achieved by following a decay chain involving only pure, well-defined \mathcal{CP} eigenstates. The approach used by this analysis reconstructs hadronic final states in order to infer $\xi_{\mathcal{CP}}$; the projection rules are given by

$$\underbrace{B^-}_{-} \rightarrow \overbrace{J/\Psi K^+}^{+} \quad \text{and} \quad \underbrace{B^+}_{+} \rightarrow \overbrace{J/\Psi K^-}^{+} \quad \downarrow \pi\pi \quad \underbrace{\pi\pi\pi}_{-}$$

where the \mathcal{CP} eigenvalue $\xi_{\mathcal{CP}}$ of the component states is shown in red. Unfortunately however, the \mathcal{CP} eigenstates K_{\pm} are not directly experimentally accessible either, they will be approximated by K_S and K_L respectively. Approximating K_+ by K_S and assuming that $K_S \rightarrow \pi\pi$ is always true introduces small corrections, $\mathcal{O}(10^{-3})$, which will be neglected; the same is true for the $K_L \mapsto K_-$ matching.

Obviously, one may choose to reconstruct the J/Ψ component from other charmonium ($c\bar{c}$) states, such as $\Psi(2S)$. A larger data sample can also be obtained by extending the \mathcal{CP} tag reconstruction to include other \mathcal{CP} -final states, however, this might impact the accuracy of the tag and the purity of the data sample. Furthermore, in some channels it is possible to include a radiated, hard photon and then perform an angular analysis of the daughter particles because the \mathcal{CP} tag is encoded in their angular distribution [5, 47].

2.3.2 Flavour Eigenstates

There are different approaches that can be used to reconstruct the decay-time flavour ξ_{Flv} of a B^0 meson. The obvious, albeit cumbersome, approach is to perform a full reconstruction of the meson and infer its flavour from the daughter particles.

Typically, this is done by reconstructing B mesons which decay semileptonically: the charge of the lepton can then be used to trace back the decay-time flavour of the B meson. The projection rules are given by

$$B^0 \rightarrow l^+ X \quad \text{and} \quad \bar{B}^0 \rightarrow l^- X'$$

where l^{\pm} denotes a charged lepton final state and X/X' is the corresponding hadronic final state⁷; Figure 7 shows the FEYNMAN diagrams of these processes for illustration.

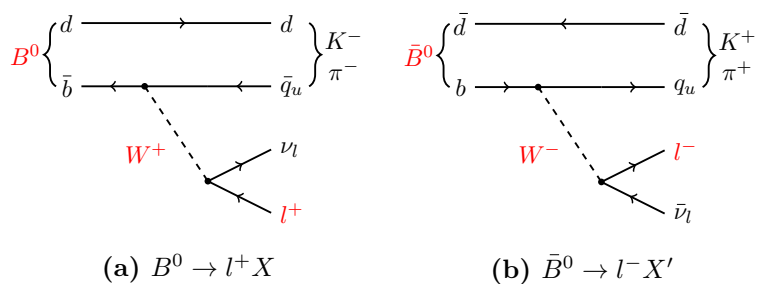


Fig. 7: Tree level FEYNMAN diagrams of semileptonic neutral B meson decay. Marked in red are particles that can be used to trace the flavour of the B meson (also marked in red).

Similarly, one may use the charge of prompt or second level kaons/pions. However, the leptonic tag provides the best tag efficiency and purity. Essentially, using the charge of hadrons instead of leptons is more prone to error because hadrons are generally more difficult to reconstruct than leptons.

⁷ Obviously this is the observed final state, i.e. neutrinos/MET are not included.

Distinguishing prompt and second-level particles is more difficult for hadrons than (charged) leptons because the momentum and track resolution is much better for the latter. Hence, correct interpretation of a decay chain such as

$$\begin{aligned} \bar{B}^0 &\rightarrow D^{+(*)} e^- \bar{\nu}_e \\ &\quad \hookrightarrow K^- \pi^+ \mu^+ \nu_\mu \end{aligned}$$

can be difficult: depending on whether the K^- and/or π^+ are interpreted as prompt (cf. Figure 7) or second level particles, the estimated flavour of the B changes.

For maximum efficiency and accuracy, different flavour tag inputs are combined using a neural network. This network then provides a likelihood for its flavour tag hypothesis, therefore propagating some measure of credibility of the tag to the user. Section 6.1 in Chapter V discusses the flavour tagging routine used by this analysis and results thereof in some more detail.

2.4 Quantifying the \mathcal{T} -Violation

In this section, the approach that shall be used to quantify the extent of the \mathcal{T} -violation is discussed.

2.4.1 Nomenclature

Each of the eight transitions building up the process pairs (cf. Table 2) that are used to measure a symmetry \mathcal{S} has a *time-dependent decay rate* $g_{\alpha,\beta}^\omega(\Delta\tau)$, where

- subscript $\alpha \in \{l^+, l^-\}$ denotes the **flavour final state**,
- subscript $\beta \in \{K_S, K_L\}$ denotes the **\mathcal{CP} final state**,
- superscript ω denotes decay order:
 - $\omega = + \Rightarrow$ decay to α occurred **before** the decay to $\beta \Rightarrow t_\alpha < t_\beta$.
 - $\omega = - \Rightarrow$ decay to α occurred **after** the decay to $\beta \Rightarrow t_\alpha > t_\beta$.

Note that the subscripts denote the measured final states rather than those of the transition used for measuring \mathcal{T} -violation, cf. Section 2.2.

The *signed proper time difference* Δt encodes the decay order of the two final states in its sign:

$$\Delta t = t_\beta - t_\alpha = \overset{\text{flavour tag}}{\downarrow} \overset{\pm}{\Delta\tau},$$

$$\overset{\text{CP-tag}}{\uparrow} \quad \quad \quad \uparrow \omega$$

where t_i are the tagging times and $\Delta\tau$ is the *actual proper time difference* (as defined in Section 2.1).

Using the signed proper time difference simplifies how measurement inefficiencies and ambiguities are considered in the minimisation that is used to determine the symmetry parameters.

2.4.2 Applying \mathcal{T}

A replacement $\omega = + \rightarrow \omega = -$ (and vice versa) corresponds to Δt exchange. However, because of the technique used to measure both the initial and final state of the transitions, cf. Section 2.2, one must switch the values held by ω , α , and β . Doing so ensures that the requirement for genuine \mathcal{T} -violation, cf. Section 3 in Chapter II, is fulfilled.

From the example given by Equation (5) and depicted in Figure 6 we can measure the decay rate for $\alpha = l^+$, $\beta = K_S$, and $\omega = +$ that is to say g_{l^+, K_S}^+ (for the left hand side). Hence, in order to measure the \mathcal{T} -conjugate transition, the right hand side's decay rate must be given by

$$\left. \begin{array}{l} \alpha = l^+ \quad \xleftarrow{\mathcal{CP}} \quad \alpha' = l^- \\ \beta = K_S \quad \xleftarrow{\mathcal{CP}\mathcal{T}} \quad \beta' = K_L \\ \omega = + \quad \quad \quad \omega' = - \end{array} \right\} \Rightarrow g_{l^+, K_S}^+ \xleftarrow{\mathcal{T}} g_{l^-, K_L}^-,$$

which is also depicted in Figure 6. For \mathcal{CP} and $\mathcal{CP}\mathcal{T}$ the swapping rules for the indices are given by:

- **For \mathcal{CP} :** This requires applying \mathcal{CP} to a reference transition, e.g. to $B^0 \rightarrow B_{\mathcal{CP}}$, where $B_{\mathcal{CP}}$ is some (approximate) \mathcal{CP} eigenstate (e.g. B_- or B_+ which obviously does not change under \mathcal{CP}). The conjugate transition $\bar{B}^0 \rightarrow B_{\mathcal{CP}}$ corresponds to an exchange of the value held by α .

- **For $\mathcal{CP}\mathcal{T}$:** Sequential application of \mathcal{CP} and \mathcal{T} to a reference transition such as $\bar{B}^0 \rightarrow B_{\mathcal{CP}}$ yields $B_{\mathcal{CP}} \rightarrow B^0$. In terms of final states this mapping corresponds to swapping the values held by β and ω , i.e. exchange of the \mathcal{CP} final state by its conjugate state and reversing of the decay order.

From this it is obvious that the combination of \mathcal{CP} and $\mathcal{CP}\mathcal{T}$ equals \mathcal{T} , which is just the result to be expected from the theory of symmetries.

2.4.3 Explicit Form

Quantum mechanics can be used to derive a general formula for the time-dependent decay rate $g_{\alpha,\beta}^{\omega}(\Delta\tau)$. This formula can be simplified assuming that $\Delta\Gamma \approx 0$ where $\Delta\Gamma$ is the decay width difference of the two oscillating neutral meson states. Since this approximation holds true for the B meson system, it can (and will) be applied in the analysis. The full formula is given in [47], the approximated solution is given by Equation (6)

$$g_{\alpha,\beta}^{\omega}(\Delta\tau) \propto e^{-\Gamma\Delta\tau} \left[1 + \underbrace{C_{\alpha,\beta}^{\omega} \cdot \cos(\Delta m\Delta\tau)}_{\text{interference from different weak and strong phases}} + \underbrace{S_{\alpha,\beta}^{\omega} \cdot \sin(\Delta m\Delta\tau)}_{\text{interference w/ and w/o mixing}} \right], \quad (6)$$

where Δm is the mass difference, Γ is the average decay width of the two neutral states, and the coefficients $C_{\alpha,\beta}^{\omega}$ and $S_{\alpha,\beta}^{\omega}$ are oscillation parameters whose sub- and superscripts follow the notation introduced previously. This formula is by no means limited to a \mathcal{T} -violation measurement: it is the generic time-resolved decay probability distribution for a state that was once part of the coherent state $|\psi\rangle$ (cf. Section 2.1).

The two oscillation components stem from different oscillation contributions, see Equation (6). In the context of the SM⁸ it can be assumed that $C_{\alpha,\beta}^{\omega} \approx 0$ and that $S_{\alpha,\beta}^{\omega}$ is significantly different from zero. For a symmetry \mathcal{S} to be an invariant transformation, the absolute value of the coefficients of two \mathcal{S} -conjugate processes must be equal. It is thus useful to construct *asymmetry parameters* $\Delta C_{\mathcal{S}}^{\omega}$ and $\Delta S_{\mathcal{S}}^{\omega}$, which are differences of the coefficients for conjugate processes. The corresponding list of asymmetry parameters is given in Table 13 in Appendix C.

If any of these asymmetry parameters significantly differs from zero, one must assume that the corresponding transformation is violated.

2.4.4 Fitting to Data

For each event which has yielded a probable flavour and \mathcal{CP} tag, the signed proper time difference Δt is calculated. The flavour and \mathcal{CP} tags form four groups of events, for each group a distribution in Δt is obtained. By performing a fit to these distributions one can obtain the asymmetry parameters from data.

The fit should be performed *simultaneously*, meaning to all four groups concurrently, in order to obtain maximal significance of the result and enable the fit to counteract the effect of mis-tags. It is expected that the number of reconstructed events will shrink significantly for increasing Δt . This is a result of the exponential dampening that can be found in Equation (6). Therefore, it is sensible to either perform an unbinned fit or collect the data in bins of varying width and use a fit routine which can operate on this data.

The natural choice for the PDF is the time dependent decay rate $g_{\alpha,\beta}^{\omega}$ given by Equation (6). However, this formula neglects experimental uncertainties, among them the mis-tag rate (and thus dilution) of the flavour tagging output. This can be solved by including the flavour mis-tag in Equation (6), thus resulting in

$$\tilde{g}_{\alpha,\beta}^{\omega}(\Delta\tau) \propto (1 - w_{\text{tag}}) g_{\alpha,\beta}^{\omega}(\Delta\tau) + w_{\text{tag}} g_{\bar{\alpha},\beta}^{\omega}(\Delta\tau), \quad (7)$$

where $\bar{\alpha}$ denotes the flavour state orthogonal to α and w_{tag} gives the fraction of flavour states reconstructed with the wrong flavour (i.e. $l^- X$ reconstructed as B^0 or $l^+ X$ reconstructed as \bar{B}^0).

⁸ I.e. assuming \mathcal{T} -violation.

Admittedly, the flavour mis-tag is not the only significant source of mis-assignment. The experimental resolution of the time difference reconstruction mechanism introduces a smearing such that the observed proper time difference $\Delta\tau_{\text{rec}}$ may take unphysical negative values. Consequentially the sign of Δt may not be used directly to infer the order of the decays (i.e. the value of ω). To overcome this limitation, the PDF is not constructed as a function of $\Delta\tau$ (or Δt for that matter) but rather as a function of the reconstructed signed time difference Δt_{rec} . Thence, the measured decay rates are convolutions of the true decay rates and a resolution function \mathcal{R} .

Following this scheme, the final PDF is no longer a function of $\Delta\tau$ but rather Δt_{rec} and is not parametrised in terms of ω . It is given by

$$\mathcal{H}_{\alpha,\beta}(\Delta t_{\text{rec}}) \propto \underbrace{\hbar_{\alpha,\beta}^+(\Delta t) \cdot H(\Delta t)}_{\text{flavour first}} \otimes \mathcal{R}(\delta t; \sigma_{\Delta t_{\text{rec}}}) + \underbrace{\hbar_{\alpha,\beta}^-(-\Delta t) \cdot H(-\Delta t)}_{\mathcal{CP} \text{ first}} \otimes \mathcal{R}(\delta t; \sigma_{\Delta t_{\text{rec}}}), \quad (8)$$

where $\hbar_{\alpha,\beta}^\omega$ is defined by Equation (7), Δt (Δt_{rec}) is the true (reconstructed) signed proper time difference, $H(\Delta t)$, is the HEAVISIDE step function, and $\mathcal{R}(\delta t; \sigma_{\Delta t_{\text{rec}}})$ is the resolution function which depends on $\delta t := \Delta t_{\text{rec}} - \Delta t$ and the estimated uncertainty on the signed proper time difference $\sigma_{\Delta t_{\text{rec}}}$. This resolution function is modelled as sum of multiple (*BABAR*: 3, *Belle*: 2) GAUSS functions [5, 47, 61]. Typically, the mean μ and width σ of the composing distributions as well as their relative contribution to \mathcal{R} can be extracted from MONTE-CARLO studies.

The first term in Equation (8) is related to the case $\Delta t > 0$, meaning that the flavour tagged meson B_{FLV} decayed before the \mathcal{CP} tagged meson $B_{\mathcal{CP}}$, while the second term is associated to the case $\Delta t < 0$, i.e. the decay to a flavour state occurred later. Consequently, the distribution $\mathcal{H}_{\alpha,\beta}(\Delta t_{\text{rec}} > 0)$ is dominated by events with $\Delta t > 0$ with a dilution of *wrong-timed events*, i.e. events with $\Delta t < 0$.⁹

⁹ Similarly $\mathcal{H}_{\alpha,\beta}(\Delta t_{\text{rec}} < 0)$ is composed predominantly of events with $\Delta t < 0$ and diluted by $\Delta t > 0$ events.

IV Experimental Apparatus

In this chapter the experimental and technical setup of the Belle experiment and the data used by this analysis are described. Section 1 briefly discusses the accelerator that facilitates the Belle experiment. In the following Section 2, the Belle detector itself — the experiment’s most fundamental core — is described. Finally, in Section 3 the experimental and simulated data samples which are used throughout this analysis are presented.

The Belle experiment is a high-energy particle physics experiment conducted by the Belle collaboration, a multinational group of scientists, based at the High Energy Accelerator Research Organisation (KEK) in Tsukuba, Japan. Although operation of the KEKB accelerator and Belle detector has ceased in 2010, ongoing and new analyses are still being performed using the collected data. The experiment’s main objective, the discovery and measurement of the \mathcal{CP} -violation in the neutral B meson system, has been long fulfilled.

1 The KEKB Accelerator Facility

The KEKB accelerator [6, 62–65] was an asymmetric e^+e^- collider which was operated by the KEK. It was located in Tsukuba, Japan, with its only interaction point (IP) being occupied by the Belle detector.

The collider had a circumference of 3 km and was operated as a synchrotron with a single linear injection accelerator. It was composed of two storage rings, one for each particle type. The electrons and positrons were accelerated to their nominal momenta prior to injection into the storage rings. At the interaction point the two beams met under a small inclination (also called beam crossing angle) of $\theta_{\text{Cross}} = 22 \text{ mrad}$; this nonzero crossing angle was chosen for purely technical reasons [66].

Figure 8 shows the schematic layout of the KEKB accelerator complex including the locations of the interaction point (Tsukuba Area) and the injection system (Fuji Area).

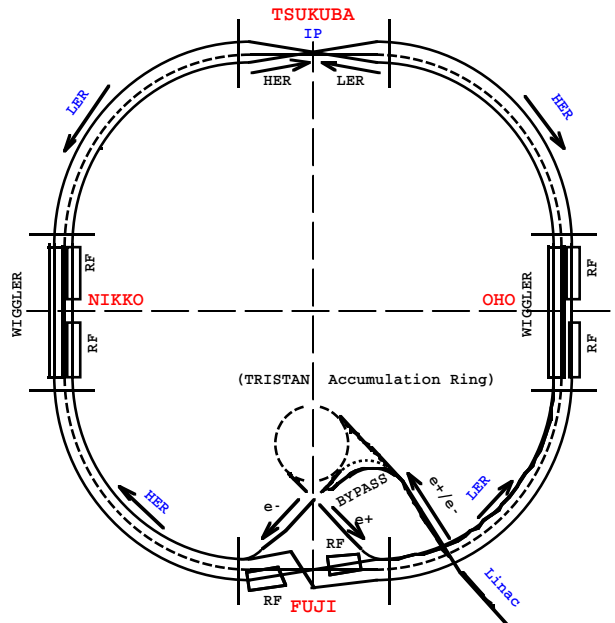


Fig. 8: Schematic drawing of the KEKB accelerator complex and its experimental areas. Figure taken from [62].

For most of its operational time the centre-of-mass energy was set to $\sqrt{s} = 10.58 \text{ GeV}$, corresponding to the mass of the $\Upsilon(4S)$ bottomonium resonance. The corresponding beam energies are $E(e^-) = 8.0 \text{ GeV}$ and $E(e^+) = 3.5 \text{ GeV}$ for the electron and positron beam respectively. Operation at this specific centre-of-mass energy is what makes the KEKB a so called B -Factory.

1.1 KEKB as B -Factory

The mass of the $\Upsilon(4S)$ is just 20 MeV above the $B\bar{B}$ production threshold. However, since the $\Upsilon(4S)$ is a bottomonium state, the branching fraction to a $B\bar{B}$ state is at more than 96% [6].

During the 1980s the CUSB and CLEO experiments extensively studied the $\Upsilon(nS)$ bottomonium resonances; Figure 9 shows a combined measurement of the hadron production cross section in e^+e^- collisions $\sigma(e^+e^- \rightarrow \text{hadrons}, s)$ as a function of the centre-of-mass energy of \sqrt{s} . The different $\Upsilon(nS)$ resonance peaks can be identified easily, note that the $\Upsilon(4S)$ peak lies above the $B\bar{B}$ production threshold and is much less distinct when compared to the lower energy resonances.

Figure 10 shows the hadronic R ratio defined as

$$R_{\text{Had}}(s) := \frac{\sigma(e^+e^- \rightarrow \text{hadrons}, s)}{\sigma(e^+e^- \rightarrow \mu^+\mu^-, s)},$$

where $\sigma(e^+e^- \rightarrow \mu^+\mu^-, s)$ is the muon production cross section in e^+e^- collisions. In this plot the y -axis is scaled such that the peak positions of the low energy resonances ($n < 4$) lie outside the plotted range. Clearly the low energy peaks are much more narrow than the $\Upsilon(4S)$ peak.

The cross section for bottomonium production at the $B\bar{B}$ production threshold ($\sqrt{s} = 10.58 \text{ GeV}$) can be determined to be 1.1 nb [67] which is roughly a third of the production cross section for $q\bar{q}$ ($q \in \{u, d, s, c\}$). Other relevant processes total to about 50 nb , these include τ -pair production, B_{HABHA} scattering, and two-photon interactions. However, the latter backgrounds can be removed efficiently using software triggering. The resulting efficiencies are 99.1% for the bottomonium and 79.5% for $q\bar{q}$; the remaining contributions are reduced such that they total about 0.2 nb of data [67].

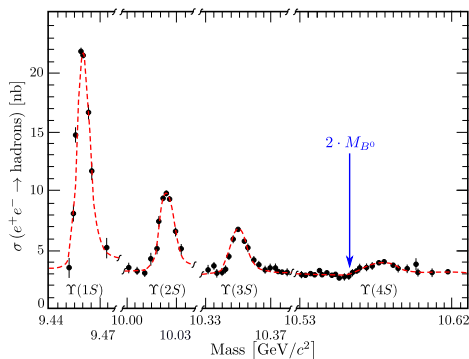


Fig. 9: Combined cross section measurement for the process $e^+e^- \rightarrow \text{hadrons}$ around the bottomonium production threshold by the CUSB and CLEO experiments [68–71]; the $B\bar{B}$ production threshold is marked in blue.

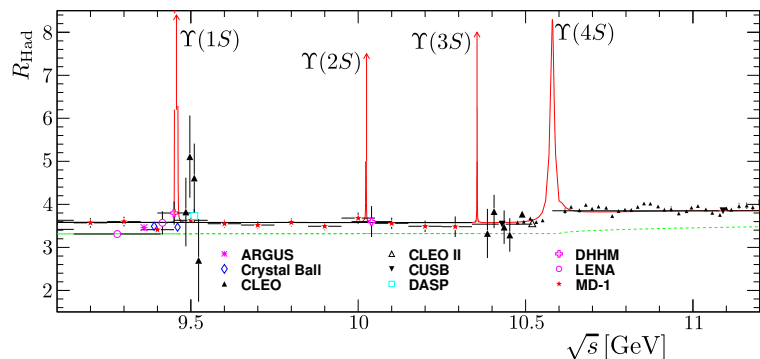


Fig. 10: $R_{\text{Had}}(s)$ ratio as function of centre-of-mass energy \sqrt{s} for the beauty threshold region. Figure taken from [6], refer to [72] for an in-depth discussion of the depicted plot.

Since the KEKB accelerator operated in this energy regime and due to the aforementioned features of doing so, it produced copious amounts of $B^0\bar{B}^0$ and B^+B^- pairs. This makes it an ideal apparatus for performing \mathcal{CP} -related studies.

1.2 Data and Luminosity

The KEKB accelerator became operational in December 1998 and data taking, by the Belle experiment, commenced about half a year later on the 1st of June 1999. During the following eleven years, Belle logged a total integrated luminosity of $\mathcal{L}_{\text{Int}} = 1040 \text{ fb}^{-1}$ (about 700 fb^{-1} on resonance) which corresponds to about $770 \cdot 10^6 B\bar{B}$ -pairs produced on-resonance. The instantaneous luminosity reached peak values of $\mathcal{L} = 2.1 \cdot 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ corresponding to roughly $20 B\bar{B}$ pairs produced per second. With these values the KEKB surpassed its design specifications and set world records for the machine with the highest instantaneous and integrated luminosity which are still unbroken to date [6].

In 2007 the KEKB received an upgrade which included installation of *crab cavities* [73]. This addition enabled harnessing the full beam power at the interaction point. It facilitated the luminosity records since it counteracts the loss of luminosity due to the nonzero crossing angle. Without crab cavities the beams do not collide truly head-on. This is caused by the shape of the particle bunches which are stretched along the direction of the beam. Therefore, two colliding bunches have a reduced geometrical cross section if they collide under an angle. The crab cavities give the particles in different regions of a

bunch a kick in opposite directions thus rotating the bunch in the plane that is transverse to the beam. At the interaction point the incident bunches are tilted with respect to the beam axis, however, doing so restores the full geometrical cross section of the collision.

The modus operandi of Belle followed the usual scheme of data runs lasting between six and nine months each year followed by shutdown periods dedicated to machine maintenance and upgrade. On the 30th of June 2010 operation of KEKB and hence Belle concluded in its final shutdown. Figure 11 shows the time evolution of the integrated luminosity logged by Belle.

Belle collected most of its data with the accelerator running on the $\Upsilon(4S)$ resonance (about 71 %, cf. [74]). Additional data runs were taken at the $\Upsilon(nS)$, $n \in \{1, 2, 3, 5\}$ resonances. During later periods of data taking, energy scans between the $\Upsilon(4S)$ and $\Upsilon(6S)$ were performed. On top of that, about 10% of the running time were dedicated to off-resonance data taking, at 60 MeV below the respective resonance peak, for non- $B\bar{B}$ background estimation and studies. These off-resonance runs were conducted about every two months in order to ensure chronological consistency of the measurement with the on-resonance data.

1.3 Asymmetric Beam Energies

The beam energies of the KEKB were chosen to be asymmetrical in order to obtain a boosted centre-of-mass system. Due to the small energy difference between the centre-of-mass energy and the $\Upsilon(4S)$ mass, the bottomonium state would be nearly at rest in the laboratory frame for a symmetric setup.

By boosting the $\Upsilon(4S)$ system and hence its daughter particles, the experiment gains a handle onto the time axis: in total the $B\bar{B}$ mesons will retain the $\Upsilon(4S)$ boost, however, since the mass difference between the bottomonium state and the $B\bar{B}$ mesons is small, it can be expected that each meson individually approximately retains the boost as well,¹ hence $\beta\gamma|_{\Upsilon(4S)} \approx \beta\gamma|_B \approx \beta\gamma|_{\bar{B}}$.

Since the boost $\beta\gamma$ is a measure of velocity, one can calculate the proper time $d\tau$ which passed in the rest frame of the particle while traversing a distance dz . Hence, having a boosted centre-of-mass system enables to perform a time-resolved measurement of the B mesons decay; the relation for calculating the proper time from the measured distance is given by

$$d\tau = \frac{dz}{c \cdot \beta \cdot \gamma}. \quad (9)$$

The $\Upsilon(4S)$ boost can be calculated from the known energies of the electron and positron beams² and the beam crossing angle θ_{Cross} using

$$\beta\gamma = \frac{|\vec{p}_{\Upsilon(4S)}|}{E(\Upsilon(4S))} \cdot \frac{E(\Upsilon(4S))}{M_{\Upsilon(4S)}} = \frac{\sqrt{E(e^+)^2 + E(e^-)^2 + 2E(e^+)E(e^-) \cdot \cos(\theta_{\text{Cross}})}}{M_{\Upsilon(4S)}}$$

to be $\beta\gamma \approx 0.425$. On average, the B mesons will decay after traversing a distance of about 200 μm measured from the $\Upsilon(4S)$ production vertex.

2 The Belle Detector

The Belle detector [66, 74] is a multi-purpose magnetic spectrometer with an angular coverage of 4π that was built around the interaction point at the experimental area Tsukuba of the KEKB accelerator.

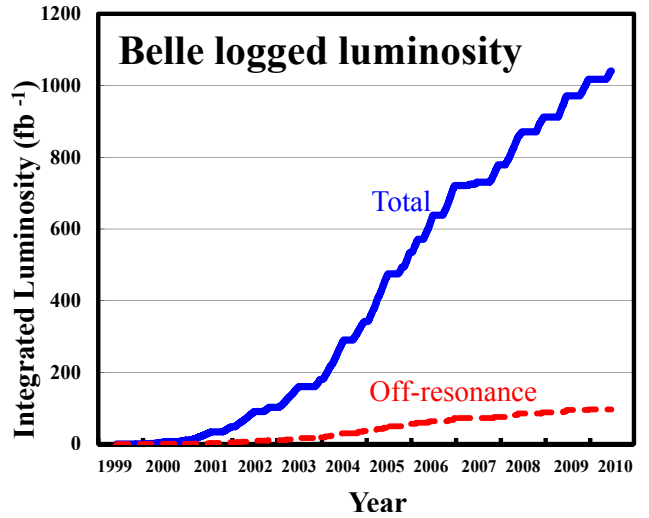


Fig. 11: Time evolution of the integrated luminosity logged by Belle. Figure taken from [74] which also contains a detailed listing of \mathcal{L}_{Int} broken down by centre-of-mass energy.

¹ The accuracy of this assumption depends on the amount of transverse momentum p_T the B mesons have.

² These energies are $E(e^-) = 8.0 \text{ GeV}$ and $E(e^+) = 3.5 \text{ GeV}$ respectively.

It was designed as a precision instrument for measuring \mathcal{CP} -violation using time-dependent analyses but has since served many analyses including meson spectroscopy and rare decay searches.

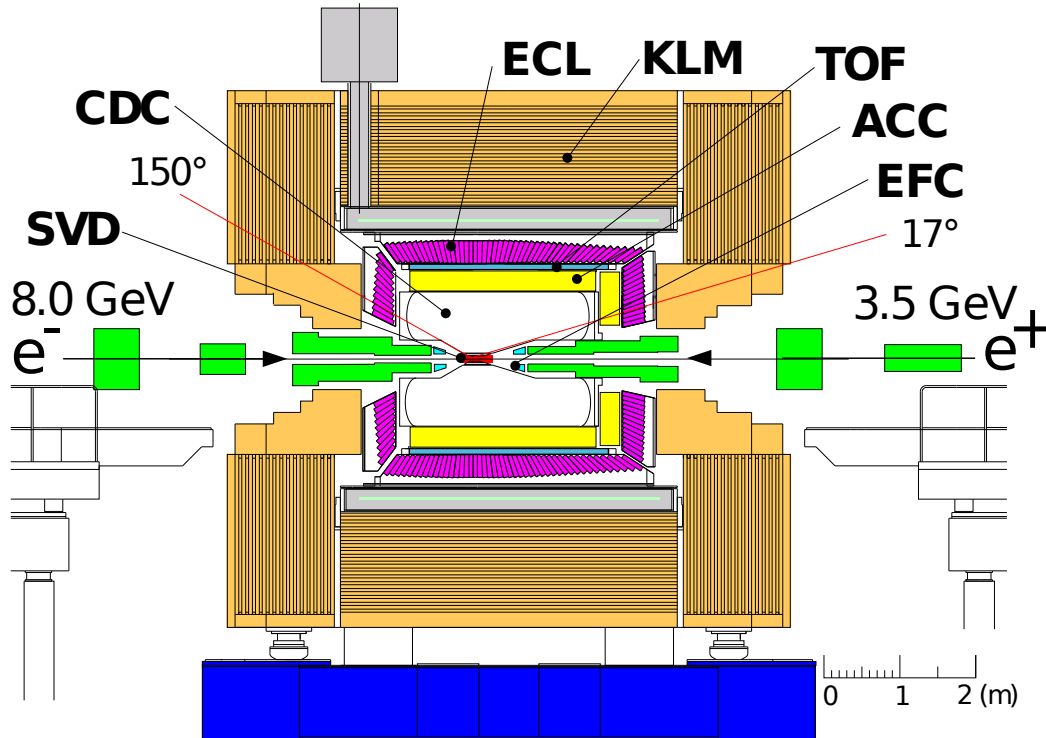


Fig. 12: Schematic drawing of the Belle detector and its subsystems [66].

Figure 12 shows a schematic drawing of the detector and its various subsystems. It is composed of barrel, forward, and rear components, however, it is not perfectly symmetric with respect to the coordinate origin: a small forward-backward asymmetry was introduced to match the expected asymmetry of the physical events resulting from the boosted centre-of-mass system.

The detector is constructed such that its solenoid's axis is aligned with the z -axis and thus in good approximation with the beam axis. This design minimises the **LORENTZ** force acting on the incident low-energy positron beam.

2.1 Detector Subsystems

The different subsystems of the detector each fulfil specialised needs of the experiment: the **Silicon Vertex Detector** (SVD) and the enclosing **Central Drift Chamber** (CDC) provide precision tracking and a fine-grained vertex measurement. Particle identification, especially for charged kaons and pions, is achieved using the dE/dx measurements taken from the CDC, a system of **Time-of-Flight Counters** (ToF), and a set of **Aerogel Cherenkov Counters** (ACC). Photons and neutral particles are reconstructed by an **Electromagnetic Calorimeter** (ECL) — extended into the end caps by the **Extreme Forward Calorimeter** (EFC) — which also helps at identifying electrons. These inner detector subsystems are bathed in the magnetic field of a superconducting solenoid which enables measuring the momenta of charged particles. Even further outwards is the **K_L/μ Detector Subsystem** (KLM) which is dedicated to the detection of the long lived K_L meson and muons.

§1. The Silicon Vertex Detector (SVD) [66, 75–77] is the innermost detector subsystem and is pivotal to the tracking of charged particles. It enables high resolution vertex reconstruction which is essential for accurate tracking resolution and time-resolved analyses. Its precise and efficient operation is crucial to this analysis.

Due to the low expected momentum of particles produced at Belle, the vertex resolution of the detector is expected to be dominated by **COULOMB**-scattering. Hence, strict limits on the material footprint of the two innermost detector components, SVD and CDC, were imposed. Furthermore, the innermost layer of the SVD was moved as close as possible to the IP in order to improve the vertex resolution.³

³ The proximity in turn increases the expected radiation dose the respective SVD layers must tolerate.

The subsystem is built from four layers⁴ of double-sided silicon strip detectors arranged to form a barrel. The layers are spaced at radii of 2.0, 4.35, 7.0, and 8.8 cm and surround the beam pipe. They are covering the polar angle region of $17^\circ < \theta < 150^\circ$,⁵ where θ is in the laboratory frame and measured relative to the z -axis. The detector strips have pitches of 75 and 73 μm along z and 50 and 65 μm for the $r\phi$ direction.

§2. The Central Drift Chamber (CDC) [66, 78, 79] is a small-cell cylindrical drift chamber. Its main purpose is to provide second level tracking and to measure the particle momentum from their track curvature.

The chamber is filled with a gas mixture acting as interaction material for impinging particles: electrons and gas-ions are produced in the chamber by ionisation as a result of the incident particle's passage; the former drift towards the anode wires where they can be detected. From the amount of signal charge measured by the readout for a given sensitive volume, the energy deposited in that volume can be inferred. The relatively large volume of the CDC allows measuring the energy loss per unit path dE/dx within several independent volume segments. This quantity can then be used for particle identification. The resolution obtained by this setup is $\sigma(dE/dx) = 6.9\%$ for minimum ionising particles.

By utilising a gas mixture of low Z -value the adverse effects of multiple COULOMB scattering are minimised while at the same time retaining the ability to perform a dE/dx measurement. The chamber is traversed by 50 layers of anode wires (including 18 layers of stereo wires) which constitute the actual detection component of the CDC. This geometrical setup provides three-dimensional measurements of the trajectories of charged particles in the same polar angle coverage region as the SVD.

The trajectory measurement from the CDC is matched to SVD measurements in order to get a greatly improved vertex resolution. Due to the strong magnetic field of the solenoid (cf. §6), charged particles traversing the CDC will be forced on a curved trajectory. By measuring the curvature of their trajectory one can then reconstruct their respective momenta.

§3. The Aerogel CHERENKOV Counter (ACC) [66, 80, 81] provides additional particle identification for high momentum incident particles (above 1.2 GeV). The main purpose of this subsystem is to provide separation between the signals of charged pions and kaons. It encloses the CDC along the barrel and in the forward direction.

Fundamentally, the ACC is a granular detector whose cells measure the CHERENKOV light emitted by impinging charged particles in the aerogel radiators. CHERENKOV light is a phenomenon which emerges when a particle traverses a material at a velocity which exceeds the speed of light in that medium.

The ACC used by Belle is a *threshold* type CHERENKOV detector which identifies a particle's species by determining whether CHERENKOV light is emitted or not. This information allows discriminating between a heavy particle hypothesis (kaon, should not radiate) and a light particle hypothesis (pion, should radiate) for a given (independently measured) incident particle momentum. By combining radiators with different refractive indices, different separation regions can be defined.⁶

This subsystem is built from 1188 aerogel blocks which have refractive indices varying between 1.01 and 1.03 depending on the polar angle. Each ACC module is fitted with a fine-mesh photo-multiplier which measures the CHERENKOV radiation emitted in the aerogel radiator module.

§4. The Time-Of-Flight System (ToF) [66, 82, 83] encloses the ACC in the barrel region. Quite like the ACC, the ToF system provides particle identification by means of estimating an impinging particle's mass from its momentum and velocity.

⁴ Up until 2003 there were only three layers arranged with a slightly tighter spacing.

⁵ Prior to the 2004 upgrade the polar angle coverage was $23^\circ < \theta < 139^\circ$.

⁶ In principle this also allows distinguishing between more than two types of particles, however, Belle uses the ACC primarily for kaon to pion separation. However, proton to pion separation has been shown to work quite well with the setup [81].

The ToF system aims to measure the time difference between the collision time at the IP and the passage of a daughter particle through the scintillator, thus measuring the daughter particle's velocity. Given an independent momentum measurement, the velocity can then be used to reject or accept a certain mass hypothesis (and thus particle type hypothesis). The time-resolution of the ToF provides three standard deviations (3σ) of separation between charged kaon and pion signals for momenta below 1 GeV and 2σ up to momenta of 1.5 GeV.

The ToF system is composed of 128 plastic scintillators; the scintillation photons are measured on both sides of the scintillator pieces by fine-mesh photo-multipliers. Angular coverage of the ToF system is $34^\circ < \theta < 120^\circ$ for the polar angle.

§5. The Electromagnetic Calorimeter (ECL) [66, 84, 85] provides calorimetry to the Belle detector: its main purpose is the measurement of the energy deposition of electrons and photons.

An impinging particle creates an electromagnetic shower within the ECL's cells. The scintillating material within the cells gets excited by the shower particles and emits some light during de-excitation. This emission is then measured by photo-diodes located at the end of each cell. The amount of light collected by the readout system is proportional to the shower size and thus the energy of the impinging particle.

The ECL subsystem is constructed from a total of 8736 CsI(Tl) crystals around the barrel and on the end caps. Each calorimeter cell is fitted with a photo-diode at its end, the readout system is installed in the solenoid's coil. The angular coverage of the ECL is the same as for SVD and CDC; the thickness of the whole ECL system is approximately 16 times the radiation length X_0 of the material.

Electrons (and positrons) are detected by bremsstrahlung and pair-production, yielding characteristic cluster shapes in the ECL. This can be used along with the dE/dx measurement from the CDC and the ratio of ECL cluster energy and absolute particle momentum to provide a highly significant electron identification scheme.

§6. The Solenoid [66] creates a homogeneous magnetic field of 1.5 T flux strength inside the barrel. This magnetic field directs charged particles along curved trajectories from which their momentum can be reconstructed. It encloses all detector subsystems except for the KLM (cf. §7) within its cylindrical volume of $3.4 \text{ m} \times 4.4 \text{ m}$ (diameter \times length).

§7. The K_L/μ Detector (KLM) [86, 87] is the outermost active part of the Belle detector and serves two purposes: 1. its metal structure serves as return yoke for the magnetic flux of the enclosed solenoid, and 2. it is a detection subsystem dedicated to measuring long-lived, deeply penetrating particles, namely the K_L and the muon.⁷

This subsystem consists of a central barrel component, which also contains flux return plates, and two end cap sections.

The working principle is similar to that of a sampling calorimeter: layers of an absorber material (4.7 cm of iron) are interleaved with a charged particle detection system (resistive plate counters, RPC). An incident particle showers in the absorber and the newly produced charged particles can be detected in the detection layer. The barrel (endcap) section consists of 14 iron layers, each with a thickness of 4.7 cm, and 15 (14) detection layers. Each detection layer, called RPC superlayer, consists of two RPC modules.

As a whole, the iron plates provide a thickness of 3.9 interaction lengths (X_0), the angular coverage is $20^\circ < \theta < 155^\circ$ in the polar angle. Unlike the ECL, the KLM does not provide a measurement of the deposited energy, hence it only allows to measure a particle's direction of movement but not its momentum. The multiplicity weighted spatial resolution of this system is 1.2 cm which gives an angular resolution of less than 10 mrad from the IP [66].

⁷ The interaction length of both muons and K_L mesons is significantly larger than that of the ordinary hadrons and leptons that are to be expected during operation.

The KLM provides separation of K_L and μ signals for momenta greater than 600 MeV. Its sandwich structure allows muon/charged hadron separation based on the signal's cluster shape (amount of transverse scattering) and the particle charge (distance before stopping in the material). Furthermore, if a charged trajectory can be matched to a KLM cluster, the hit is assumed to stem from a muon.

3 Data Samples

The following section discusses the data samples that are used by this analysis. Generally, the data are grouped in experiments — each corresponding to a data taking period — of vastly varying size. Both the MONTE-CARLO (MC) and experimental data used by this analysis come from the experiments 31–51, 55–65.⁸ The analysis does not use data samples taken earlier than experiment 31 because those data were taken before installation of the SVD upgrade (called *SVD II* [76]) which greatly increased the vertexing performance.⁹ Because a high-precision vertex reconstruction is crucial to this analysis, the loss of statistics should be countered by the better resolution. All data samples have been preprocessed using the most recent version (b20090127_0910) of the Belle reconstruction software.

3.1 MONTE-CARLO Data

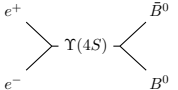
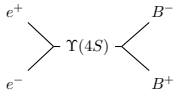
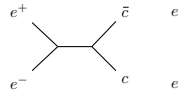
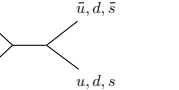
The analysis is developed and evaluated using official Belle generic MC data samples. Data generation is essentially a three-step process

1. The bare e^+e^- collision and the resulting physical decays are simulated using the *EvtGen* package (also called *CLEO QQ generator*) [88]. This program models individual decay types using *decay tables* which encode the possible decays, their respective branching fractions, and the decay specific model (e.g. whether to include photon radiation etc.).
2. Using the *GEANT 3* [89, 90] package, the interactions of the particles produced in step 1. with the detector material are simulated. Afterwards additional background sources not related to the actual particle collision, such as beam bremsstrahlung, are introduced into the data set. These beam induced backgrounds are not simulated but rather taken from experimental data [66]. This step records the detector response of the Belle apparatus in the same way it would be recorded during normal operation with experimental data.
3. Finally the reconstruction and filtering algorithms (the preprocessing stage) are applied to the data. Hence, most inefficiencies and inaccuracies resulting from these algorithms should be found in the MC data to the same extent as in experimental data. All remaining differences are absorbed as systematic uncertainties.

In Section 1.1 the basic composition of the preprocessed data was already discussed; to recapitulate, the data are composed nearly exclusively of $q\bar{q}$, $q \in \{u, d, s, c, b\}$ events.

The MC samples are accordingly categorised in four main contributions, Mixed, Charged, Charm, and *uds*. Table 3 lists the sample types along with a graphical representation of the corresponding physical process and the relative contribution f to the total data sample.

Obviously the signal is a subcategory of the Mixed sample type, it will be denoted by the label *Signal*. The non-signal component of the Mixed type data will be denoted by the label *Mixed* or $B^0\bar{B}^0 \rightarrow J/\Psi K^0 + X$.

Samples			
resonant		continuum	
Mixed	Charged	Charm	uds
			
12.2 %	12.2 %	29.0 %	46.7 %

Tab. 3: List of data sample types used in this analysis. Depicted are illustrations of the physical process, the flow of time in these diagrams is top-to-bottom. Also listed are the approximate relative contributions f of the different types to a measured data set. The final particles for each type may decay generically except for the b quark of the bottom mesons which always decays to a c quark (charge conjugates implied).

⁸ Even experiment numbers are not used.

⁹ This choice reduces the amount of data available to the analysis from the values quoted in Section 1.2 to $\mathcal{L}_{\text{int}} \approx 560$ fb corresponding to roughly $619 \cdot 10^6$ $B\bar{B}$ pairs.

3.2 Experimental Data

The experimental data have been pre-filtered (*skimmed*) using the `psi_skim` which targets events involving $B \rightarrow J/\Psi X$ transitions. In Belle a *skim* is a data subset which has a reduced background contribution without loss of signal events. By using a skim the CPU time required to run the full analysis can be reduced without loss of generality because the skimming should never remove events which pass the full analysis.

Among other things, the `psi_skim` imposes that the event contains two oppositely charged signal-like leptons (electrons, muons). That is to say, their four-momentum sum must correspond to a J/Ψ -like candidate, i.e. its invariant mass must be compatible with that of the J/Ψ . This is the name-giving skim criterion since it enriches events with a J/Ψ . Some more filtering criteria are: at least three *good*¹⁰ charged tracks, a total visible energy which exceeds 20 % of the centre-of-mass energy, and a continuum suppression cut (cf. Section 8 in Chapter V).

¹⁰ That is to say with $p_T > 0.1$ GeV and radial respectively lateral distance to Interaction Point (IP) of $dr < 2.0$ cm and $dz < 4.0$ cm.

V

The Event Reconstruction

This chapter documents the bulk of the analysis: the overall process and the methods involved in the reconstruction of the observables required for performing a measurement of the \mathcal{T} -violation.

In the first section, the concrete implementation of the method that introduced in Chapter III will be outlined briefly. The specifics of the object reconstruction and the problems encountered during its implementation are then discussed in Sections 2 to 6. Section 7 discusses the final reconstruction step, the data set refinement. Finally, Section 8 briefly discusses continuum suppression, which is used by the data refinement.

1 Recipe

Following the abstract methodology presented in Chapter III, the concrete reconstruction recipe can be formulated as follows.

1. Reconstruct candidates for the J/Ψ , K_S , and K_L particles.
2. For each pairing $J/\Psi K_S$ and $J/\Psi K_L$ reconstruct a B_{CP} candidate. Filter this collection of B_{CP} candidates to select one final object. By doing so, the CP tag ξ_{CP} of the event is determined. Furthermore, reconstruct the decay vertex \mathcal{V}_{CP} of the B_{CP} .
3. Use the remaining particle objects of the event to infer the flavour tag ξ_{Flv} of the B_{Flv} . Then reconstruct the decay vertex \mathcal{V}_{Flv} of the B_{Flv} .
4. Reconstruct the signed proper time difference $\Delta\tau$ from the decay vertices \mathcal{V}_{CP} and \mathcal{V}_{Flv} using Equation (9).
5. Refine the data set, that is to say reduce the amount of background events to a reasonably low contribution.

This analysis is implemented using the Belle Analysis Software Framework (BASF) [91–93] and the ROOT framework [94–96].¹ Some design decisions are based on specifics of the BASF or the Belle experiment, refer to Chapter IV for a discussion of the Belle experiment, its detector, and the data samples used in this analysis.

The data processed by the reconstruction algorithm can be accessed from so called *data tables*, each table corresponds to a different type (e.g. charged particle, photon candidate, etc.) of reconstructed object.

2 Reconstructing the J/Ψ

Since the J/Ψ is a neutral particle it cannot be measured directly: fundamentally the detector systems can only measure the energy deposition of charged particles.² Instead, it must be reconstructed from its daughter particles, hereby restricting the possible decay modes to those which result in charged particles.

¹ Using the versions b20090127_0910 (final version) and 5.34.18 respectively.

² In practice, neutral particles can be measured if they interact on a microscopic scale with the detector material hereby producing charged particles.

In order to simplify the reconstruction procedure and increase purity, the analysis is restricted to use leptonically decaying J/Ψ mesons.³ The required decay signature is given by

$$J/\Psi \rightarrow l^+l^- \quad \text{for } l \in [e, \mu].$$

The preprocessed data samples provide a collection of particles reconstructed from charged tracks, the `Mdst_charged` data table. This collection is used as input to the routines reconstructing electrons and muons. The reconstruction of positrons and anti-muons is completely analogous to that of electrons and muons; the determining difference lies in the sign of the reconstructed charge.

Reconstruction of charged particles is performed by [97, 98]

- 1st reconstruction of charged tracks in the CDC by matching individual hits to form a trajectory,
- 2nd matching of CDC tracks to SVD hits, and
- 3rd extrapolation of the tracks into the outer detectors to their respective stopping point or their escape from the detector.

2.1 Electron Reconstruction

An electron appears as a charged track in the inner subdetectors. Due to the external magnetic field the electron track is curled with its curvature depending on the electron's momentum, its track may even be helical.

As a first stage filtering it is required that the fitted track originates close to the IP: the separation along z and r may not exceed 4 cm respectively 2 cm.

Afterwards the electron hypothesis is evaluated using the `eid` tagging routine [82, 97]. This tagging algorithm builds a normalised electron likelihood L_{eid} using various input variables created from the CDC, ACC, and ECL subdetector signals.⁴ Figure 13 shows the distribution of L_{eid} for an input sample of electrons and charged pions. Using this algorithm an electron efficiency of $(92.4 \pm 0.4) \%$ at a π^\pm fake rate of $(0.25 \pm 0.02) \%$ is achieved [97].

2.2 Muon Reconstruction

Muons deposit significantly less energy in the inner detectors when compared to electrons. Due to the low interaction probability of muons, the ECL can be used as high-efficiency muon identifier. This is done by selecting tracks which pass through the complete ECL and deposit an amount of energy which is roughly equal to the minimum-ionising deposit [98]. Furthermore, any charged tracks reconstructed in the KLM are likely of muonic origin.

Muonic tracks are required to fulfil the same requirements on the proximity to the IP as electron tracks, cf. Section 2.1.

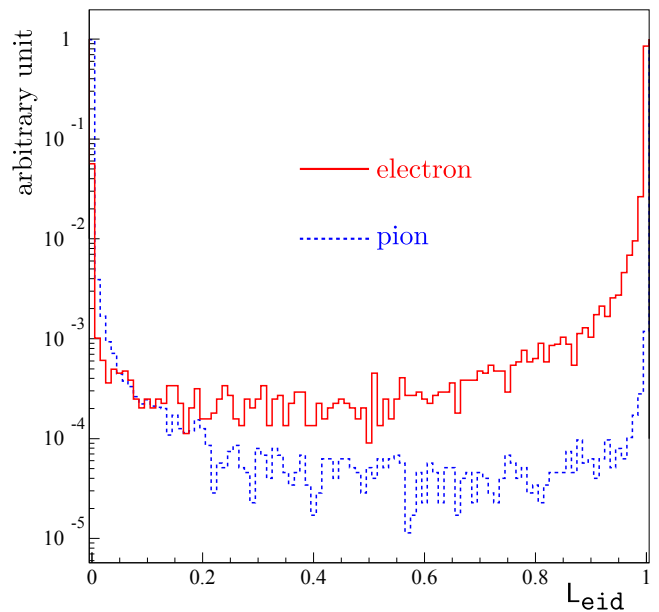


Fig. 13: Typical distribution of the L_{eid} likelihood ratio [97].

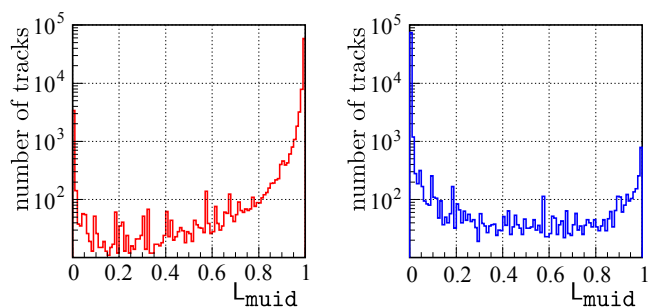


Fig. 14: Distribution of the L_{muid} likelihood for muons and pions, shown in red and blue respectively [82].

³ Leptons can be detected and reconstructed with much better accuracy than hadrons.

⁴ Such as the reduced chisquare goodness-of-fit classifier χ_{red}^2 of the electron track fit, the energy-to-momentum ratio E/p , the ECL shower shape, the observed dE/dx , and the number of measured photo-electrons per refractive index of the ACC [82].

The muon hypothesis is evaluated using the `muid` tagging routine [82, 87, 98]. First level muon identification comes from a track-to-KLM-cluster matching routine. For each remaining candidate track a normalised muon likelihood L_{muid} , cf. Figure 14, is calculated using the track’s range and transverse scattering in the KLM as inputs. Using this algorithm an averaged muon selection efficiency of roughly 90% at a pion (kaon) fake rate of about 1.4% (1.7%) is achieved for momenta between 1.0 and 3.0 GeV [98].

2.3 Forming and Selecting Dilepton Candidates

For each pair of oppositely charged electrons (and muons respectively) a J/Ψ candidate is formed by combining the lepton four-momenta. At the same time a vertex fit, using the `kvertexfitter` package of the `BASF`, is performed using the lepton tracks and the detector geometry (specifically the magnetic field) as input. Using the newly reconstructed vertex, the J/Ψ four-momentum origin (i.e. its decay vertex $\mathcal{V}_{J/\Psi}$) is recalculated. Only candidates with a successful vertex fit are considered for further selection.

Afterwards kinematic cuts on the mass and the absolute value of the three momentum of the dilepton candidate are used to filter the collection. The candidate’s mass, for example, must match the nominal mass $M_{J/\Psi}$ within a few hundred MeV, refer to Appendix D for a complete list of cuts.

Figure 15 shows the invariant mass distribution of the reconstructed J/Ψ candidates (i.e. the dilepton-mass). One can clearly identify the mass peak with its maximum around the nominal J/Ψ mass and an asymmetric tail towards lower values. This is a result of unaccounted energy due to measurement inefficiencies, photon radiation, and other effects. The sudden increase in the distribution at $M_{J/\Psi} \approx 3.036$ GeV is a result of the fact that the lower invariant mass cut for J/Ψ candidates of muonic origin is larger than that for candidates reconstructed from electrons, cf. Table 14 in Appendix D. Following this reconstruction scheme, a fake rate (i.e. the contribution of candidates which are non- J/Ψ) of 17.05% is achieved. Note that the distribution and the resulting numbers have not been corrected for MC deficiencies and/or acceptance.

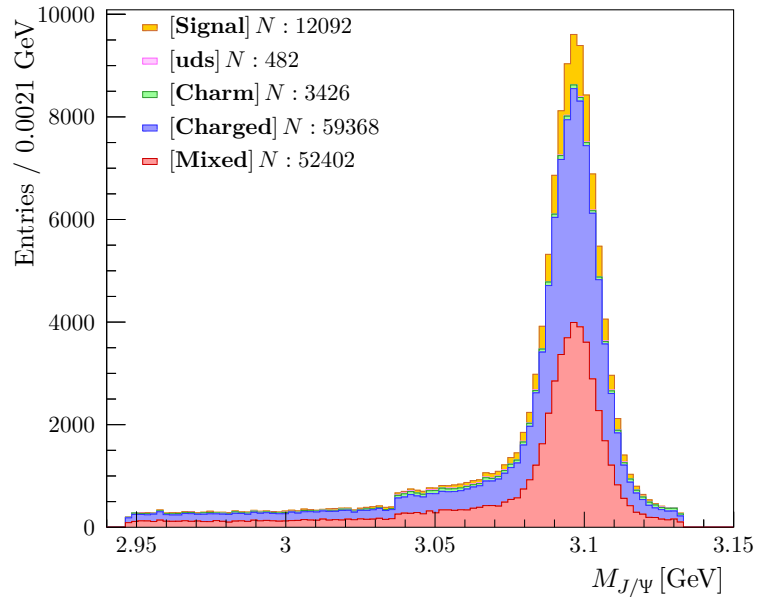


Fig. 15: Invariant mass distribution of the reconstructed J/Ψ candidates shown for the different MC samples.

3 Reconstructing the K_S

The K_S candidates in this analysis are reconstructed from the two-pion decay mode. However, in order to simplify the procedure and increase both purity and efficiency, only the decay to charged pions is considered. Candidates are taken from the `Mdst_vee2` table which contains particle candidates that are reconstructed as so called V^0 particles.

3.1 The V^0 Particle

A V^0 particle is a neutral pseudo-particle that is reconstructed from two oppositely charged tracks which form a shape similar to the letter V . This shape would be expected from a two-prong decay of a signal K_S to two pions since the two daughters are expected to share the K_S momentum evenly and their respective momentum is expected to be large. Other particles which exhibit a similar event shape are the Λ and $\bar{\Lambda}$ baryons, which decay to a $p\pi^-$ ($\bar{p}\pi^+$ respectively) pair, and converted photons, i.e. $\gamma \rightarrow e^+e^-$.

The complete set of V^0 candidates is obtained by combining all possible pairs of oppositely charged tracks in the event. Afterwards a vertex fit using the two input tracks is performed to determine the V^0 decay vertex. Different V^0 hypotheses can then be evaluated by combining the daughter particles using their respective mass hypothesis. That is to say the K_S hypothesis is evaluated by **1.** reconstructing the four-momentum of the positive (negative) track by imposing that its mass should be equal to that of the positive (negative) pion, **2.** calculating the K_S four-momentum by combining the partial four-momenta, and **3.** requiring that the resulting invariant mass be a good match for a K_S . The χ_{red}^2 goodness-of-fit classifier of the reconstructed vertex and its position relative to the IP can be used for additional filtering.

3.2 Identifying the K_S

A first level identification scheme for the K_S was already discussed in the previous section: the mass hypothesis selection of the V^0 . After this preliminary selection stage the K_S hypothesis is further evaluated using the particle ID package `FindKs` [82, 99]. This package applies cuts on geometric variables such as the flight path length of the K_S candidate and the azimuthal angle between the reconstructed K_S momentum vector and the fitted decay vertex vector. The cuts have been optimised w.r.t. the ratio $S/\sqrt{S+B}$ — where S (B) is the number of signal (background) events passing the selection — using MC data. This optimisation provides an overall signal-to-noise ratio S/N of 23.5 (13.7) for MC (data) at an efficiency of 75.6 % (74.0 %) [99].⁵

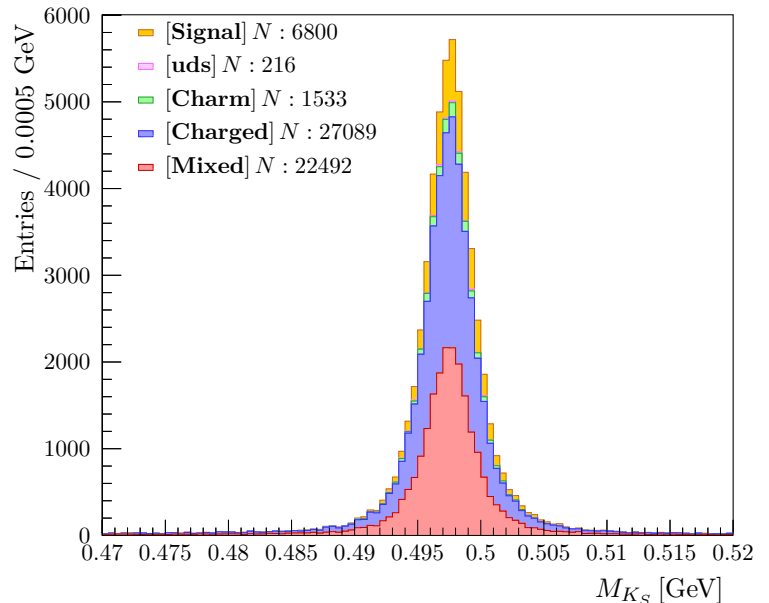


Fig. 16: Invariant mass distribution of the reconstructed K_S candidates shown for the different MC samples.

The reconstructed invariant mass distribution of the K_S candidates is shown in Figure 16. This distribution shows a clear peak centred on the nominal K_S mass and has, as is expected, symmetric tails.⁶ The fake rate (i.e. the contribution of candidates which are non- K_S) of this reconstruction scheme is at an excellent 3.8%. Note that the distribution and the resulting numbers have not been corrected for MC deficiencies and/or acceptance.

4 Reconstructing the K_L

The K_L meson is a very long lived particle which makes it strenuous to detect it properly. Using the minimum truth-level momentum of the K_S and K_L candidates and their respective mass and mean lifetime, one can calculate their mean expected survival distance to be 6.5 cm and 37.56 m respectively. The latter figure is well outside the farthest detector component, whereas the former resides safely within the SVD (cf. Figure 12 and Section 2 in Chapter IV in general).

Hence, unlike the J/Ψ and K_S , the K_L cannot be reconstructed explicitly from its daughter particles; detection must rely on calorimetry, i.e. detection by destruction. The KLM subdetector is built such it poses a barrier which is thick enough that most K_L candidates will come to rest within its corpus. Naturally, this approach limits the possible accuracy of the reconstruction:

- Without daughter particle tracks, the flight-path (and thus vertex) resolution suffers greatly.
- The volume of the detection cells of ECL and (more importantly) KLM are significantly larger

⁵ For data these numbers have been estimated from fits to the distributions of kinematic variables.

⁶ No significant photon radiation by the pions is to be expected.

than those of the inner detector components.⁷ Comparing a residual of 1.2 cm (KLM) to pixel pitches of at most 75 μm (SVD) makes the loss of accuracy rather obvious.

- Due to the method that is used for detection, the spatial resolution degrades even further: by forcing the incident particle to shower in the absorber material, its original trajectory is disturbed and the shower size is significantly larger than the ionisation cloud of a particle traversing a silicon pixel sensor.
- Because there is no track based momentum reconstruction, the only chance to obtain an energy estimate for the incident particle comes from the deposited energy. However, the energy resolution of the ECL detector cannot compete with a track based reconstruction. Moreover, the KLM does not provide an energy estimate.

Fundamentally however, reconstruction of a K_L candidate is simple:

- 1st combine KLM hits to form clusters,
- 2nd reject clusters which can be matched to a charged track extrapolated from the CDC within 15°, and
- 3rd for all remaining clusters the track is determined as a straight line pointing from the IP to the centre-of-gravity of the respective cluster.

Optional information from the ECL can be added if it can be matched and there is no conflicting track. With the `Mdst_klong` collection the `BASF` provides a collection of candidates that have been reconstructed in this manner.

Additional filtering is then performed by applying cuts on the number of hit layers in the KLM, the angular separation to the next charged track, the angular separation between the KLM and ECL tracks (if available), and the ECL signal (also if available); the complete list of cuts can be found in Appendix D. Following this scheme a fake rate (non- K_L particles in the K_L sample) of 46 % is obtained.

For further study and selection it is beneficial to obtain an estimate of the K_L momentum: the following section discusses the method used by this analysis. Note that the approach requires a reconstructed J/Ψ candidate and expects the event to be of signal-type.

4.1 Estimating the Four-Momentum

The only properly measured kinematic property of the K_L is the direction of its three-momentum $\hat{\vec{p}}_{K_L}$ in the laboratory frame.⁸ In order to obtain an estimate of the K_L four-momentum, two K_L quantities must be inferred: its energy in the lab frame E_{K_L} and the absolute value of its three-momentum in the lab frame p_{K_L} .

The estimate of the K_L four-momentum assumes that

- 1 The candidate is truly a K_L , hence its invariant mass $\sqrt{P_{K_L}^2}$ must be equal to the reference value of the K_L mass M_{K_L} .
- 2 There are no additional daughter particles of the B_{CP} , hence

$$P_{B_{CP}}^2 \stackrel{!}{=} \left(P_{J/\Psi} + K_L \right)^2 \stackrel{!}{=} M_{B^0}^2.$$

- 3 The beam energy in the centre-of-mass frame is split evenly between the two B mesons, therefore

$$\frac{1}{2} E_{\text{Beam}}^{[\text{c.m.}]} \stackrel{!}{=} E_{B_{CP}}^{[\text{c.m.}]} \quad \text{and thus with } \mathbf{2} \quad \frac{1}{2} E_{\text{Beam}}^{[\text{c.m.}]} \stackrel{!}{=} E_{K_L}^{[\text{c.m.}]} + E_{J/\Psi}^{[\text{c.m.}]}$$

By rearranging Equation (10) the absolute three-momentum in the lab frame p_{K_L} can be expressed in terms of the nominal mass and lab frame energy (following assumption 1); the unphysical negative solution is rejected.

$$M_{K_L}^2 = E_{K_L}^2 - \vec{p}_{K_L}^2 = E_{K_L}^2 - p_{K_L}^2 \tag{10}$$

$$\Rightarrow p_{K_L} = +\sqrt{E_{K_L}^2 - M_{K_L}^2} \tag{11}$$

⁷ Because they are further out — and thus occupy a much larger volume — instrumenting them with the same granularity is simply too expensive.

⁸ Henceforth known as *lab frame*.

The generic boost formula given by Equation (12), where $\vec{\beta}$ is the boost-vector and γ the LORENTZ factor, can be rewritten using $\vec{p}_{K_L} := p_{K_L} \cdot \hat{p}_{K_L}$ and Equation (11) to yield Equation (13).

$$E_{K_L}^{[c.m.]} = \gamma \left(E_{K_L} + \vec{\beta} \cdot \vec{p}_{K_L} \right) \quad (12)$$

$$= \gamma E_{K_L} + \gamma \vec{\beta} \cdot \hat{p}_{K_L} \sqrt{E_{K_L}^2 - M_{K_L}^2} \quad (13)$$

$$= \frac{1}{2} \left(E_{\text{Beam}}^{[c.m.]} - E_{J/\Psi}^{[c.m.]} \right) \quad (14)$$

Using assumption 3 the centre-of-mass energy of the K_L must be given by Equation (14).

The combination of Equations (13) and (14) is fully determined using the measured K_L and J/Ψ data as well as the known centre-of-mass energy and boost. Solving the resulting equation for E_{K_L} yields two solutions, however, MC tests revealed that the negative solution provides a more accurate estimate.

Figure 17 shows the distribution of the difference between the truth-level absolute three-momentum $p_{K_L, \text{Truth}}$ and its estimated lab frame counterpart p_{K_L} for the different MC samples. Further examination shows that for 82.8% of the signal events — which are treated using the above method — the resulting value for p_{K_L} lies within 200 MeV of the true value. This number should be compared to the mean and RMS of the truth-level p_{K_L} distribution which are 1.86 GeV and 360 MeV respectively.

It can be found that the estimate typically overestimates p_{K_L} whenever the absolute difference

$$|p_{K_L, \text{Truth}} - p_{K_L, \text{Reco}}|$$

is larger than 200 MeV. This can be attributed to the assumptions 2 and 3 of the above listing: the overestimate is a result of some unaccounted energy and/or momentum component in the above equations. Possible sources are photon radiation or a nonzero p_T of the B_{CP} .

Clearly the distributions for the non-signal contributions are not zero-centred and tend to have large, negative values. However, this is not surprising because the assumptions that are used for the calculation cannot be correct for the non-signal samples.

5 Reconstructing the B_{CP}

The steps taken to create the CP -even and -odd B_{CP} candidates are outlined in Sections 5.1 and 5.2 respectively. Following that, the procedure for selecting a final B_{CP} candidate and thus obtaining the event's CP tag is discussed in Section 5.3.

5.1 Obtaining CP -even Candidates

A list of B_+ candidates is created by forming the pairwise combinations of all J/Ψ and K_L candidates. For each candidate the expected K_L direction \vec{e}_{K_L} is calculated from the reconstructed J/Ψ four-momentum $P_{J/\Psi}$; the calculation follows a simple procedure:

- 1st The J/Ψ four-momentum $P_{J/\Psi}^{[c.m.]}$ is calculated by boosting $P_{J/\Psi}$ using the well-known centre-of-mass boost vector $\vec{\beta}$, then
- 2nd the sign of the spatial components of the boosted vector are reversed, and finally

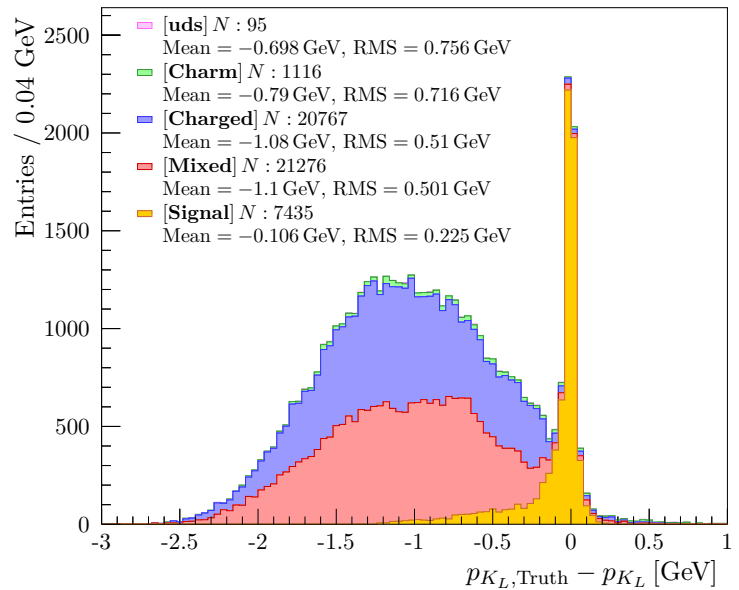


Fig. 17: Difference between estimated and truth-level absolute value of three-momentum of the K_L in the lab frame for the different MC samples.

3rd the resulting four-vector is boosted back to the lab frame; \vec{e}_{K_L} is given as the unit-length three-momentum of the four-vector obtained by the previous steps.

For the expected K_L direction to be correct, one must assume that there is in fact a signal decay $B \rightarrow J/\Psi K_L$ — which is obviously a two-body decay — and that the B carries half of the momentum of the $\Upsilon(4S)$.

Following the calculation, the angular separation $\Delta\varphi_{K_L}$ between \vec{e}_{K_L} and the reconstructed K_L three-momentum direction \hat{p}_{K_L} is used as a cut variable. Additional cuts are applied on the absolute values of the J/Ψ and $B_{\mathcal{CP}}$ three-momenta, and the ΔM_{BC} variable. Refer to Appendix D for a complete list of cuts and definitions of the cut variables.

While in the development stage various additional kinematic variables have been trialled for filtering of the $B_{\mathcal{CP}}$ candidates. However, detailed study of the performance of the cut based selection has shown that the resulting signal purity and selection efficiency is rather low. It proved more efficient to use a *multivariate* selection method. Hence, the cuts applied at this stage of the selection procedure are relatively loose in order to obtain a larger data sample which is then filtered more thoroughly and efficiently using the multivariate selection. The method that was used and its performance are discussed in Section 7.

5.2 Obtaining \mathcal{CP} -odd Candidates

For each pair of J/Ψ and K_S a B_- candidate is created by calculating the four-vector sum. This candidate list is trimmed by applying cuts on the absolute values of the three-momenta of the J/Ψ and the resulting B_- . Refer to Appendix D for a listing of the cut values.

During development of the analysis several additional selection criteria based on kinematic variables — such as the B_- invariant mass, the ΔE and M_{BC} variables, and others — have been implemented and their performance was evaluated. However, using those variables at this stage of the analysis procedure and within a classical cut-based selection scheme was discontinued in favour of a *multivariate* selection method; this method is discussed in Section 7.

5.3 Selecting the Final Candidate

If there is one single $B_{\mathcal{CP}}$ candidate that has matched the B_+ or B_- quality cuts, it is accepted without further selection. In the case of multiple $B_{\mathcal{CP}}$ candidates, the event is vetoed if the candidates do not share the same \mathcal{CP} eigenvalue $\xi_{\mathcal{CP}}$. Otherwise, the candidates are filtered, yielding a single $B_{\mathcal{CP}}$ candidate; the filtering method depends on the candidates' \mathcal{CP} eigenvalue $\xi_{\mathcal{CP}}$:

$\xi_{\mathcal{CP}} = +$	Due to the poor reconstruction of the K_L , a cut on the invariant mass of the $B_{\mathcal{CP}}$ is impractical. Hence, the quality of the candidate is evaluated using the $\Delta\varphi_{K_L}$ variable, cf. Appendix D, of its K_L daughter. The $B_{\mathcal{CP}}$ candidate which corresponds to minimal $\Delta\varphi_{K_L}$, i.e. whose K_L daughter has the minimal angular deviation from its expected trajectory, is selected.
$\xi_{\mathcal{CP}} = -$	The candidate whose reconstructed mass has the smallest deviation from the nominal B^0 meson mass is selected.

Implementation of the selection as an event veto in the case of heterogeneous \mathcal{CP} tags surely is not the optimal approach. However, the studies performed as part of this analysis have not produced a viable approach to be used for reducing the ambiguity of a differing \mathcal{CP} tag.

The relative data reductions resulting from applying the dual- \mathcal{CP} tag veto are listed in Table 4.

Data Type	Veto Result
Signal	Reduction by 9.5 %
Charged	Reduction by 3.8 %
Charm	Reduction by 4.6 %
Mixed	Reduction by 3.8 %
uds	Reduction by 4.4 %

Tab. 4: Result of applying dual- \mathcal{CP} tag veto on the different data types.

Obviously the reduction by the veto is most severe for the signal sample. This effect can be attributed to the fact that a signal event should actually behave like the expected type of event whereas a background event should not. Hence, it is more likely for a signal event to be interpreted in a way that provides two candidates which match the expectation — and thus trigger the veto — than it is for a background-type event.

A dedicated study or a subsequent analysis should further investigate the selection method and specifically the dual- CP tag veto. By analysing the MC truth record the above hypothesis can be verified or replaced by a better explanation. Doing so surely is worthwhile even if this is done only in order to better understand the sources of the wrong-sign CP tag.

5.4 The \mathcal{V}_{CP} Vertex

In this analysis, the decay vertex \mathcal{V}_{CP} of the B_{CP} is set to the vertex $\mathcal{V}_{J/\Psi}$ reconstructed for the J/Ψ . This decision was based on the following aspects:

1. Given the largest truth-level momentum observed for any J/Ψ candidate of this analysis (which is roughly 3.6 GeV) and the known J/Ψ mass and lifetime [6], the expected distance traversed by a J/Ψ is about $2.44 \cdot 10^{-6} \mu\text{m}$. This distance is well below any realistically achievable vertex resolution, hence the truth-level displacement of $\mathcal{V}_{J/\Psi}$ and \mathcal{V}_{CP} is negligible.
2. Inclusion of the K_S in the vertex fit does not yield a significant increase in accuracy. Inclusion of the K_L in the fit is not sensible altogether because its kinematic properties can only be reconstructed poorly: the track resolution for the K_L is at the order of a centimetre.
3. By excluding the K_S the procedure is more consistent: the accuracy of the J/Ψ vertex should not depend on the CP eigenvalue of its K^0 sibling. By specifically excluding the $K_{S,L}$ it is guaranteed that the performance of the B_{CP} vertex reconstruction does not depend on its CP eigenvalue either.

Figure 18 shows the residual distribution of the B_{CP} vertex. Only the signal contribution of the MC data was used for this plot because z_{CP}^{tru} is not defined for the background. The distribution shows no apparent asymmetry and is sufficiently well zero-centred.

The mean value of the z_{CP}^{tru} distribution is roughly -1.5 mm at a root mean square (RMS) of 3.5 mm . Comparing the RMS of the residual distribution — which is an estimate for the error on the \mathcal{V}_{CP} vertex measurement — to those numbers shows that the expected uncertainty is roughly 1.4% of the RMS of the quantity to be estimated. Obviously the CP -vertex can be reconstructed with acceptable accuracy.

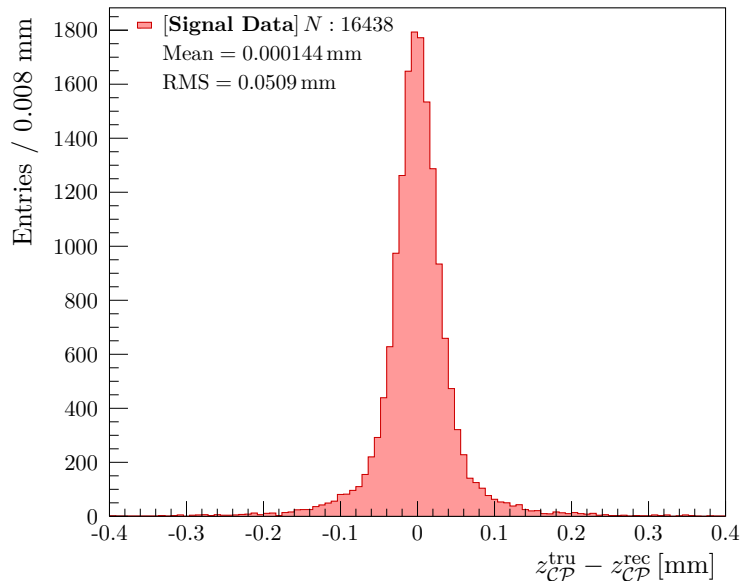


Fig. 18: Distribution of residual (truth value minus reconstructed value) of the z -coordinate of the B_{CP} vertex.

6 Reconstructing the B_{Flv}

The B_{Flv} reconstruction consists of two steps: 1. estimation of the event's flavour tag — discussed in Section 6.1 — and 2. reconstruction of the B_{CP} decay vertex — presented in Section 6.2.

6.1 Flavour Tagging

The flavour tag is performed using the `Hamlet` [100, 101] package of the `BASF`.

6.1.1 Tagging Procedure

Using the pseudo-particle obtained from the B_{CP} reconstruction, a CP arm of the event is defined; this is marked in green in Figure 19. The `Hamlet` routines use the other arm of the event (i.e. the whole event except for the B_{CP} particle and its daughters, marked in yellow in the illustration) and infer a flavour tag. By construction, this flavour tag should coincide with the flavour of the B_{Flv} meson. Naturally the reconstruction level flavour-arm can only contain visible particles, i.e. neutrinos are only considered as missing four-momentum; neutral particles may contribute if they decay to charged particles within the detector acceptance or are measured by the calorimeter.

The basic flavour tagging principle was already discussed in Section 2.3.2 in Chapter III. In this analysis, the `Hamlet` package is configured to use a multi-dimensional likelihood approach for flavour tag determination. The flavour tagging strategy used internally was designed to maximise the effective tagging efficiency e_{eff} (defined by Equation (15) further below).

Fundamentally, `Hamlet` performs a lookup of a discriminator variable from a multidimensional lookup table, which was calculated from MC data. The lookup requires inputs from various particle ID tagging routines and kinematic variables (all of this is done internally). A list of transitions which are used by `Hamlet` to infer a flavour-tag can be found in [101]. Unfortunately the flavour tagging scheme cannot determine the correct flavour of the B mesons from the measured final state particles in all cases. Among other things, this may be caused by [101]:

- Inefficiencies in the particle detection (detector acceptance, etc.), reconstruction (e.g. software trigger levels), and identification.
- A B meson decaying through a decay chain which does not encode its flavour eigenvalue, e.g. $\bar{D}^0\pi^0 \rightarrow K^0\pi^0\pi^0$.
- Decays through processes producing final state particles which have little information about the initial state flavour.

Furthermore the flavour tagging may fail because the CP -side of the event has been reconstructed incorrectly: by removing flavour-encoding particles from the tagging arm, which then also retains particles from the CP arm, the flavour tag becomes unreliable.

6.1.2 Tagging Performance

An important quantity for evaluating the tagging algorithm's performance is the *wrong tag probability* w_{tag} — that is the fraction of wrongly tagged events — which can be used to calculate the dilution factor \mathcal{D} :

$$\mathcal{D} := 1 - 2 \cdot w_{\text{tag}}.$$

The dilution factor relates the true process amplitude $\mathcal{A}_{\text{True}}$ with the *diluted* (measured) amplitude $\mathcal{A}_{\text{Meas}}$:

$$\mathcal{A}_{\text{Meas}} = \mathcal{D} \cdot \mathcal{A}_{\text{True}}.$$

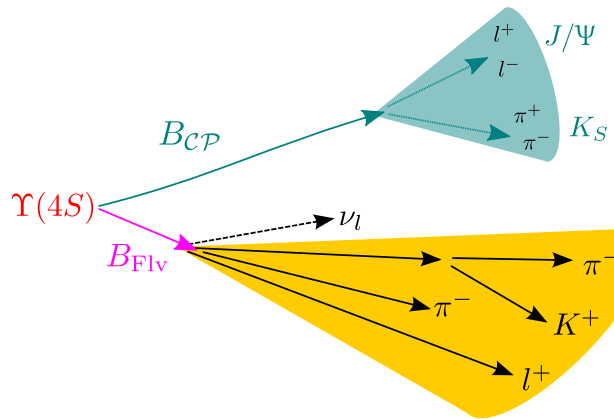


Fig. 19: Illustration of the CP - and flavour-arm of a B_- signal event. The event configuration of a B_+ can be defined analogously.

Precise knowledge of the dilution factor is crucial for a competitive measurement. Moreover, it can be used to weight events according to the credibility of their flavour tag. Another variable quantifying the performance of the tagging is the *tagging efficiency* ϵ_{tag} , which is defined as

$$\epsilon_{\text{tag}} := \frac{\text{Number of events tagged as } X}{\text{Number of events truly being } X \text{ in input}}.$$

For both groups excellent tagging efficiencies of roughly $\epsilon_{\text{tag}} = 99.8\%$ are achieved which is in perfect agreement with the expectation [101]. A combined performance discriminator is the *effective efficiency* e_{eff} given by

$$e_{\text{eff}} := \epsilon_{\text{tag}} \cdot (1 - 2 \cdot w_{\text{tag}})^2 = \epsilon_{\text{tag}} \cdot \mathcal{D}^2. \quad (15)$$

In order to optimise the flavour tagging algorithm one must optimise it for maximum e_{eff} . It has been found that this can be done more efficiently if events that fall in different ranges of w_{tag} are treated differently [101].

Typically, the quantity w_{tag} can only be determined from MC data. However, the flavour tagging algorithm has been designed such that it allows extracting this value on an event-by-event basis from data. A thorough discussion of the flavour tagging algorithm optimisation procedure and its evaluation can be found in [101]. In the following discussion, the variable w_{tag} refers to the *estimated* wrong tag probability, this estimated quantity should be interpreted as a measure of likelihood of the tagging hypothesis provided by the algorithm.

Using the mean wrong tag probability for signal events $\hat{w}_{\text{tag}} = 26.2\%$, the calculated tagging efficiency, and Equation (15), the effective efficiency can be calculated for data to be $e_{\text{eff}} = 23.0\%$. Evaluating the flavour tagging performance on background events does not really make sense: for starters, the B_{CP} cannot be correctly reconstructed, hence in a strict sense the flavour tagging must fail by construction. Furthermore, since the background may not contain a neutral B meson, flavour tagging may be pointless altogether.

In Figure 20 different plots used for evaluation of the flavour tagging performance are presented:

Figure 20a shows the mis-tag probability w_{tag} as obtained from different types of MC data. Note that the plot shows the histograms for different \mathcal{CP} eigenvalues with a different sign; for $\xi_{\mathcal{CP}} = +1$ ($\xi_{\mathcal{CP}} = -1$), the histogram grows upwards (downwards). The distributions for different values of $\xi_{\mathcal{CP}}$ and the different input categories exhibit no apparent asymmetry. It should be noted that very low mis-tag probabilities (i.e. the first bin) are slightly favoured, however, even more pronounced is the region around $w_{\text{tag}} \sim 0.455$, which is the least meaningful tagger result.

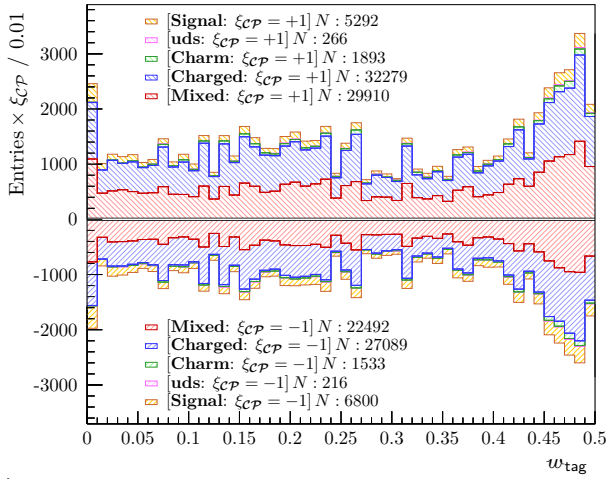
Figure 20b compares the mis-tag probability w_{tag} for *correct-sign* and *wrong-sign* tags. As is to be expected, the former category peaks for $w_{\text{tag}} \sim 0$, whereas the latter peaks close to $w_{\text{tag}} = 0.5$. Note that the data that are used for this plot come from the *Signal* subset. This restriction was employed because there can be no correct-sign tag for non-signal events.

Figure 20c shows the correct-tag fraction; it can be found that the maximal contribution of correctly tagged events occurs close to $w_{\text{tag}} = 0$, peaking for a percentage of 84.8%. The minimal correct-tag fraction of 49.9% is found close to $w_{\text{tag}} = 0.5$. Note that the *Signal* subset is the only data sample used for creation of this plot.

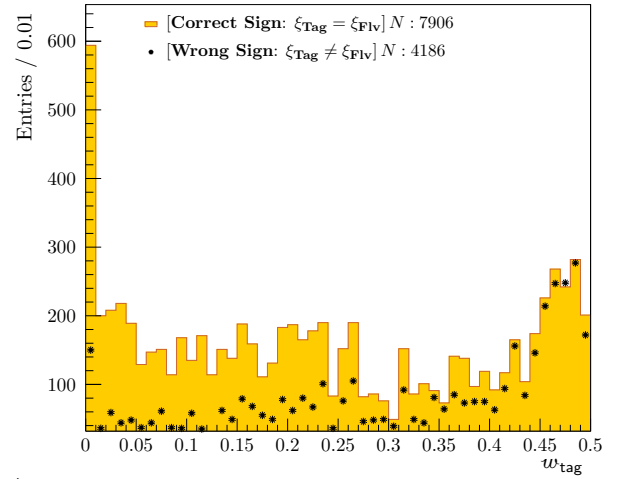
6.2 Estimating the \mathcal{V}_{Flv} Vertex

Since neither the analysis itself nor the flavour tagging algorithm explicitly reconstruct the B_{Flv} , there is no pseudo-particle which can be used to obtain the \mathcal{V}_{Flv} vertex. Because the tagging algorithm obtains its result from a likelihood using numerous inputs instead of a concrete reconstruction scheme, it does not really make sense to try to match the flavour tag by combining the remaining particles in some ordered manner.

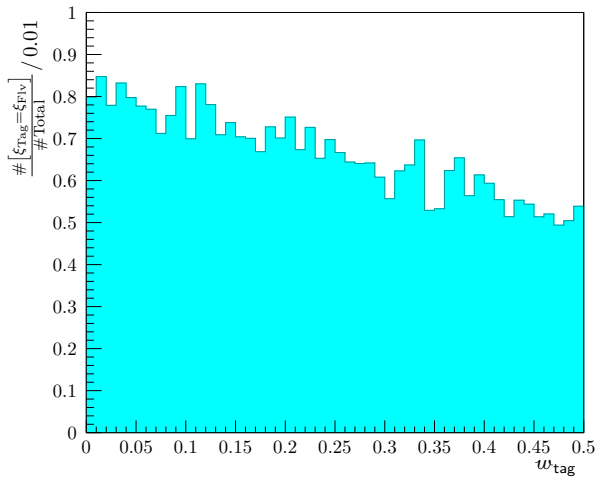
Hence, the \mathcal{V}_{Flv} is estimated by combining the tracks of **all** remaining charged particles — called the *restcharges* — regardless of whether they can be used to create a physical state with the estimated flavour tag.



(a) Mis-tag probability comparison for the different MC sample types.



(b) Mis-tag probability comparison for correct-sign and wrong-sign tags.



(c) Correct-sign fraction as function of mis-tag probability.

Fig. 20: Performance evaluation plots for the flavour tagging.

- Figure (a) shows the mis-tag probability w_{tag} for each input sample separately as a stacked plot. The histogram for the data subset selected by $\xi_{CP} = +1$ ($\xi_{CP} = -1$) grows upwards (downwards).
- Figure (b) compares the evolution of the abundance distribution of the correct- (filled area) and wrong-sign (stars) flavour-tags as a function of the mis-tag probability w_{tag} .
- Figure (c) shows the correct-sign fraction

$$\frac{\#[\xi_{\text{Tag}} = \xi_{\text{Flv}}]}{\#\text{Total}},$$

where $\#[\xi_{\text{Tag}} = \xi_{\text{Flv}}]$ and $\#\text{Total}$ are the number of correctly-tagged respectively total events, as a function of the mis-tag probability w_{tag} .

Note that the data used for creation of Figures (b) and (c) comes exclusively from the *Signal* subcategory.

In order to combine the particles properly they must be identified; this follows a simple scheme which is illustrated by Figure 21.

Basically, the particles are subjected to the particle identification routines `muid` (for muon identification [82, 87, 98]), `eid` (for electron identification [82, 97]), and `atc` (for kaon identification [82]) in that order.

Whenever the output of the respective classifier L_i exceeds the cut value C_i applied at that step, the subsequent routines are skipped and the candidate is selected to be of the type of that classifier. Candidates which are not selected by any of the previous classifiers are assumed to be pions since they are the dominant hadronic contribution at Belle anyways. The cut values C_i are given in Table 18 in Appendix D.

Using the four-momenta and tracks obtained from the restcharge reconstruction, the vertex of the B_{Flv} is fitted. This fit is performed using the `TagVK` package of the `BASF`, which is essentially an extended wrapper around the `kvertexfitter` package already used for fitting the J/Ψ decay vertex. In addition to the restcharge collection, the `TagVK` algorithm utilises the measured interaction point and beam profile as well as the reconstructed B_{CP} vertex in order to obtain an optimal fit result.

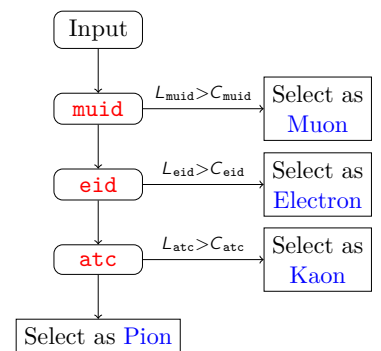


Fig. 21: Illustration of tagging procedure for the restcharge particles.

Figure 22 shows the residual distribution of the z -component of the \mathcal{V}_{FLV} vertex; naturally only signal MC events are used for this plot.

Obviously the distribution is not zero-centred, furthermore it is not symmetric: it has a nonzero, negative skewness. Hence, the distance of the B_{FLV} decay vertex along the z -axis relative to the IP is slightly overestimated by the reconstruction scheme. However, this bias is still relatively small at about $20\ \mu\text{m}$ which is less than one percent of the RMS of the $z_{\text{FLV}}^{\text{tru}}$ distribution.

The RMS of the residual distribution is slightly more than twice that of the residual distribution for the \mathcal{V}_{CP} vertex, however, at about 3.2% of the RMS of the $z_{\text{FLV}}^{\text{tru}}$ distribution, the expected \mathcal{V}_{CP} vertex uncertainty is still sufficiently small.

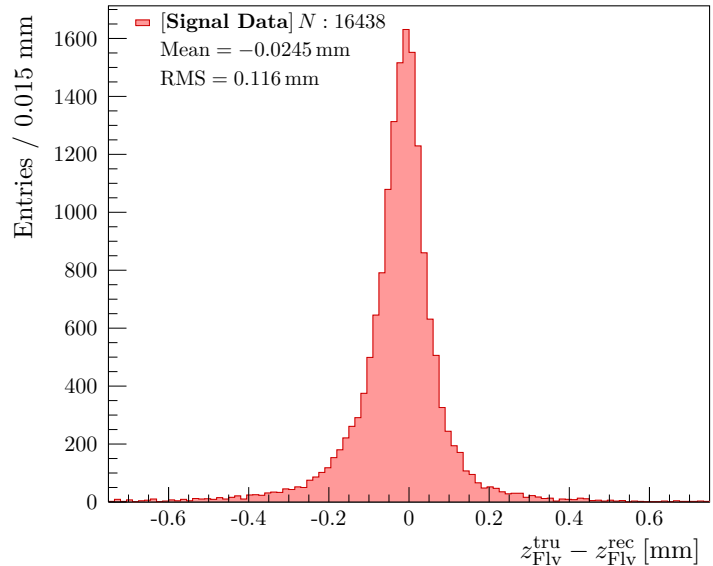


Fig. 22: Distribution of residual (truth value minus reconstructed value) of the z -coordinate of the B_{FLV} vertex.

the expected \mathcal{V}_{CP} vertex uncertainty is still sufficiently small.

7 Refining the Data Set

In the previous sections a reconstruction and selection scheme has been outlined which is the result of an extensive development period. As part of the development process, the cuts applied at every stage have been varied and additional cuts have been trialled in order to determine a routine which maximises signal purity and efficiency. However, while it is not difficult to achieve acceptable levels for the B_- mode, the poorly reconstructed K_L prevents proper selection of a B_+ sample: using a standard cut-based approach a purity of roughly 85% was achieved for the B_- mode, on the other hand for B_+ the maximal value obtained was 34%.

Hence, a *multivariate* selection is used in order to benefit from the full signal-background discrimination power encoded in the event observables. The following two sections are meant as a very brief introduction into *Multivariate Analyses*, a more complete discussion of multivariate analysis techniques, how their performance can be evaluated, and which pitfalls one should avoid can be found in [102]. Section 7.1 outlines the concepts of a multivariate analysis technique, afterwards Section 7.2 briefly introduces *Boosted Decision Trees*, which are the technique used by this analysis. Finally, Section 7.3 discusses the concrete application for this analysis.

7.1 Multivariate Analysis Techniques

A selection approach based on the principles of *univariate analysis* will apply a classical rectangular cut selection, that is a strictly sequential application of cuts on the given observables (variables). The Multivariate Analysis (MVA) based selection method on the other hand applies a selection on the multidimensional variable space of all the observables simultaneously rather than for each observable separately.

Where the rectangular cut approach yields a true-false response (whether all requirements have been fulfilled or not), the MVA approach typically attributes a probability to the true-false hypothesis instead of dealing in absolutes. At the fundamental level, the MVA combines the discriminating power encoded in each individual variable of a given event into a single variable, the **MVA Classifier Output**.

Figure 23 shows an example of the output distributions for signal and background events. Obviously, the background contamination can be reduced by moving the selection cut towards larger y at the cost of reducing the signal efficiency. For a given cut value two error rates can be determined: the *background*

acceptance rate (Type-I Error) and the *signal rejection* rate (Type-II Error).⁹ Application of an MVA approach to a problem involves balancing the trade-off between these two types of errors.

Basically the increased power of an MVA when compared to a classical approach comes from the following factors:

1. The MVA can harness the separation power encoded in the correlations between different variables. If properly setup, an MVA can provide very strong separation from a handful of weakly separating variables. To the MVA not only a variable's value counts but also how it relates to the values of the other variables.
2. By operating on all variables simultaneously the MVA approach is less prone to reject events which appear background-like in some variables as long as the overall appearance is still signal-like.

In order to be able to provide separation power, the MVA must learn to distinguish signal- and background-like events through a process called *supervised machine learning*.¹⁰ Naturally the exact form of the internal mapping function used by the MVA to reduce from the n -dimensional variable space to the scalar classifier output space must be determined. The learning phase can be seen as a process, where the actual mapping function is built by minimising a loss function.

This loss function encodes the amount of signal (background) events which are rejected (accepted) using an input sample of events with known class membership, i.e. it encodes both Type-I and -II errors. Hence, in order to train the MVA, a large sample of pre-classified MC event data is required.

Figure 24 shows a toy example involving a data set of two variables, where each event belongs to either the signal or background class. After the training procedure, the MVA's internal function has produced a classifier boundary which is visualised by the red line.

Typically, four steps are involved when using an MVA based approach for event selection:

1. Using a pre-classified set of *signal* and *background* events, the MVA classifier is trained. The composition of the input data should be equal to the composition of the actual data sample that will be classified later.
2. A second statistically independent pre-classified data set is used to test the performance of the training. Ensuring that sufficient separation power is achieved on an independent set and validating that no *overtraining* has occurred are the two main objectives of this step. Overtraining refers to the feature that an MVA becomes sensitive to the statistical fluctuations within a data sample. If overtraining is present it should manifest in poor separation along the classifier boundary since an independent input sample should not have the same fluctuations. A simple overtraining check is to calculate the compatibility of the MVA classifier output distributions for the training sample and

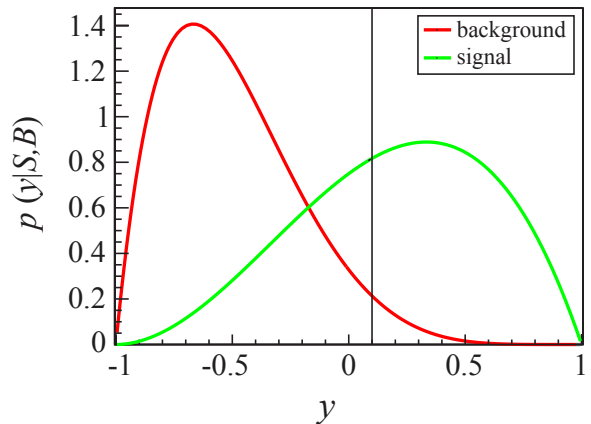


Fig. 23: Example of MVA output distribution for signal and background events; the vertical line denotes the selection cut. Modified illustration taken from [102].

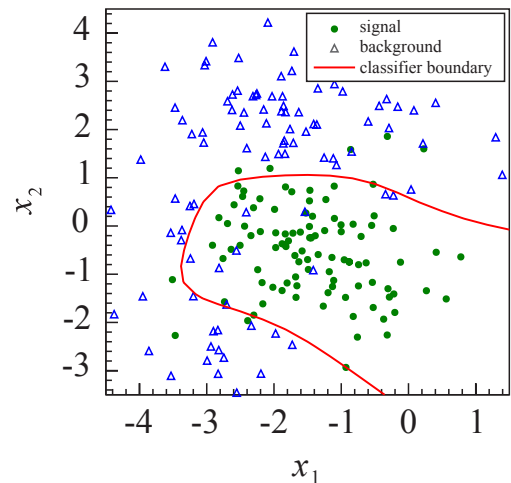


Fig. 24: Example data in two variables for signal and background classes along with the trained classifier boundary. Modified illustration taken from [102].

⁹ Which can be determined from $\int_{y_{\text{Cut}}}^{\infty} p(y,B) dy$ and $\int_{-\infty}^{y_{\text{Cut}}} p(y,S) dy$ respectively.

¹⁰ This process is called *training*; there are other techniques for training an MVA, however, they are not important for this discussion.

a second statistically independent sample. A small compatibility can be caused by overtraining. References [102, 103] include a more complete discussion of the matter.

3. Using a third statistically independent pre-classified data set (one may also use the same data set that was used for the overtraining check) a classifier output cut \mathcal{C}_{MVA} is determined which fulfills the desired requirements on the Type-I and -II errors.
4. Finally the trained classifier is applied to the input sample, i.e. the data sample where the class of each event is not known a priori. After application the sample is filtered using the cut \mathcal{C}_{MVA} that was previously determined: all events that pass this cut are assumed to be signal-like.

7.2 Boosted Decision Trees

A general *decision tree* takes some input \vec{x} and evaluates a decision making function f_{D} at each node. This decision making function determines to which child of the current node the input is passed next. Typically, each node has none or exactly two children (it is a binary decision tree) and the decision making function is unique to the current node.¹¹

Eventually, the input is passed to a node without children (a so called *leaf node*) where it stops propagating. The set of nodes which led to this leaf is called *branch*. After the initial training each leaf gets assigned to a class, depending on the signal-to-background ratio of the set of events that got propagated to that leaf. If the signal-to-background ratio is less than a certain critical value R_{crit} , the node is marked to be background otherwise signal. This ends the training phase of the decision tree classifier; a depiction of the principle is shown in Figure 25.

The performance of a decision tree can be enhanced by a process called *boosting*: instead of growing a single tree, a number of different trees are created sequentially. After creation of a tree is finished all those events that have been mis-classified (e.g. signal events which propagated to a background leaf) receive a higher *weight*¹² and therefore will be more important in the next iteration of tree creation. A method called *pruning* can be used to remove parts of the tree(s) which either have low sensitivity or are sensitive to statistical fluctuations in the data.

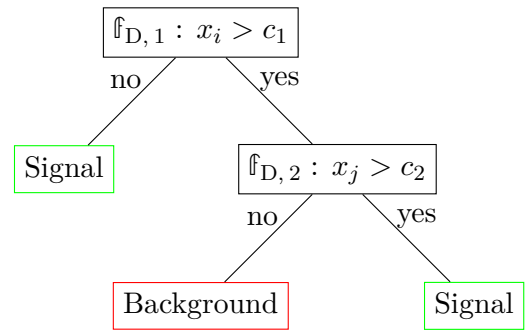


Fig. 25: Illustration of decision tree principle.

The MVA discriminant is calculated from the properties of the propagation of the individual events through the trees grown in the training phase; these properties include the number of times tagged as signal, the total sum of weights, and others. Usually the contributions from different trees are weighted such that the separation power of the classifier is maximised. This reduces the contribution of trees which result in poor separation power.

7.3 Application for this Analysis

The MVA based event selection performed as part of this analysis uses the TMVA software package [103–105].¹³ The goal of using the MVA selection is to maximise the signal purity and efficiency of the collections of B_+ and B_- events that have been reconstructed previously.

In order for the MVA selection to work properly, the events of a given class must be similar, hence, two differently trained MVAs are used for B_+ and B_- events. The B_+ events are further grouped by the available detector information: some K_L candidates include an ECL cluster (cf. Section 4), thus additional observables can be used by the MVA. Table 5 lists the distinct MVA instances and their corresponding subsample type.

MVA Label	Subsample Type
minus	B_-
plus_KLM	B_+ w/o ECL Cluster
plus_KLM_ECL	B_+ w/ ECL Cluster

Tab. 5: List of subsamples using a distinct MVA instance.

¹¹ I.e. there are N (possibly different) variants $f_{\text{D},i}$ for $i \in N$, where N is the number of non-leaf tree nodes.

¹² Most boosting algorithms slightly decrease the weight of correctly classified events after each trained tree.

¹³ Using version 4.2.0.

Naturally, this three-fold split makes evaluation of the MVA more cumbersome because all steps of evaluation and validation have to be performed for the three distinct instances separately. However, using no sub-grouping altogether or grouping only by \mathcal{CP} tag results in a significant loss of separation power.

The kinematic variables used as input to the different MVAs are listed in Table 20 in Appendix F.1, the distributions of all the variables used by each of the MVAs are given in Appendices F.2 to F.4; for the sake of completeness the correlation matrices of the input variables are included as well. No transformation of the input variables has been performed. Refer to Table 19 in Appendix E for a list of parameter settings passed to TMVA for the training of the BDTs.

7.3.1 Overtraining Check

Figure 26 shows the overtraining check plots for the MVAs. The statistical compatibility of the distributions obtained from the training and testing sample has been evaluated using a KOLMOGOROV-SMIRNOV [102] test, the resulting probabilities p_{KS} are listed in Table 6.

MVA Label	p_{KS} (sig)	p_{KS} (bgr)
minus	0.513	0.973
plus_KLM	0.025	0.190
plus_KLM_ECL	$0.25 \cdot 10^{-3}$	0.813

Tab. 6: Results of overtraining check for the MVAs.

minus: Very good statistical compatibility for both signal and background contributions, cf. Table 6. Additional checks such as comparing the signal and background efficiencies for the training/testing samples to those of a reference sample show no indication of overtraining.

plus_KLM: Acceptable statistical compatibility for the background and (less compatible) signal contributions is observed, cf. Table 6. Additional check plots (e.g. the error fraction as function of the tree training index) show that convergence is reached after about one third of the trained trees. However, reduction of the number of trees that are used also reduces the separation power gained from the MVA (in the testing as well as the reference sample).

plus_KLM_ECL: Very good statistical compatibility for the background but very poor agreement for the signal component, cf. Table 6. It is found again that convergence is reached early on (for this MVA at about one sixth of the trees), however, here too reduction of the number of trees to be trained results in reduced separation power.

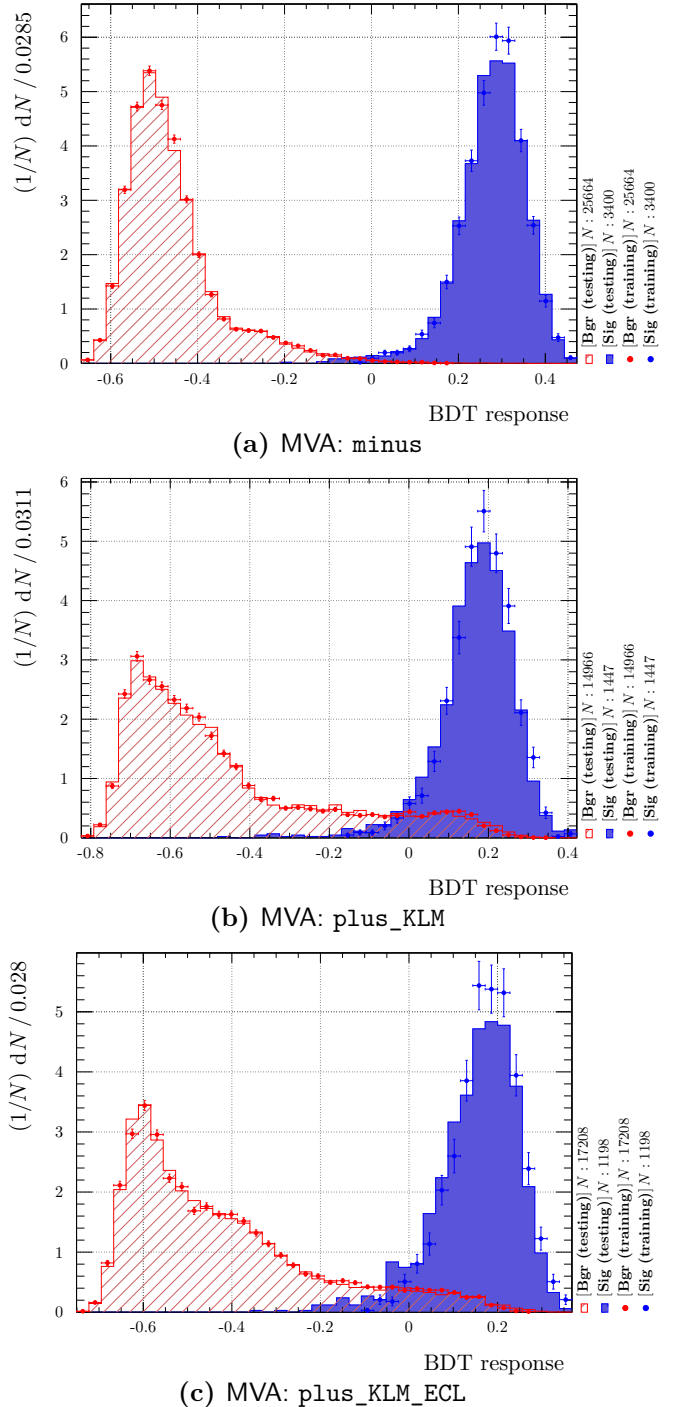


Fig. 26: Overtraining check plots for the MVAs.

In order to combat overtraining several training parameters¹⁴ of the TMVA algorithm have been tweaked with the aim of reducing overtraining while at the same time maintaining separation power.

Pruning was disabled when training the MVAs, because during the initial trials enabling pruning has consistently reduced the power of the BDTs. However, pruning is suggested to be of use when faced with the challenge of overtrained BDTs [103]. Nonetheless, since the MVAs perform consistently well for different reference input samples, the indications for overtraining discussed above should not be taken as showstopper. Furthermore, since the cut values on the respective MVA classifier outputs are determined from a third, statistically independent data sample — which also shows good separation between signal and background, cf. Section 7.3.2 — the effects of possible overtraining on the MVA performance is not severe.

Another important tool for evaluating the performance of an MVA is the Receiver Operating Characteristic (ROC) curve [102, 103, 106], which is a scan of the signal selection efficiency \mathcal{E}_{Sig} vs. the background rejection efficiency \mathcal{E}_{Bgr} . Quite like the conundrum of trading a smaller Type-I error for a larger Type-II error, the two efficiencies are intertwined: it is best to maximise both, however, if a lower limit is imposed on one of them (e.g. by requiring a minimal purity \mathcal{P}) the other becomes fixed.

A measure for reducing the information stored in the ROC curve into a single number is its integral \int_{ROC} (i.e. area below the curve). The hypothetical perfect MVA selects all signal events and is able to reject all background class events, hence it must have a ROC integral equal to unity. Table 7 lists the ROC curve integrals obtained for the three MVAs. During development the ROC curve integral has been taken as figure-of-merit for optimising the MVA training settings and the active input variables.

MVA Label	\int_{ROC} [%]
minus	99.95
plus_KLM	97.22
plus_KLM_ECL	97.84

Tab. 7: ROC curve integrals for the different MVAs.

7.3.2 Selecting the MVA Cut Value

Figure 27 shows the classifier cut efficiency plots for all MVAs. Quite like the ROC curve, these plots can be used to evaluate how well the MVA separates signal and background. A classifier cut efficiency plot shows the efficiencies for selecting signal and background as a function of the cut applied on the MVA classifier output. Additionally, further dependent variables such as the signal purity and the significance¹⁵ — which is used as figure-of-merit for determining the optimal MVA cut value \mathcal{C}_{MVA} — are shown.

These quantities allow for a straightforward evaluation of the MVA performance and, more importantly, provide a simple means for determining the optimal MVA cut value. A steeply falling background efficiency and a similarly steeply rising signal efficiency is a sign for a well-performing MVA. However, since those quantities describe the relative change for the two classes it is of utmost importance to check the signal purity curve. Given a background contribution which is significantly larger than the signal contribution, the background efficiency must become very small in order for a high-quality (that is to say *pure*) output sample. For an efficient MVA the purity curve will rise early on and reach a large number (in an optimal setup it reaches unity).

¹⁴ Specifically the number of trees to be trained, the tree splitting fraction, and the tree depth.

¹⁵ Which is defined as $N_{\text{Sig}}/\sqrt{N_{\text{Sig}} + N_{\text{Bgr}}}$, where N_{Sig} (N_{Bgr}) is the number of signal (background) events which pass the MVA cut.

minus: Figure 27a shows a very well performing MVA: the background efficiency is well below 10 % before there is even the slightest change in the signal efficiency.

Unsurprisingly, the signal purity rises swiftly and reaches a value of 95 % at a signal efficiency of about 99 %. The optimal MVA classifier cut value $\mathfrak{C}_{\text{MVA}}$ and the resulting data properties are listed in Table 8.

plus_KLM: The plus_KLM MVA performs considerably worse than that for minus, however, this is not a surprise considering the difficulties associated with the K_L reconstruction. The most notable feature of Figure 27b is that the background efficiency falls much less steeply and consequently the purity does not increase as abruptly.

Instead of showing a broad peak with steeply falling flanks, the significance increases very gradually, has a more narrow peak, and drops very abruptly afterwards. The fluctuations of the purity for very large $\mathfrak{C}_{\text{MVA}}$ are a result of low statistics in combination with a binned calculation; the optimal $\mathfrak{C}_{\text{MVA}}$ is again listed in Table 8.

plus_KLM_ECL: For plus_KLM_ECL the MVA performance is slightly worse when compared to plus_KLM which is a bit surprising because the former data subset contains more detector observables which should in principle provide additional separation power.

The ECL cluster observables are used as MVA input variables, because it is expected that the cluster shape of non- K_L particles should be different from that observed for a K_L . Apart from an overall reduction of the purity, the discussion of the plot in Figure 27c is the same as for plus_KLM. The best cut value is once again listed in Table 8.

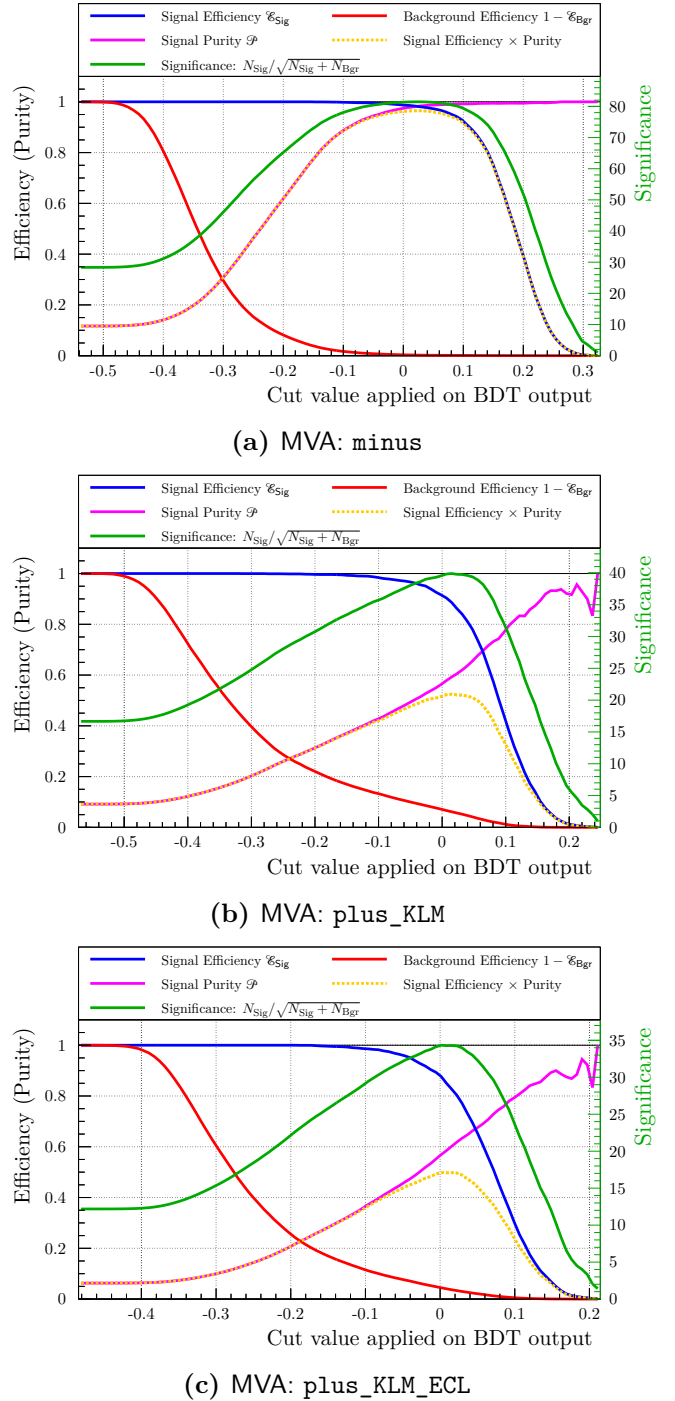


Fig. 27: Classifier cut efficiency plots for the MVAs; these plots have been obtained from a statistically independent reference MC data sample.

MVA Label	$\mathfrak{C}_{\text{MVA}}$	\mathcal{E}_{Sig} [%]	\mathcal{E}_{Bgr} [%]	\mathcal{P} [%]	Significance
minus	0.0	98.1	0.2	98.4	81.47
plus_KLM	0.015	88.6	6.1	59.2	39.90
plus_KLM_ECL	0.0	87.7	4.5	56.9	34.34

Tab. 8: Optimal cut values determined for the MVAs and resulting data properties. The dependent quantities have been evaluated at the optimal MVA cut value $\mathfrak{C}_{\text{MVA}}$.

7.3.3 Study of the Remaining Background

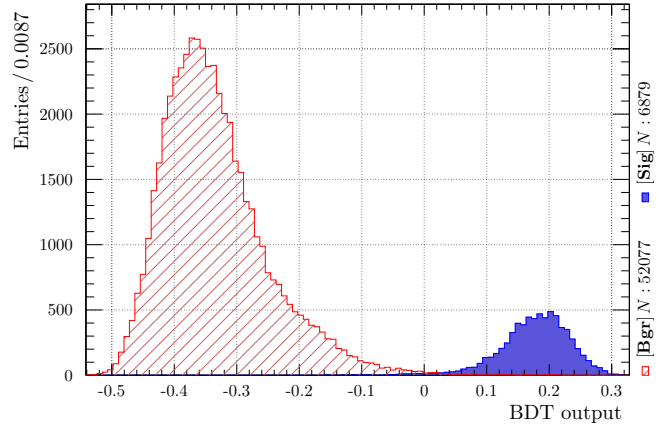
Figure 28 shows the unnormalised classifier output distributions for all MVAs obtained from a statistically independent reference sample. Without normalisation the drastic difference in the abundance of the two event classes becomes apparent: there are around ten times as many background events in the MVA input sample as there are signal events.¹⁶

As is to be expected from the previous discussion, there is very little overlap between the signal and background contributions for the minus MVA (Figure 28a). Separation of the two contributions poses no challenge when using the MVA based approach. Both the plus_KLM and the plus_KLM_ECL MVA produce signal and background distributions which have significant overlap, this is consistent with the previous discussion and the plots shown in Figure 27.

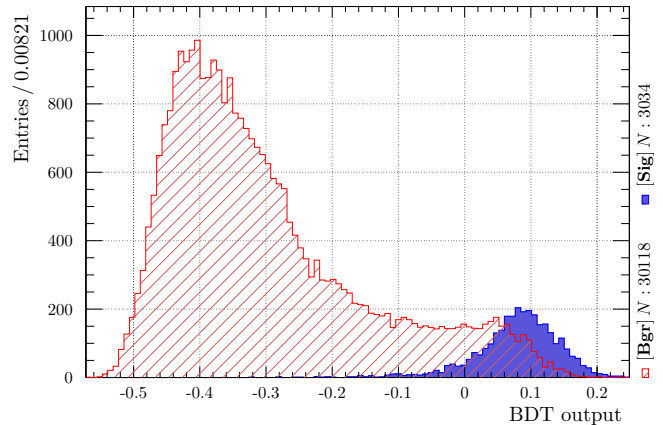
It should be obvious that understanding the sources of the background contributions is key to any further optimisation in the signal selection. Figures 29 and 31 show the distributions of two kinematic variables for the B_- and B_+ subsamples broken down by different background sample categories. Obviously the dominant background sources are generic B^+B^- decays (i.e. the charged sample) and, contributing slightly less, generic non-signal decays of $B^0\bar{B}^0$ (i.e. the mixed sample). The charm and uds contributions are insignificant and — even for the difficult K_L mode — can be reduced to a negligible contribution by applying a simple cut on M_{BC} and $p_{B,Est}^{[c.m.]}$ respectively.

The plots in Figures 30a and 32a show the same data, however, the histograms are broken down by different types of signature fakes. Obviously all reconstructed events should be signal-like to some extent. Therefore, it is to be expected that a significant portion of the background has the correct decay signature ($J/\Psi K_S$ and $J/\Psi K_L$ respectively) but the event is not truly a signal event. To reiterate, there should be a large contribution of events with true J/Ψ and $K_{S,L}$ particles which are, however, not siblings and/or are not the child of a B^0/\bar{B}^0 .

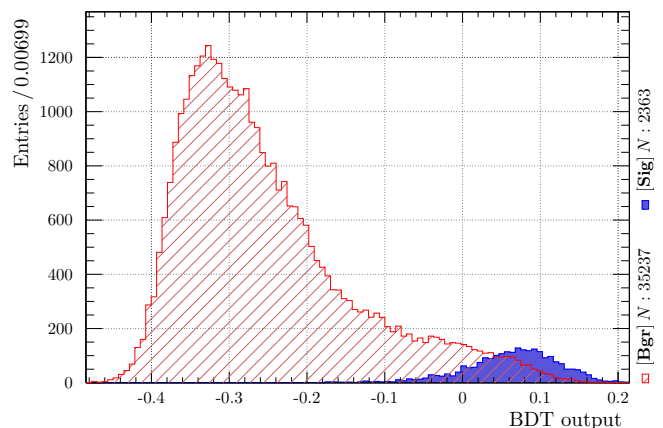
For the B_- mode, the zero-fake contribution (both J/Ψ and K_S are correctly reconstructed) is most dominant at slightly more than 75 %, which is not surprising considering that both particles can be reconstructed



(a) MVA: minus



(b) MVA: plus_KLM



(c) MVA: plus_KLM_ECL

Fig. 28: Classifier output distributions for the MVAs. Note that the histograms have not been normalised, hence, the relative contributions correspond to the expected data class fractions. The plots have been created using a statistically independent reference MC data sample.

¹⁶ The number varies from about 7.5 for minus to nearly 15 for plus_KLM_ECL.

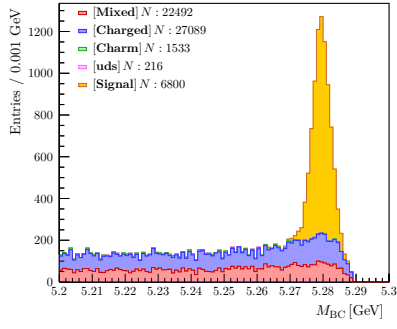
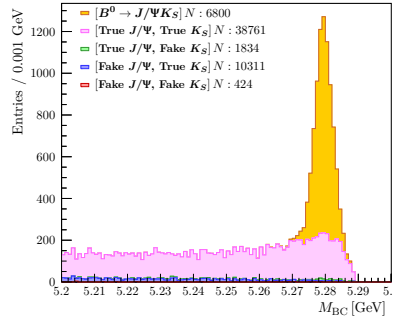
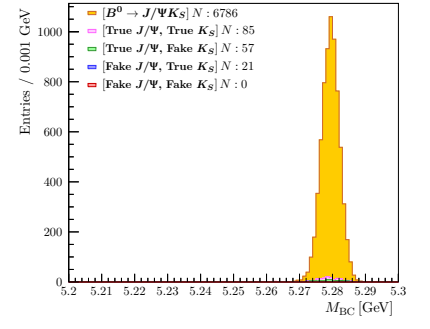
(a) Without imposing \mathcal{C}_{MVA} .(b) After imposing \mathcal{C}_{MVA} .

Fig. 29: Distribution of beam constrained mass M_{BC} for B_- mode; shown are different sample contributions corresponding to signal and different samples of background events.

Fig. 30: Distributions of beam constrained mass M_{BC} for B_- mode before and after application of the MVA classifier cut. Shown are different sample contributions corresponding to signal and different types of fake events.

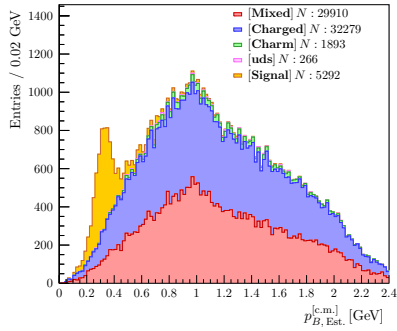
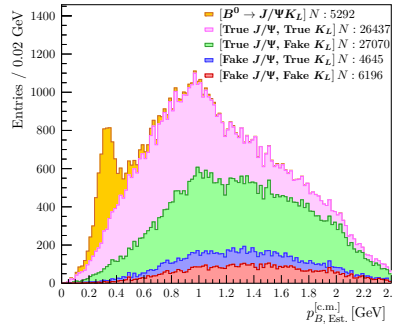
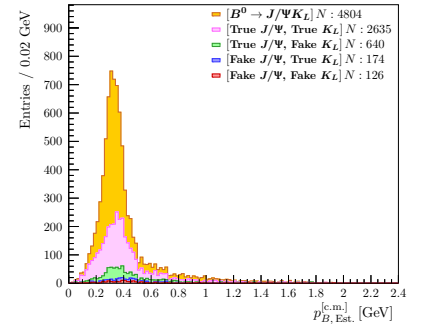
(a) Without imposing \mathcal{C}_{MVA} .(b) After imposing \mathcal{C}_{MVA} .

Fig. 31: Distribution of the estimated B_{CP} absolute three-momentum in centre-of-mass system $p_{B, Est.}^{[c.m.]}$ for B_+ mode; shown are different sample contributions corresponding to signal and different samples of background events.

Fig. 32: Distributions of the estimated B_{CP} absolute three-momentum in centre-of-mass system $p_{B, Est.}^{[c.m.]}$ for B_+ mode before and after application of the MVA classifier cut. Shown are different sample contributions corresponding to signal and different types of fake events.

very well. It can be found that these background events come from generic $B\bar{B}$ decays. The second largest contribution is the “fake J/Ψ true K_S ” subset, which is not surprising either, considering the individual fake rates for J/Ψ and K_S discussed in Sections 2 and 3.

Unsurprisingly, the zero-fake contribution is much smaller in the B_+ mode at about 41%. The most dominant contribution comes from “true J/Ψ fake K_L ” events, which is fully consistent with the expectations and the discussion in Section 4.

Of much greater interest, however, are the distributions shown in Figures 30b and 32b. These show the same type of plot after the MVA cut has been applied to the data. Obviously there is an insignificant amount of background left in the B_- mode and most of it is difficult to reduce because it contains a true J/Ψ . More interesting is the situation for the B_+ mode, where the remaining zero-fake contribution amounts to more than 73%. This might explain why the MVA failed to further increase the background efficiency: because the estimated four-momentum of the K_L is a guess at best, there is no direct handle onto a B^0/\bar{B}^0 discriminating variable which in turn makes the zero-fake background an effectively irreducible background.

Further reduction of the background contribution to the B_+ might be feasible if new variables are found which allow evaluation of the signal-likeness of the $J/\Psi K_L$ pair.

8 Continuum Suppression

Typically, analyses at B factories face the challenge to cope with the so called *continuum background* (cf. Table 3 in Chapter IV), that is the combined background of $e^+e^- \rightarrow q\bar{q}$, $q \in \{u, d, s, c\}$. However, as was discussed in the previous section, the continuum contribution is rather small even before application of the MVAs. Nonetheless, three variables which are typically used for continuum reduction have been included in the MVAs, hence, the basics of continuum suppression will be discussed in this section.

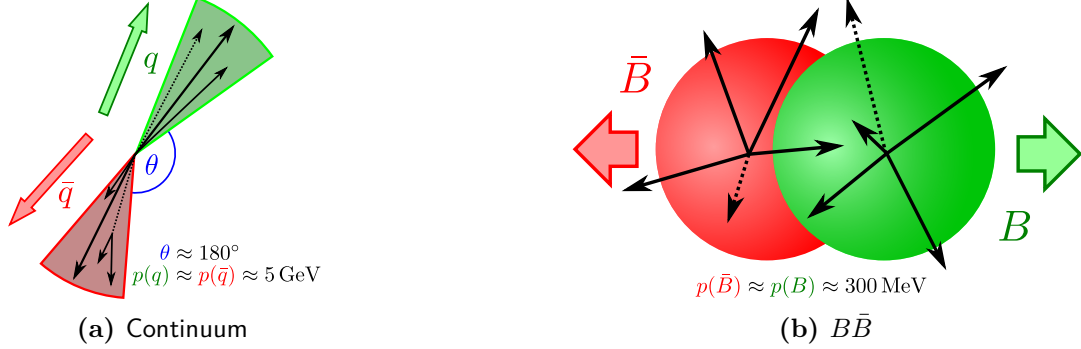


Fig. 33: Illustration of the typical expected event shape for continuum and $B\bar{B}$ events.

The typical event shape of continuum and $B\bar{B}$ events is distinctly different, Figure 33 illustrates this. Due to their significantly smaller mass and hence lesser amount of energy required for production, the continuum event state decays back-to-back with large absolute momenta. Perfect events should show very clear separation of the two hemispheres of the $q\bar{q}$ decay, cf. Figure 33a. All secondary particles (e.g. the hadrons produced by the fragmentation) are expected to have only a small momentum component perpendicular to the original quark's flight direction. Due to its compressed yet elongated shape this type of event shape is called a *jet*.

Naturally the $B\bar{B}$ system decays back-to-back as well, however, there is much less remaining energy: the $B\bar{B}$ system is nearly at rest in the centre-of-mass system after production. Because the B is a pseudoscalar, there is no preferred decay direction for its daughters. Hence, their decay direction distribution should be isotropic. As a result of these properties the event shape is expected to be much more terse and somewhat spherical. Furthermore, the two hemispheres should overlap to some extent, cf. Figure 33b.

Conceptually, every continuum suppression variable is designed to take advantage of these event features. Two approaches which are part of the standard Belle toolkit, the FOX-WOLFRAM moments H_i and the cosine of the thrust angle $\cos\theta_{\text{Thrust}}$, are used by this analysis; they will be discussed below.

§1. FOX-WOLFRAM Moments: The FOX-WOLFRAM moments have been introduced as observables for the description of the event shape in e^+e^- events in the late 1970s [107, 108]. Later their utility as continuum suppression classifiers was discovered. The k^{th} FOX-WOLFRAM moment H_k is defined as

$$H_k := \sum_{ij} \frac{|\vec{p}_i||\vec{p}_j| \cdot P_k(\cos\theta_{ij})}{E_{\text{Vis}}^2},$$

where \vec{p}_i (\vec{p}_j) is the momentum of the i^{th} (j^{th}) particle, $P_k(\cos\theta_{ij})$ is the k^{th} LEGENDRE polynomial, with θ_{ij} being the angle between the momenta of the i^{th} and j^{th} particle, and E_{Vis} is the total visible (measured) energy of the event. The ratios $R_i := H_i/H_0$, $i > 0$ are called the *normalised FOX-WOLFRAM moments*, which are commonly used for continuum suppression.

This analysis uses the 2nd and 4th FOX-WOLFRAM moments which are calculated using the `FoxWolfram` module of the `BASF`. Since they are used as input variables to the MVAs, the resulting distributions are shown in Appendix F.

§2. Cosine of Thrust Angle: The notion of the thrust of an event — first defined in 1977 [109] — became a widely used quantity in the 1980s as a tool to find, reconstruct, and eventually quantify jets. Since continuum suppression is all about determining jet-likeness, the thrust is a sensible starting point.

The concept of thrust comes with two quantities, the *thrust axis* \vec{T} and the (scalar) *thrust* T . The former is defined as the direction which maximises the longitudinal component of the momenta of all particles, i.e.

$$\vec{T} := \left\{ \vec{T} \text{ such that } |\vec{T}| = 1 \text{ and } \max \sum_i |\vec{p}_i \cdot \vec{T}| \right\},$$

the latter is then given as

$$T := \frac{\sum_i |\vec{p}_i \cdot \vec{T}|}{\sum_i |\vec{p}_i|}.$$

For highly isotropic events the thrust T should approach 0.5 whereas for events with a strong directional preference (and hence alignment with the thrust axis) the thrust should approach a maximum of 1. In order to tackle the problem at hand, the variable $\cos \theta_{\text{Thrust}}$ is introduced: this is the angle between the thrust axes calculated for the reconstructed B candidate and the complete remainder of the event.

Following the above reasoning, for a signal event the thrust axis of a reconstructed B should point in some random direction, hence $|\cos \theta_{\text{Thrust}}|$ should be distributed uniformly for $B\bar{B}$ events. Continuum events on the other hand have a strong directional preference and the (wrongly) reconstructed B is likely located within one of the jets with its thrust axis aligned with the jet's axis. Hence, for continuum events the distribution of $|\cos \theta_{\text{Thrust}}|$ is expected to increase towards unity.

This quantity is calculated using the **Thrust** module of the **BASF**, it is used as input variable to the MVAs, hence the calculated variable distributions can be found in Appendix **F**.

VI Measuring the \mathcal{T} -Violation

Prior to being able to perform the actual \mathcal{T} -violation measurement fit one final ingredient is missing: the *time resolution function*. The approach which was used to obtain an estimate of this object is discussed in Section 1. Next, the \mathcal{T} -violation fit method, its implementation, and the result of the fit are discussed in Section 2. As part of the presentation of the fit implementation, the specific (algorithmic) objects composing the complete implementation are listed. This is done to provide a complete reference of the algorithmic procedure that is used by this analysis.

1 Estimating the Time Resolution

A precise estimate of the time (difference) resolution function is crucial to this analysis. In principle a measurement must extract the true Δt from the measured Δz by unfolding the vertex resolution from the reconstructed value while at the same time correcting for possible bias and mis-reconstruction effects [110]. Because the average Δz is of the same order as the vertex resolution, an accurate estimate of the vertex resolution is crucial to any time-resolved measurement [110].

Several time-resolved studies have been performed by Belle [1, 74, 111, 112] and a dedicated study has investigated the time resolution function [110], however, there exists no official Belle tool for obtaining and using the time resolution function. Hence, as part of this analysis, a custom approach was designed to model and extract the time resolution.

In principle it is possible to extract the resolution function from MC data by calculating the distribution of the difference between the true and reconstructed time difference.¹ This distribution can then either be used as a lookup-table for a binned, approximate (but model independent) resolution function or a fit of a model function to the data can be performed yielding a proper function. Obviously, since the latter approach uses a fit model, it is prone to error due to mis-modelling.

Following the approach used by the *BABAR* collaboration in their \mathcal{CP} -violation studies [3], a triple-Gaussian function is used to model the time resolution. The PDF used in the fit is given by

$$\begin{aligned} \mathcal{R}(\Delta t_{\text{Tru}} - \Delta t_{\text{Rec}}) \propto & \mathcal{G}_{\text{Core}}(\mu_{\text{Core}}, \sigma_{\text{Core}}) \\ & + f_{\text{Medium}} \cdot \mathcal{G}_{\text{Medium}}(\mu_{\text{Medium}}, \sigma_{\text{Medium}}) \ , \\ & + f_{\text{Tail}} \cdot \mathcal{G}_{\text{Tail}}(\mu_{\text{Tail}}, \sigma_{\text{Tail}}) \end{aligned} \quad (16)$$

where $\mathcal{G}(\mu, \sigma)$ is the GAUSS function. The mean values μ of all three components are expected to be around zero: the fit parameter range is restricted to $[-1, +1]$. Their widths have been sectioned such that there are no overlapping ranges: $\sigma_{\text{Core}} \in [0, 1]$, $\sigma_{\text{Medium}} \in [1, 4]$, $\sigma_{\text{Tail}} \in [4, 12]$. Apart from the three mean values and widths, two scale factors f_{Medium} and f_{Tail} tune the relative contributions of the individual functions. This fit function has been implemented in the *Roofit* framework [113, 114] and was fitted to the $\Delta t_{\text{Tru}} - \Delta t_{\text{Reco}}$ distributions, the fit result is shown in Figure 34

Note that the data set that was used for fitting is statistically independent from the other data samples used by this analysis. Furthermore, only the signal contribution has been used, because calculation of Δt_{True} is only possible for signal MC. Obviously this introduces some mis-modelling, however, it is the most consistent approach: there is no clear definition of time resolution for events which are not of

¹ Naturally, this assumes that the MC models the detector resolution and inefficiencies sufficiently well.

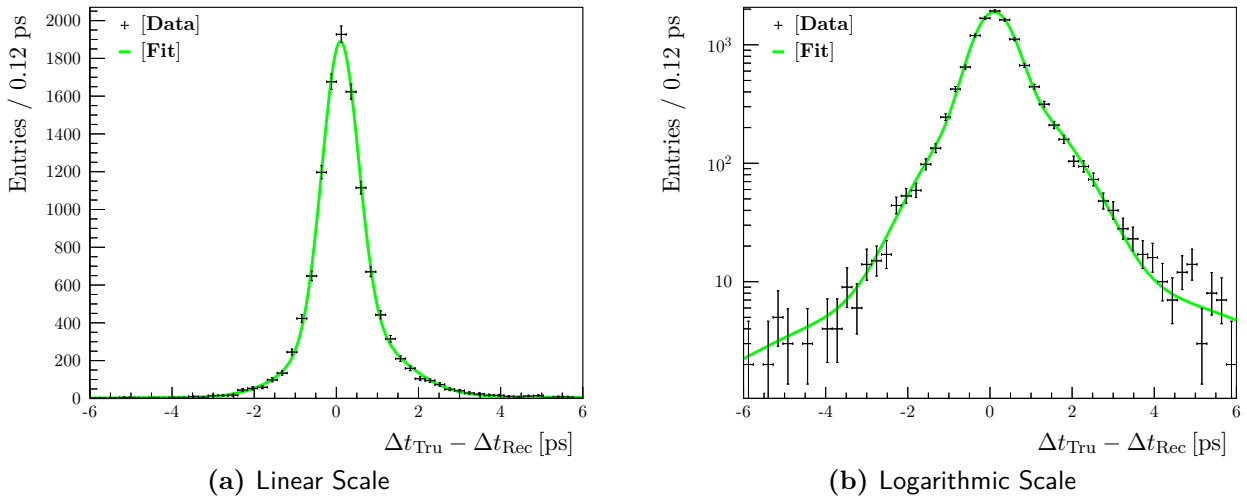


Fig. 34: Result of the time resolution extraction fit. The data are signal-only MC events taken from a statistically independent data set.

signal type. A more correct but also more time-consuming approach would require detailed study of the time difference distributions for background events. The knowledge gained by such a study could then be used to introduce a background time distribution component into the signal PDF, that is the PDF given by Equation (8). In addition to a background time difference study this would require a detailed signal-background composition study of the final data set since the actual composition fractions become part of the final fit. Due to lack of time and since there are not indications suggesting that the time resolution function is poorly modelled, this approach has not been considered any further.

Following the discussion in Section 5 in Chapter V it is expected that the time resolution function should be equivalent for \mathcal{CP} -odd and -even tagged events. Indeed, if the fit is performed separately for those two classes, the parameter values that are obtained are in agreement with each other and the fit value obtained for the combined data. Hence, in the subsequent \mathcal{T} -violation fit the fit result obtained from the combined data set is used.

The values obtained from the fit for the time resolution function parameters are listed in Table 9. Note that the quoted uncertainties are fit uncertainties as calculated by RooFit. Evaluation of systematic uncertainties has not been performed for this fit; this includes the model uncertainty. That is to say, the correctness of the three-Gauss resolution model — i.e. the PDF given by Equation (16) — has not been evaluated. The uncertainty resulting from assuming this specific model has not been assessed.

The overall agreement of the fit functions and the data points is quite good: data points and fit function agree within the statistical uncertainty over most of the fit range. Using RooFit a measure of statistical compatibility of data and fit can be calculated,² the resulting value is $\chi^2_{\text{Red}} = 1.1$ which is an excellent agreement.

Variable	Fit Result
μ_{Core}	0.094 ± 0.008
σ_{Core}	0.44 ± 0.01
μ_{Medium}	0.34 ± 0.03
σ_{Medium}	1.17 ± 0.03
μ_{Tail}	0.99 ± 0.04
σ_{Tail}	4.00 ± 0.04
f_{Medium}	0.36 ± 0.02
f_{Tail}	0.036 ± 0.003

Tab. 9: Results obtained for the parameters of the time resolution function — cf. Equation (16) — after performing a fit to the MC distributions, cf. Figure 34.

2 Evaluating the \mathcal{T} -Violation

In Section 2.1 the implementation of the fitting procedure is briefly summarised. Following that, the fit application routine is outlined in Section 2.2. Finally, in Section 2.3 the results obtained from the fitting program are presented and discussed.

² Essentially this quantity is the reduced chisquare compatibility, i.e. χ^2 divided by the number of degrees of freedom, of the two components after filling them into a binned data object.

2.1 Implementation

The fit to the Δt distributions is implemented using the `Roofit` framework [113, 114]. Recalling that the base PDF given by Equation (6) is parametrised by ω , α , and β , a `C++` class is created that models this PDF.³ Because the class is parametrised in ω , α , and β too, a total of eight instantiations is required to fully describe all signal components.

For each of these instantiations the convoluted PDF of $g_{\alpha,\beta}^{\omega} \otimes \mathcal{R}(\delta t, \sigma_{\Delta t_{\text{rec}}})$ is calculated using the `RoofitConvPdf` class which implements the convolution using a *Fast-Fourier-Transform*. By prioritising the convolution — which is introduced only in the last step of the derivation in Chapter III — the consumption of computational resources during the fit is reduced significantly.

When performing the fit, the internal fitting routines must treat the result of a convolution as a black box. That is to say, individual components of the function going into the convolution cannot be separated by the fit routine after performing the convolution. Hence, the later the convolution is performed, the more components are combined into the box. Therefore, it is computationally favourable to include the convolution at the innermost possible level of the final PDF. Of course this only holds true if the function to be convoluted is composed of contributions which can be separated when performing a convolution. Furthermore, actual gain is only obtained if the contributions are fit targets to different phase-space areas because otherwise the fit targets are required at each fitting step anyways and thus the convolution must be recalculated for each fitting step too.

Afterwards the instantiations which correspond to the same \mathcal{CP} tag and time-ordering (shared values for β and ω) are combined⁴ using the prescription given by Equation (7). This yields a set of eight `C++` classes, called *tagged* PDFs, each of which includes modelling of the time resolution and the event-by-event mis-tag rate. The *tagged* PDFs are pairwise combined⁴ such that four PDFs are created which include both time-orderings for a given \mathcal{CP} and flavour tag.⁵

Finally, the four PDFs are combined into a `RoofitSimultaneous` object — which is also given the required flavour and \mathcal{CP} tag categories for each of the four input PDFs. This enables `Roofit` to perform the fit in a simultaneous manner and to all expected types of (signal) data equivalently. Furthermore, this class selects only that component of the PDF at each step of the fit which is appropriate given the event's \mathcal{CP} - and flavour-tag. No background modelling contribution is included in the fit.

2.2 Application

Prior to performing the fit, the data samples obtained from applying the MVA cut (cf. Section 7.3.2 in Chapter V) are reweighted using a global type-dependent event weight. For each of the \mathcal{CP} tag categories a global efficiency

$$\varepsilon_{\text{Global}}(\xi) := \frac{\# \text{ of selected events with } \xi_{\mathcal{CP}} = \xi}{\# \text{ of input events with } \xi_{\mathcal{CP}} = \xi}$$

is calculated. Then the input event weight to the `Roofit` routines for \mathcal{CP} -odd events w_- is set to unity, while the \mathcal{CP} -even events are weighted with $w_+ = \varepsilon_{\text{Global}}(+1)/\varepsilon_{\text{Global}}(-1) = 1.138$.

The fit range is restricted to $\Delta t \in [-7, 7]$ ps in order to reduce the influence of the low-statistics phase space areas on the fit result. This requirement reduces the (unweighted) number of input events from 11 444 to 10 973 (a reduction of 4.1 %). However, it should be noted that about 40 % of the removed events have unphysical values for the reconstructed Δt (i.e. values much larger than $|\Delta t| = 100$ ps).

Finally, the actual fit is performed using the `fitTo` routine of the `RoofitSimultaneous` class with the `Extended` option activated. Because the data are read using the `RoofitDataSet` class, the fitting is performed in an unbinned manner. The fit results will be discussed in the following section.

2.3 Results

In Figure 35 the *raw time asymmetry* $\mathcal{A}_{\mathcal{T}}$ for the four independent sets is shown. Each independent set corresponds to one of the combinations of reference and \mathcal{T} -conjugate transitions that are listed in

³ The class derives from `RoofitAbsPdf` in order to take advantage of the fitting routines provided by `Roofit`.

⁴ Using the `RoofitAddPdf` class.

⁵ I.e. pairs with shared values for α and β but different values for ω are combined.

Table 12a in Appendix B (cf. also Section 2.2 in Chapter III). Hence, the raw asymmetry for the first combination is defined by

$$\mathcal{A}_{\mathcal{T}}^{\bar{B}^0 B^-}(\Delta t) := \frac{\mathcal{D}_{l^-, K_L}^-(\Delta t) - \mathcal{D}_{l^+, K_S}^+(\Delta t)}{\mathcal{D}_{l^-, K_L}^-(\Delta t) + \mathcal{D}_{l^+, K_S}^+(\Delta t)},$$

where \mathcal{D}_{l^-, K_L}^- (\mathcal{D}_{l^+, K_S}^+) is the time-dependent frequency distribution⁶ of events where a \bar{B}^0 and a K_L (a B^0 and a K_S) have been reconstructed and the decay to the \mathcal{CP} final state occurred before (after) the decay to the flavour final state. The three remaining asymmetries can be constructed similarly. These asymmetry plots provide a visual cue of the asymmetry w.r.t. to \mathcal{T} that is to be found in the data. A perfectly symmetric data input should result in a flat distribution at zero whereas \mathcal{T} -violation manifests in a sinusoidal distribution.

The plots in Figures 35a and 35b clearly show an asymmetry w.r.t. to \mathcal{T} . However, the plots in the bottom row show a significantly less asymmetric behaviour: for Figure 35c nearly half of the data points are compatible with zero (10 out of 22), for Figure 35d that number is even higher at 15 out of 22.

Figure 36 shows the time difference distributions for each of the four $\mathcal{CP} \times \text{flavour}$ -tag categories along with the (fitted) signal PDF contribution modelling that specific category; the resulting parameter values are listed in Table 10.

Variable	Fit Result	Variable	Fit Result
$S_{l^+ K_L}^+$	-0.860 ± 0.005	$C_{l^+ K_L}^+$	0.4 ± 0.1
$S_{l^+ K_L}^-$	0.60 ± 0.02	$C_{l^+ K_L}^-$	0.700 ± 0.004
$S_{l^+ K_S}^+$	-0.988 ± 0.004	$C_{l^+ K_S}^+$	0.688 ± 0.009
$S_{l^+ K_S}^-$	0.200 ± 0.003	$C_{l^+ K_S}^-$	0.155 ± 0.006
$S_{l^- K_L}^+$	1.198 ± 0.008	$C_{l^- K_L}^+$	0.05 ± 0.1
$S_{l^- K_L}^-$	-0.598 ± 0.007	$C_{l^- K_L}^-$	0.31 ± 0.08
$S_{l^- K_S}^+$	-0.573 ± 0.007	$C_{l^- K_S}^+$	0.13 ± 0.07
$S_{l^- K_S}^-$	1.195 ± 0.003	$C_{l^- K_S}^-$	-0.44 ± 0.06

Tab. 10: Parameter values obtained from the \mathcal{T} -violation fit. The listed uncertainties correspond to the values estimated by RooFit as part of the fitting routine.

Generally the fitted functions show very good agreement for the left flank of the respective distributions. However, the agreement in the peak region is generally rather poor which is further discussed below. Several of the remaining data points (i.e. those at the peak and the right flank) lie at a significant distance from the fit result curve, however, most of them are still in agreement with the fit curve within the error box. For Figure 36d the deviations are most severe: nearly one third of the data points do not agree within their error box. The reduced chisquare χ_{Red}^2 , calculated by RooFit for the combinations of data and fit result, varies from 2.7 (acceptable compatibility) to about 8.3 (poor compatibility).

The parameter values obtained from the fit are not in agreement with the values published by *BABAR*, this holds true for both their MC study [47] as well as the actual analysis [5]. More specifically, the cosine parameters $C_{\alpha, \beta}^\omega$ — which are expected to be zero in the SM — are mostly far off from zero: only one parameter is compatible with zero. The sine parameters $S_{\alpha, \beta}^\omega$ are expected to be around ± 0.7 which is not matched by a single parameter result: the fit result fails to match the SM expectation.

In order to investigate the stability and aptitude of the fit routine, the starting values of the fit parameters have been varied. Table 11 lists the resulting parameter values for the variables S_{l^+, K_L}^- and $C_{\alpha, \beta}^-$ for different starting values along with the relative change of the result (w.r.t. the nominal value). It can be seen that the fit of the cosine parameter is highly unstable: a relative change of nearly 60% can be observed. Similarly, the sine parameter shows instability, albeit the maximum observed relative change

⁶ I.e. a lookup function for the respective time difference distribution. The notation follows the same scheme used already previously to denote the data subgroups, cf. Section 2.4.1 in Chapter III.

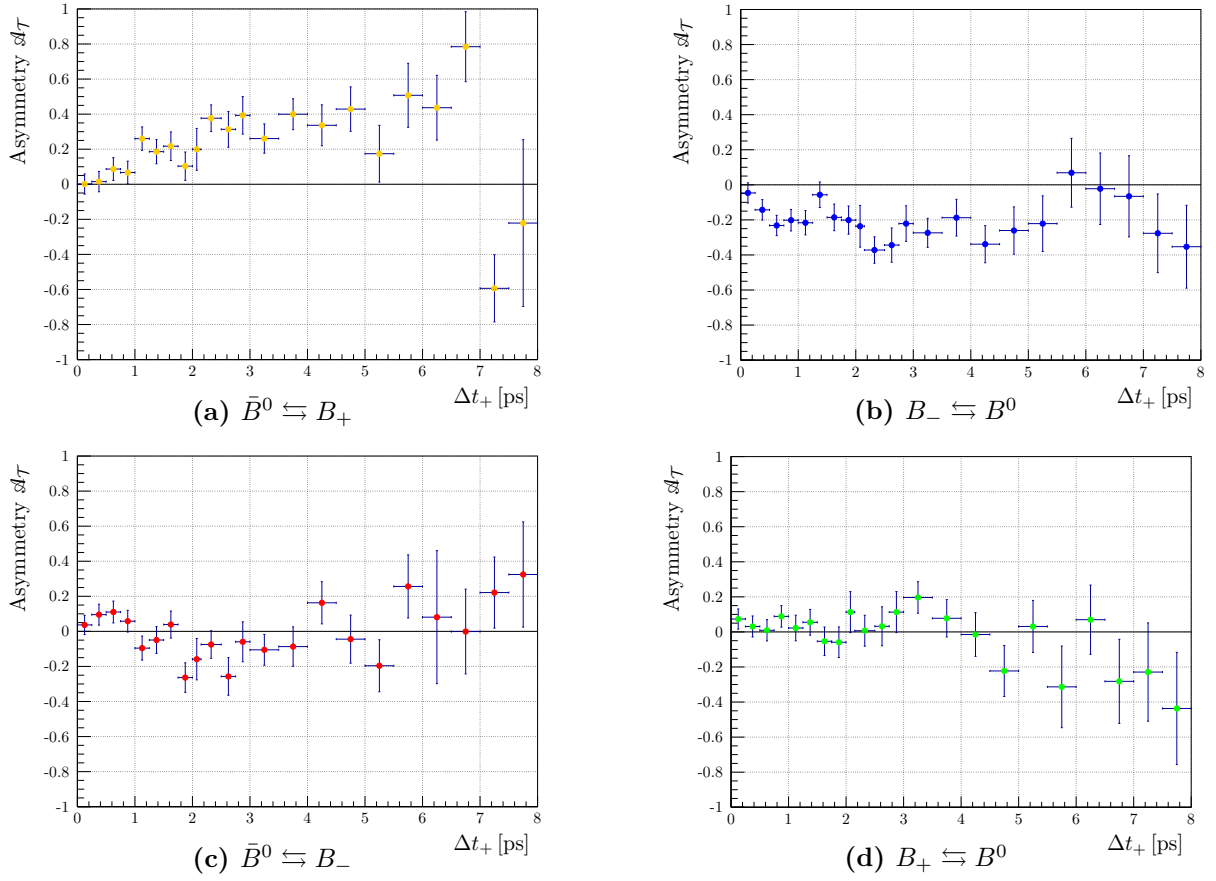


Fig. 35: The raw \mathcal{T} -asymmetries shown for the four independent sets corresponding to the four possible comparisons between the reference and \mathcal{T} -conjugate transitions. The horizontal error bars denote the bin width, the vertical bars the statistical error which has been propagated from the inputs given to the asymmetry calculation.

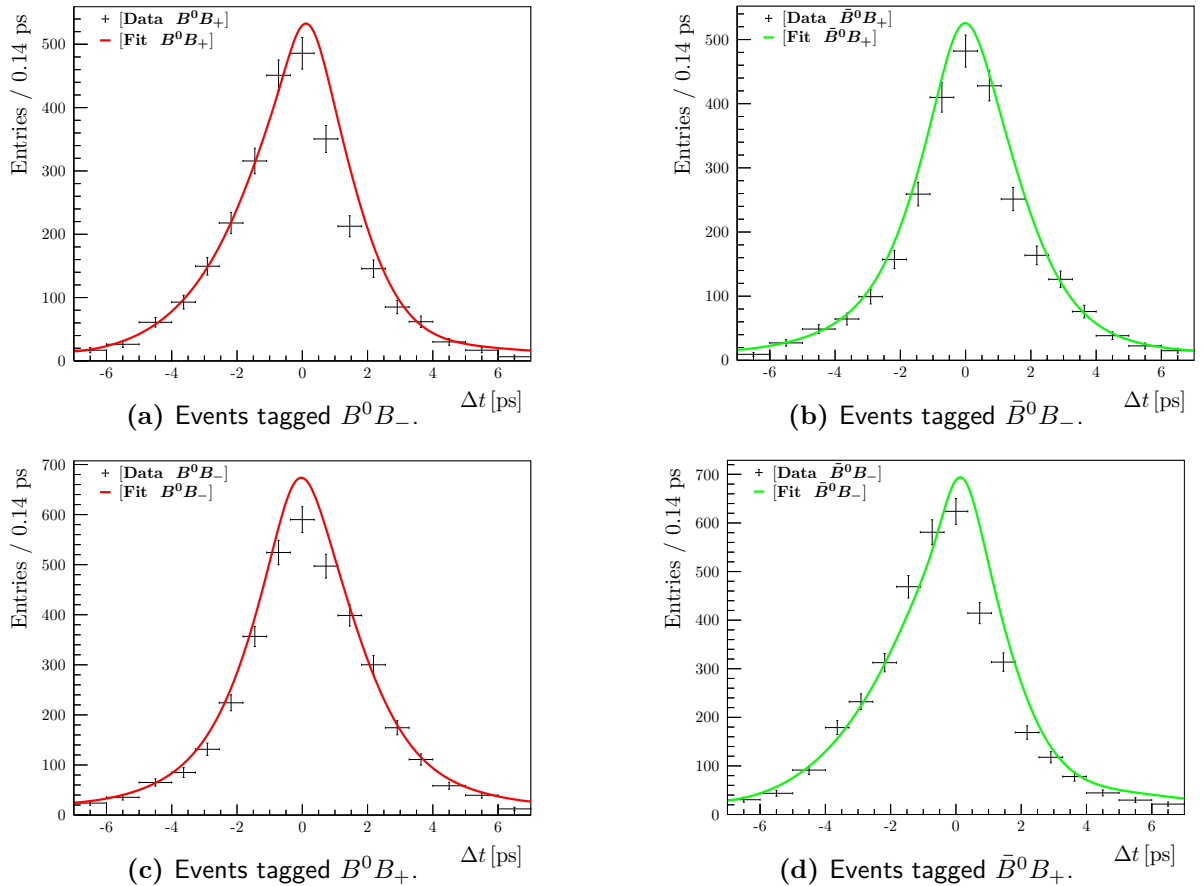


Fig. 36: Plots showing the result of the \mathcal{T} -violation fits to the time difference distributions. The horizontal error bars denote the bin width (of the distribution created for plotting), the vertical bars the statistical error of the data.

is much smaller at about 13%. Clearly, there is a significant dependence of the fit result on the choice of initial parameters.

Restriction	Fit Result		Relative Change [%]	
	$S_{l^+K_L}^-$	$C_{l^+K_L}^-$	$S_{l^+K_L}^-$	$C_{l^+K_L}^-$
Normal	-0.860 ± 0.005	0.4 ± 0.1	—	—
$S_{\text{Init}} = \pm 0.3$	-0.971 ± 0.002	0.237 ± 0.006	12.9	45.6
$S_{\text{Init}} = \pm 0.9$	-0.938 ± 0.009	0.352 ± 0.002	9.1	20.7
$C_{\text{Init}} = +0.5$	-0.867 ± 0.006	0.4 ± 0.1	0.8	4.6
$C_{\text{Init}} = -0.5$	-0.97 ± 0.06	0.2 ± 0.1	13.2	59.3

Tab. 11: Results of varying the starting values S_{Init} (C_{Init}) of the $S_{\alpha,\beta}^\omega$ ($C_{\alpha,\beta}^\omega$) PDF parameters. The fit ranges $S_{\alpha,\beta}^\omega \in \pm[0.2, 1.2]$ and $C_{\alpha,\beta}^\omega \in [-0.7, 0.7]$ have been used, the label Normal refers to $S_{\text{Init}} = \pm 0.7$, $C_{\text{Init}} = 0$.

The allowed parameter ranges used for the fits shown in Figure 36 have been chosen to be very wide⁷ in order to avoid biasing the fit by the SM expectation. However, since the fit result matches both data and expectation rather poorly additional, more restrictive, fit runs have been performed.

Figure 37 shows plots comparing the different fit results obtained for varying degrees of restriction of the allowed parameter space. The case shown in Figure 37a exhibits the smallest change resulting from the variation, whereas the one shown in Figure 37b changes the most. Taking a closer look at Figure 37b it becomes obvious that any gain in compatibility with the peak region results in a degradation of the agreement on the right flank. The reduced chisquare goodness-of-fit classifier χ_{Red}^2 (listed in the plots as calculated by RooFit) of the different fit result curves shows that the *Normal* setting produces the best results in terms of statistical compatibility.

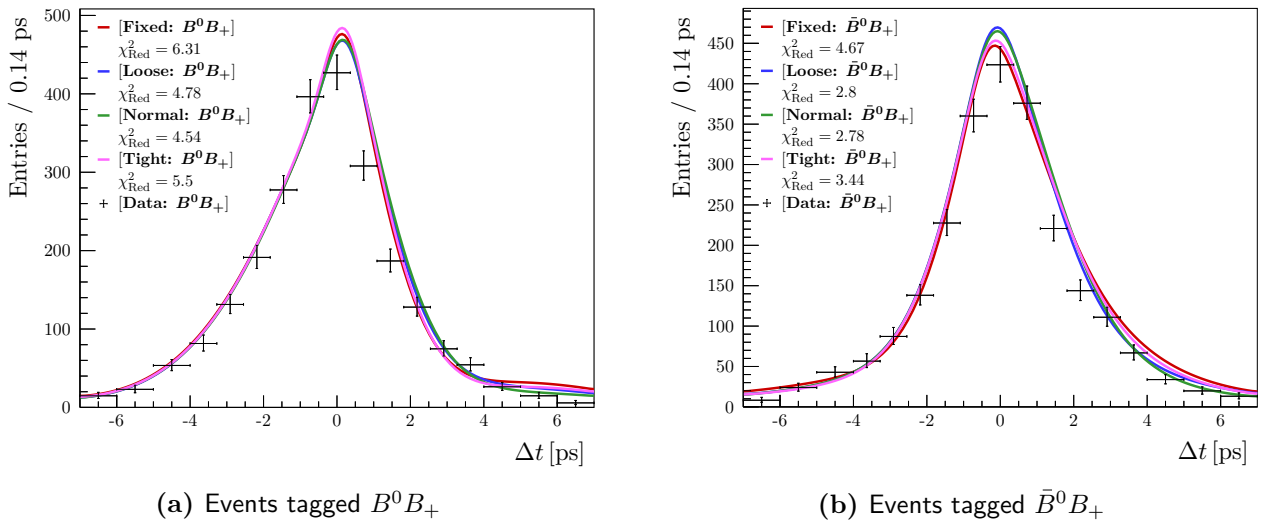


Fig. 37: Plots comparing the effects of varying the ranges of the fit parameters onto the fit result distributions. Shown are the results for the two data subgroups where the resulting change is smallest (a) respectively largest (b). The **Normal** label refers to the parameter range used previously; for **Loose** the range on the sine parameters has been widened to $S_{\alpha,\beta}^\omega \in [-1.4, 1.4]$, the $C_{\alpha,\beta}^\omega$ ranges have been left unchanged; the **Tight** and **Fixed** sets use the same range for the sine parameter as the **Normal** set, however, the cosine parameter ranges have been reduced to $C_{\alpha,\beta}^\omega \in [-0.2, 0.2]$ for **Tight** and set to zero for **Fixed**.

However, the more stunning finding from these two plots is how little the fitted curves change for such a significant change in the allowed fit parameter range. Along with the results of the initial parameter value variations presented previously, this supports the hypothesis that the fit routine as a whole, or the data itself, has a low sensitivity on the physical \mathcal{T} -violation parameters.

⁷ The cosine coefficients $C_{\alpha,\beta}^\omega$ are ranged from -0.7 to 0.7 , the sine coefficients $S_{\alpha,\beta}^\omega$ from 0.2 to 1.2 (or -1.2 to -0.2 if they are expected to be negative).

Naturally, this poses the question which factors are responsible for the poor modelling and sensitivity; the following factors come to mind:

§1. Improper Data The reconstruction scheme has been discussed in great detail in Chapter V and there is no obvious reason not to trust these results. Various tests have been performed to ensure that the reconstructed data is understood sufficiently well. However, if the vertex estimates (and thus time difference measurements) or the \mathcal{CP} or flavour tags are not credible the resulting distributions need not be compatible with the signal PDF.

Another data related factor — strictly speaking it is not an effect of *improper data* — is the overall sensitivity of the data set. That is to say, the fact that the data sample may be too small, or the reconstruction scheme too inefficient or inaccurate in order to allow for a sensitive measurement.

It is unlikely that this is the cause for the poor modelling because the fundamental approach is sound and follows the methods used by previous \mathcal{CP} -violation studies and the *BABAR* \mathcal{T} -violation analysis. As will be discussed in the following chapter, the number of events obtained by this analysis and the data purity is in good agreement (or even better) than that of the data used by *BABAR*.⁸

§2. Poor Time Resolution Estimate Since the convolution is a very important contribution to the final signal PDF, a poorly estimated time resolution function may degrade the fit result. However, one may argue that this source is not very likely the root cause of the problems because of the excellent agreement of the fitted time resolution function with the distribution taken from MC data (cf. Section 1).

Unfortunately there is no official Belle approach to estimating and modelling the time resolution which then could be used as a reference. Hence, evaluation of this effect is somewhat difficult.

§3. Problematic Fit Implementation The most likely cause can be found in the implementation of the fit, possible sources coming into play are

- a) numerical instabilities,
- b) an improper fit approach or implementation, and
- c) bugs in the code.

Essentially, there are two possible reasons for numerical instabilities: either the mathematical PDF is shaped such that it is prone to minimisation failure⁹ or the algorithmic PDF (i.e. the object passed to RooFit's fitting routine) suffers from minimisation failure. Naturally if the algorithmic PDF is at fault, the true origin of the numerical instabilities is source type **b**), i.e. an improper fit implementation.

Application of the fit program to the truth time difference distributions¹⁰ shows that the fit fails to correctly model the distributions yet again. The connection between the two sides of the PDF is discontinuous and not smooth but rather step-like. This leads to the hypothesis that the step-like combination of the two sides of the signal PDF is not implemented in a way that allows RooFit to fit it properly to data.

Naturally, if the modelling of the transition from $\Delta t < 0$ to $\Delta t > 0$ is inadequate, numerical instabilities may be introduced due to poor description of the peak region. The basic PDF given by Equation (6) shows that a continuous connection between both sides can only exist if $C_{\alpha,\beta}^+ = C_{\alpha,\beta}^-$. However, forcing those parameters to be equal does not result in a properly continuous fit result for neither reconstruction level nor truth level MC data. This is not entirely surprising because the **Fixed** fit configuration

⁸ On top of that: the total Belle data sample appropriate for this analysis is larger than that available to the *BABAR* analysis.

⁹ For example because there are many local minima, the local minima are much more favourable than the surrounding phase-space, or the numerical difference between local and global minima is small.

¹⁰ More concretely, instead of using the reconstructed time difference, the MC truth value is used as input to the fit. The time resolution convolution is disabled. Hence, the only reconstruction error comes from the \mathcal{CP} and flavour mis-tag and the truth level time resolution (which should be zero). Background events are not included because they do not provide a truth level time difference.

(cf. Figure 37) fixes all $C_{\alpha,\beta}^\omega$ parameters to be zero (the SM expectation), however, it does not result in a significantly improved fit result.

Apart from the fact that it does not solve the problem, forcing the cosine parameters to be pairwise equal impairs the generality of the fit. A better solution would be to introduce additional contributions to the likelihood function which minimise the difference $|C_{\alpha,\beta}^+ - C_{\alpha,\beta}^-|$.

Finally, there may be bugs in the code which are responsible for the failure of the fit. The code has been checked several times for consistency and correctness. No striking reason for supporting the assumption that the presence of bugs is responsible for its result has been found, hence this cause will not be discussed any further.

VII Conclusion

1 Synopsis of the Achievements

Over the course of the last year an extensive analysis has been developed. A rigorous scheme for efficiently obtaining a signal data sample with very low pollution has been devised and was tested extensively. Much effort and time has been spent on this first stage of the analysis, because only proper input data can provide the foundations required for an efficient and accurate analysis result. This effort and its results have been discussed in detail in Chapter V.

Comparing the final sample composition to that of the \mathcal{T} -violation study performed by *BABAR* [5] verifies the prosperity of the effort: the purities of both the \mathcal{CP} -odd and -even data subsets exceed the values quoted by the *BABAR* paper.

For the \mathcal{CP} -odd mode, the *BABAR* analysis uses additional charmonium modes apart from the J/Ψ , which decreases their overall purity (by at most about 11% relative to our data) but enhances the number of signal events (by about 12% relative to our data). The true gain of the MVA based selection can be seen in the \mathcal{CP} -even mode, where we achieve a purity that is only about 1.5% better, however, the number of signal events is increased by nearly 48% when compared to the numbers quoted by *BABAR*.

Further study has shown that the remaining background contributions are difficult to separate from the signal component due to their signal-likeness.

The time difference resolution function has been estimated using independent MC data. Since there is no official Belle toolkit or method and there are many approaches to estimate the time difference resolution, evaluation of the performance of the approach used by this analysis is somewhat difficult. However, the shape of the function that is used in the fit is well-motivated by previous analyses both by Belle and *BABAR*. More importantly, because the data and fit result are in excellent agreement there is no striking reason to doubt the approach used by this analysis.

The time difference distributions obtained from the data clearly show an asymmetric behaviour, cf. Figure 35. However, the quantification of the extent of the asymmetry — and thus violation of \mathcal{T} — remains elusive. It should be stated without doubt that the proper implementation of the \mathcal{T} -violation fit is a challenge on its own. Although this analysis could not bring the final \mathcal{T} -violation fit to fruition, it provides the data and has laid the groundwork for a subsequent study to focus exclusively on implementing, testing, and — finally — applying the fit.

2 Discussion of the Shortcomings

For the most part, the basic event reconstruction scheme is well-tested and well understood. Although there remain some unanswered questions or hypotheses which require further validation (such as the dual- \mathcal{CP} tag veto, cf. Section 5.3 in Chapter V), there is no doubt that the approach is working and shows no unexpected behaviour.

However, in order to make a complete analysis, the data refinement procedure requires some further validation and cross-checks. Additional work should be put into determining how much overtraining has

occurred, whether it can be reduced, and how it affects the data selected by the MVA. Some work has been put into investigating whether the MVA selection introduces a bias,¹ however, a subsequent study should investigate this further. Doing so also enables to use a more fine-grained (and thus accurate) event weight (e.g. depending on Δt or \mathcal{C}_{MVA}) instead of the plain, global event weight that is used by this analysis, if that is necessary.

The overall performance of the flavour tagging is slightly worse than anticipated, cf. Section 6.1 in Chapter V. It might be worthwhile for a subsequent study to investigate the origin of the inferior performance. Quite possibly, the performance could be enhanced by optimising the flavour tagging algorithm for the specific problem at hand instead of using a rather general scheme. Nevertheless, the expected gain from doing so is rather small: the increase in the effective signal tagging efficiency using a more optimised approach is not expected to exceed a few percent considering what previous optimisation studies have achieved [101].

However, probably the most considerable shortcoming of the current state of the analysis is the final \mathcal{T} -violation fit, cf. Section 2 in Chapter VI. As was already discussed, it is probable that the implementation of the fit is not fully correct. This may either be a fault of the implementation itself or the interaction of the fit implementation and the fitting routine (i.e. `Roofit`). In principle, fixing an error in the implementation is straightforward once that mistake is found. The latter cause on the other hand is more difficult to correct: it requires rewriting the fit model in a way that is specifically tailored to `Roofit`, hence doing so requires extensive knowledge of the inner workings of `Roofit`.

Getting the fit to work, albeit producing a poor result, was a time-consuming challenge on its own. Hence, due to lack of time and familiarity with the internals of `Roofit` this analysis fails to provide a truly sound \mathcal{T} -violation fit.

3 Prospects and Outlook

In order to push the analysis to the next level, all subsequent work should start with clearing up any remaining doubt in the aptitude of the selected data. In the previous section and in Chapter V some open questions have already been discussed, however to recapitulate, below follow the most important aspects:

- Additional background studies should verify the hypothesised mechanisms responsible for the remaining background contribution. Doing so might also enable a further reduction of the background contamination.
- When the question of systematic uncertainties comes up, the MVA performance should be investigated further. Specifically, how much overtraining is present, whether it affects the selection, and whether the MVA based selection introduces bias effects should be studied.
- The overall data and MC agreement should be investigated. This cross-check has only been performed on a few kinematic variables (such as the J/Ψ mass) which have shown no staggering deviations. However, specifically the modelling of the time difference measurement and the time difference resolution should be verified. It is highly unlikely that this is not modelled correctly in MC since it is crucial to many of the analyses performed by Belle. Nevertheless, particularly because there is no official scheme for estimation of the time difference resolution this must be treated with great care, hence also validated with the very same care.

Furthermore, this analysis does not use the complete appropriate Belle data sample but only the *SVD II* part, cf. Section 3 in Chapter IV. An increase in statistics might improve the overall performance. However, it should be noted that the time difference resolution is expected to be worse for the elided data because the *SVD I* data uses an inferior vertexing system. Therefore, inclusion of *SVD I* data necessitates using different time difference resolution functions for the two data groups. By including the *SVD I* data, the total data available is increased by about 24.5%, cf. Chapter IV.

The previous discussions should have made it clear, that the \mathcal{T} -violation fit implemented as part of this analysis is flawed. Whether this is due to false assumptions, false implementation, or incorrect steering

¹ E.g. checking that there is no shape dependence of the Δt distributions on MVA cut value.

of `RootFit` could not be determined to satisfaction. It should also be obvious that the contribution of the fit implementation to the overall analysis is rather heavyweight in terms of the associated technical difficulty and the time required for implementation and testing. However, the difficulties encountered with the fit do not diminish the prospects for this measurement itself. Both, the fundamental methodology — cf. Chapter III — and the data reconstruction scheme are sound.

My personal recommendation for a proceeding analysis is to work very closely with the Belle \mathcal{CP} -violation group. Furthermore, I strongly advise against reusing any part of my \mathcal{T} -violation fit routines. A subsequent analysis should utilise the reconstruction and selection scheme as well as the time difference resolution function created as part of this thesis; the actual \mathcal{T} -violation fit should be rewritten from scratch.

The final fit provides plentiful challenges, hence, subsequent work should be performed in close contact with fellow scientists who have extensive experience in fitting and know the ins and outs of doing a time-resolved \mathcal{CP} -violation analysis.

In hindsight I would recommend to start by performing a time-resolved \mathcal{CP} -violation study using the existing analysis tools (used by analyses such as [1, 74, 111, 112]) and the data set obtained by my analysis. By feeding the data into the analysis machinery used by the published studies, it can be verified that the data are indeed reconstructed correctly. Furthermore, because the \mathcal{CP} -violation PDF is symmetric around $\Delta t = 0$ and not formed by combining pieces of subfunctions, it is much less difficult to model it algorithmically. Hence, reimplementing the \mathcal{CP} -fit PDF is a good starting point for the development of the fit routine eventually used for a \mathcal{T} -violation fit.

Appendices

A References

- [1] I. ADACHI *et al.* (The Belle collaboration). *Precise measurement of the \mathcal{CP} violation parameter $\sin(2\phi_1)$ in $B^0 \rightarrow (c\bar{c})K^0$ decays.* Phys. Rev. Lett. **108**, 171802 (2012). ↑ 1, 49, 59.
- [2] K. ABE *et al.* (The Belle collaboration). *Observation of Large \mathcal{CP} Violation in the Neutral B Meson System.* Phys. Rev. Lett. **87**, 091802 (2001). ↑ 1, 9.
- [3] B. AUBERT *et al.* (The BABAR collaboration). *Measurement of time-dependent \mathcal{CP} asymmetry in $B^0 \rightarrow c\bar{c}K^{(*)0}$ decays.* Phys. Rev. D **79**, 072009 (2009). ↑ 1, 12, 49.
- [4] B. AUBERT *et al.* (The BABAR collaboration). *Observation of \mathcal{CP} Violation in the B^0 Meson System.* Phys. Rev. Lett. **87**, 091801 (2001). ↑ 1, 9.
- [5] J. P. LEES *et al.* (The BABAR collaboration). *Observation of Time-Reversal Violation in the B^0 Meson System.* Phys. Rev. Lett. **109**, 211801 (2012). ↑ 1, 12, 15, 18, 52, 57.
- [6] J. BERINGER *et al.* *Review of Particle Physics.* Phys. Rev. D **86**, 010001 (2012). ↑ 2, 3, 7–9, 12, 13, 19, 20, 34.
- [7] M. THOMSON. *Modern Particle Physics.* Cambridge University Press, 2013. ISBN: 9781107292543. ↑ 3–5.
- [8] D. GRIFFITHS. *Introduction to Elementary Particles.* Wiley, 2008. ISBN: 9783527618477. ↑ 3, 5.
- [9] M. E. PESKIN and D. V. SCHROEDER. *An Introduction To Quantum Field Theory.* Westview Press, 1995. ISBN: 0201503972. ↑ 3, 4, 7.
- [10] WIKIPEDIA. *Standard Model — Wikipedia, The Free Encyclopedia.* http://en.wikipedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg (visited on 02/09/2014). ↑ 3.
- [11] P. W. HIGGS. *Broken Symmetries and the Masses of Gauge Bosons.* Phys. Rev. Lett. **13** (16: 1964), p. 508. ↑ 4.
- [12] F. ENGLERT and R. BROUT. *Broken Symmetry and the Mass of Gauge Vector Mesons.* Phys. Rev. Lett. **13** (9: 1964), p. 321. ↑ 4.
- [13] G. S. GURALNIK, C. R. HAGEN and T. W. B. KIBBLE. *Global Conservation Laws and Massless Particles.* Phys. Rev. Lett. **13** (20: 1964), p. 585. ↑ 4.
- [14] The CMS collaboration. *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC.* Phys. Lett. B **716** (2012), p. 30. ↑ 5.
- [15] The ATLAS collaboration. *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC.* Phys. Lett. B **716** (2012), p. 1. ↑ 5.
- [16] S. L. GLASHOW. *Partial-symmetries of weak interactions.* Nucl. Phys. **22** (4: 1961), p. 579. ↑ 6.
- [17] A. SALAM and J. WARD. *Electromagnetic and weak interactions.* Phys. Lett. **13** (2: 1964), p. 168. ↑ 6.
- [18] S. WEINBERG. *A Model of Leptons.* Phys. Rev. Lett. **19** (21: 1967), p. 1264. ↑ 6.

- [19] G. ARNISON *et al.* (The UA1 collaboration). *Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s} = 540$ GeV.* Phys. Lett. B **122** (1: 1983), p. 103. [↑ 6.](#)
- [20] M. BANNER *et al.* (The UA2 collaboration). *Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the CERN pp collider.* Phys. Lett. B **122** (5: 1983), p. 476. [↑ 6.](#)
- [21] G. ARNISON *et al.* (The UA1 collaboration). *Experimental observation of lepton pairs of invariant mass around $95 \text{ GeV}/c^2$ at the CERN SPS collider.* Phys. Lett. B **126** (5: 1983), p. 398. [↑ 6.](#)
- [22] P. BAGNAIA *et al.* (The UA2 collaboration). *Evidence for $Z^0 \rightarrow e^+e^-$ at the CERN pp Collider.* Phys. Lett. B **129** (1983), p. 130. [↑ 6.](#)
- [23] THE LEP ELECTROWEAK WORKING GROUP (The ALEPH, DELPHI, L3, and OPAL collaborations). *Electroweak Measurements in Electron-Positron Collisions at W-Boson-Pair Energies at LEP.* Phys. Rep. **532** (2013), p. 119. [↑ 6.](#)
- [24] THE LEP ELECTROWEAK WORKING GROUP, SLD ELECTROWEAK GROUP, AND SLD HEAVY FLAVOUR GROUP (The ALEPH, DELPHI, L3, OPAL, and SLD collaborations). *Precision electroweak measurements on the Z resonance.* Phys. Rep. **427** (2006), p. 257. [↑ 6.](#)
- [25] N. CABIBBO. *Unitary Symmetry and Leptonic Decays.* Phys. Rev. Lett. **10** (12: 1963), p. 531. [↑ 6.](#)
- [26] M. KOBAYASHI and T. MASKAWA. *CP-Violation in the Renormalizable Theory of Weak Interaction.* Progr. Theor. Phys. **49** (2: 1973), p. 652. [↑ 6.](#)
- [27] S. W. HERB *et al.* *Observation of a Dimuon Resonance at 9.5 GeV in 400 GeV Proton-Nucleus Collisions.* Phys. Rev. Lett. **39** (5: 1977), p. 252. [↑ 6.](#)
- [28] F. ABE *et al.* (The CDF collaboration). *Observation of Top Quark Production in $\bar{p}p$ Collisions with the Collider Detector at Fermilab.* Phys. Rev. Lett. **74** (14: 1995), p. 2626. [↑ 6.](#)
- [29] S. ABACHI *et al.* (The DØ collaboration). *Search for High Mass Top Quark Production in $\bar{p}p$ Collisions at $\sqrt{s} = 1.8$ TeV.* Phys. Rev. Lett. **74** (13: 1995), p. 2422. [↑ 6.](#)
- [30] L. WOLFENSTEIN. *Parametrization of the Kobayashi-Maskawa Matrix.* Phys. Rev. Lett. **51** (21: 1983), p. 1945. [↑ 7.](#)
- [31] R. D. PECCEI. *Discrete and global symmetries in particle physics.* In: *Lecture Notes in Physics, Berlin Springer Verlag.* Ed. by L. MATHELITSCH and W. PLESSAS. **521**. 1999. [↑ 7–10.](#)
- [32] R. G. SACHS. *The Physics of Time Reversal.* University Of Chicago Press, 1987. ISBN: 0226733319. [↑ 7, 9, 10.](#)
- [33] G. LÜDERS. *On the Equivalence of Invariance under Time Reversal and under Particle-Antiparticle Conjugation for Relativistic Field Theories.* Dan. Mat. Fys. Medd. **28** (5: 1954). [↑ 8.](#)
- [34] W. PAULI, L. ROSENFELD and V. WEISSKOPF, Eds. *Niels Bohr and the Development of Physics: Essays Dedicated to Niels Bohr on the Occasion of His Seventieth Birthday.* McGraw-Hill, 1955. [↑ 8.](#)
- [35] O. W. GREENBERG. *CPT Violation Implies Violation of Lorentz Invariance.* Phys. Rev. Lett. **89**, 231602 (2002). [↑ 8.](#)
- [36] The ATLAS collaboration. *Measurement of the mass difference between top and anti-top quarks in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector.* Phys. Lett. B **728** (2014), p. 363. [↑ 8, 14.](#)
- [37] The CMS collaboration. *Measurement of the mass difference between top and antitop quarks.* J. High Energy Phys. **6** (2012), p. 109. [↑ 8, 14.](#)
- [38] I. BIGI and A. SANDA. *CP Violation.* Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology. Cambridge University Press, 2009. ISBN: 9781139478243. [↑ 8.](#)
- [39] T. D. LEE and C. N. YANG. *Question of Parity Conservation in Weak Interactions.* Phys. Rev. **104** (1: 1956), p. 254. [↑ 8.](#)

- [40] C. S. WU *et al.* *Experimental Test of Parity Conservation in Beta Decay*. Phys. Rev. **105** (4: 1957), p. 1413. ↑ 8.
- [41] R. L. GARWIN, L. M. LEDERMAN and M. WEINRICH. *Observations of the failure of conservation of parity and charge conjugation in meson decays: the magnetic moment of the free muon*. Phys. Rev. **105** (4: 1957), p. 1415. ↑ 8.
- [42] J. H. CHRISTENSON *et al.* *Evidence for the 2π Decay of the K_2^0 Meson*. Phys. Rev. Lett. **13** (4: 1964), p. 138. ↑ 8, 9.
- [43] V. FANTI *et al.* (The NA48 collaboration). *A new measurement of direct \mathcal{CP} violation in two pion decays of the neutral kaon*. Phys. Lett. B **465** (1999), p. 335. ↑ 8, 9.
- [44] A. ALAVI-HARATI *et al.* (The KTeV collaboration). *Observation of Direct \mathcal{CP} Violation in $K_{S,L} \rightarrow \pi\pi$ Decays*. Phys. Rev. Lett. **83** (1: 1999), p. 22. ↑ 8, 9.
- [45] F. LOW. *Symmetries and elementary particles*. Documents on modern physics. GordonBreach, 1967. ↑ 9.
- [46] H. R. QUINN. *Time reversal violation*. J. Phys: Conf. Ser. **171**, 012001 (2009). ↑ 10, 12.
- [47] J. BERNABEU, F. MARTINEZ-VIDAL and P. VILLANUEVA-PEREZ. *Time Reversal Violation from the entangled $B^0\bar{B}^0$ system*. J. High Energy Phys. **1208** (2012), p. 064. ↑ 10–15, 17, 18, 52, 67.
- [48] E. ABOUZOID *et al.* (The KTeV collaboration). *Precise measurements of direct \mathcal{CP} violation, \mathcal{CPT} symmetry, and other parameters in the neutral kaon system*. Phys. Rev. D **83**, 092001 (2011). ↑ 10.
- [49] A. ANGELOPOULOS *et al.* (The CPLEAR collaboration). *\mathcal{T} -violation and \mathcal{CPT} -invariance measurements in the CPLEAR experiment: a detailed description of the analysis of neutral-kaon decays to $e\pi\nu$* . EPJ C **22** (2001), p. 55. ↑ 10.
- [50] A. ANGELOPOULOS *et al.* (The CPLEAR collaboration). *First direct observation of time reversal noninvariance in the neutral kaon system*. Phys. Lett. B **444** (1998), p. 43. ↑ 10, 12.
- [51] G. D'AMBROSIO and G. ISIDORI (The KLOE collaboration). *Determination of \mathcal{CP} and \mathcal{CPT} violation parameters in the neutral kaon system using the BELL-STEINBERGER relation and data from the KLOE experiment*. J. High Energy Phys. **12** (2006), p. 11. ↑ 10.
- [52] B. AUBERT *et al.* (The BABAR collaboration). *Limits on the Decay-Rate Difference of Neutral B Mesons and on \mathcal{CP} , \mathcal{T} , and \mathcal{CPT} Violation in $B^0-\bar{B}^0$ Oscillations*. Phys. Rev. Lett. **92**, 181801 (2004). ↑ 10.
- [53] B. AUBERT *et al.* (The BABAR collaboration). *Limits on the decay-rate difference of neutral B mesons and on \mathcal{CP} , \mathcal{T} , and \mathcal{CPT} violation in $B^0-\bar{B}^0$ oscillations*. Phys. Rev. D **70**, 012007 (2004). ↑ 10.
- [54] N. C. HASTINGS *et al.* (The Belle collaboration). *Studies of $B^0-\bar{B}^0$ mixing properties with inclusive dilepton events*. Phys. Rev. D **67**, 052004 (2003). ↑ 10.
- [55] L. WOLFENSTEIN. *Violation of Time Reversal Invariance in K^0 Decays*. Phys. Rev. Lett. **83** (5: 1999), p. 911. ↑ 12.
- [56] H.-J. GERBER. *Evidence for time-reversal violation?* EPJ C **35** (2: 2004), p. 195. ↑ 12.
- [57] M. BAÑULS and J. BERNABÉU. *Studying indirect violation of \mathcal{CP} , \mathcal{T} and \mathcal{CPT} in a B -factory*. Nucl. Phys. B **590** (1: 2000), p. 19. ↑ 12.
- [58] M. BAÑULS and J. BERNABÉU. *\mathcal{CP} , \mathcal{T} and \mathcal{CPT} versus temporal asymmetries for entangled states of the B_d -system*. Phys. Lett. B **464** (1: 1999), p. 117. ↑ 12.
- [59] L. WOLFENSTEIN. *The search for direct evidence for time reversal violation*. Int. J. Mod. Phys. E **08** (06: 1999), p. 501. ↑ 12.
- [60] A. EINSTEIN, B. PODOLSKY and N. ROSEN. *Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?* Phys. Rev. **47** (10: 1935), p. 777. ↑ 13.

- [61] K. ABE *et al.* (The Belle collaboration). *Observation of Mixing-induced CP Violation in the Neutral B Meson System*. Phys. Rev. D **66**, 032007 (2002). ↑ 18.
- [62] S. KUROKAWA and E. KIKUTANI (The Belle collaboration). *Overview of the KEKB Accelerators*. Nucl. Instr. Meth. Phys. Res. A **499** (1: 2003), p. 1. ↑ 19.
- [63] T. ABE *et al.* (The Belle collaboration). *Achievements of KEKB*. Progr. Theor. Exp. Phys. (3: 2013). ↑ 19.
- [64] K.-E. B. KENKYUJO (The Belle collaboration). *KEKB B-factory design report*. National Laboratory for High Energy Physics Ibaraki-ken, 1995. ↑ 19.
- [65] K. AKAI *et al.* (The Belle collaboration). *Commissioning of KEKB*. Nucl. Instr. Meth. Phys. Res. A **499** (1: 2003), p. 191. ↑ 19.
- [66] A. ABASHIAN *et al.* (The Belle collaboration). *The Belle Detector*. Nucl. Instr. Meth. Phys. Res. A **479** (1: 2002), p. 117. ↑ 19, 21–25, 70.
- [67] B. CASEY (The Belle collaboration). *HadronB*. Belle Note #390. KEK, 2001. http://belle.kek.jp/secured/belle_note/gn390/bn390_012901.ps.gz. ↑ 20.
- [68] D. BESSON *et al.* *Observation of New Structure in the e^+e^- Cross Section above the $\Upsilon(4S)$* . Phys. Rev. Lett. **54** (5: 1985), p. 381. ↑ 20.
- [69] R. GILES *et al.* *Total cross section for electron-positron annihilation into hadron final states in the upsilon energy region*. Phys. Rev. D **29** (7: 1984), p. 1285. ↑ 20.
- [70] E. RICE *et al.* *Search for Structure in $\sigma(e^+e^- \rightarrow \text{hadrons})$ between $\sqrt{s} = 10.34$ and 11.65 GeV*. Phys. Rev. Lett. **48** (14: 1982), p. 906. ↑ 20.
- [71] V. PAVLUNIN. *Evidence for the B_s meson in the $\Upsilon(5S)$ data*. J. Phys: Conf. Ser. **9** (1: 2005), p. 87. ↑ 20.
- [72] V. EZHELA, S. LUGOVSKY and O. ZENIN. *Hadronic Part of the muon $g-2$ Estimated on the $\sigma_{total}^{2003}(e^+e^- \rightarrow \text{hadrons})$ Evaluated Data Compilation*. IHEP-2003-35. 2003. arXiv: 0312114 [hep-ph]. ↑ 20.
- [73] M. ARINAGA *et al.* (The Belle collaboration). *Progress in KEKB Beam Instrumentation Systems*. Progr. Theor. Exp. Phys. (3: 2013). ↑ 20.
- [74] J. BRODZICKA *et al.* (The Belle collaboration). *Physics achievements from the Belle experiment*. Progr. Theor. Exp. Phys. (1: 2012). ↑ 21, 49, 59.
- [75] Z. NATKANIEC *et al.* (The Belle collaboration). *Status of the Belle Silicon Vertex Detector*. Nucl. Instr. Meth. Phys. Res. A **560** (1: 2006), p. 1. ↑ 22.
- [76] H. AIHARA *et al.* (The Belle collaboration). *Belle SVD2 vertex detector*. Nucl. Instr. Meth. Phys. Res. A **568** (1: 2006), p. 269. ↑ 22, 25.
- [77] H. AIHARA (The Belle collaboration). *The Belle Silicon Vertex Detector*. Nucl. Instr. Meth. Phys. Res. A **466** (2001), p. 268. ↑ 22.
- [78] H. HIRANO *et al.* (The Belle collaboration). *A high-resolution cylindrical drift chamber for the KEK B-factory*. Nucl. Instr. Meth. Phys. Res. A **455** (2: 2000), p. 294. ↑ 23.
- [79] M. AKATSU *et al.* (The Belle collaboration). *Cathode image readout in the Belle central drift chamber*. Nucl. Instr. Meth. Phys. Res. A **454** (2: 2000), p. 322. ↑ 23.
- [80] T. IJIMA *et al.* (The Belle collaboration). *Aerogel Cherenkov Counter for the Belle Detector*. Nucl. Instr. Meth. Phys. Res. A **453** (1: 2000), p. 321. ↑ 23.
- [81] T. SUMIYOSHI *et al.* (The Belle collaboration). *Silica aerogel Cherenkov counter for the KEK B-factory experiment*. Nucl. Instr. Meth. Phys. Res. A **433** (1999), p. 385. ↑ 23.
- [82] E. NAKANO (The Belle collaboration). *Belle PID*. Nucl. Instr. Meth. Phys. Res. A **494** (1: 2002), p. 402. ↑ 23, 28–30, 37.
- [83] H. KICHIMI *et al.* (The Belle collaboration). *The Belle TOF system*. Nucl. Instr. Meth. Phys. Res. A **453** (1: 2000), p. 315. ↑ 23.

- [84] K. MIYABAYASHI (The Belle collaboration). *Belle Electromagnetic Calorimeter*. Nucl. Instr. Meth. Phys. Res. A **494** (1: 2002), p. 298. ↑ 24.
- [85] H. IKEDA *et al.* (The Belle collaboration). *A detailed test of the CsI(Tl) calorimeter for Belle with photon beams of energy between 20 MeV and 5.4 GeV*. Nucl. Instr. Meth. Phys. Res. A **441** (3: 2000), p. 401. ↑ 24.
- [86] A. ABASHIAN *et al.* (The Belle collaboration). *The K_L/μ detector subsystem for the Belle experiment at the KEK B-factory*. Nucl. Instr. Meth. Phys. Res. A **449** (1: 2000), p. 112. ↑ 24.
- [87] A. ABASHIAN *et al.* (The Belle collaboration). *Muon identification in the Belle experiment at KEKB*. Nucl. Instr. Meth. Phys. Res. A **491** (1: 2002), p. 69. ↑ 24, 29, 37.
- [88] D. J. LANGE. *The EvtGen particle decay simulation package*. Nucl. Instr. Meth. Phys. Res. A **462** (1: 2001), p. 152. ↑ 25.
- [89] R. BRUN *et al.* *GEANT 3.21* (1987). CERN Report DD/EE/84-1. ↑ 25.
- [90] R. BRUN and F. CARMINATI. *CERN Program Library Long Writeup W5013*. (unpublished). 1993. <http://wwwasdoc.web.cern.ch/wwwasdoc/pdfdir/geant.pdf>. ↑ 25.
- [91] R. ITOH (The Belle collaboration). *BASF User's Manual. An introduction to the Belle data analysis system*. Belle Note #161. KEK, 1996. http://belle.kek.jp/secured/belle_note/gn161/basf.ps.gz. ↑ 27.
- [92] S. NAGAYAMA (The Belle collaboration). *PANTHER – Reference Manual*. Belle Note #131. KEK, 1996. http://belle.kek.jp/secured/belle_note/gn131/panth2.ps.gz. ↑ 27.
- [93] S. NAGAYAMA (The Belle collaboration). *PANTHER – User's Manual*. Belle Note #130. KEK, 1996. http://belle.kek.jp/secured/belle_note/gn130/panth1.ps.gz. ↑ 27.
- [94] *ROOT: A Data Analysis Framework*. <http://root.cern.ch/>. ↑ 27.
- [95] I. ANTICHEVA *et al.* *ROOT — A C++ framework for petabyte data storage, statistical analysis and visualization*. Comput. Phys. Commun. **180** (12: 2009), p. 2499. ↑ 27.
- [96] R. BRUN and F. RADEMAKERS. *ROOT — An object oriented data analysis framework*. Nucl. Instr. Meth. Phys. Res. A **389** (1: 1997), p. 81. ↑ 27.
- [97] K. HANAGAKI *et al.* (The Belle collaboration). *Electron identification in Belle*. Nucl. Instr. Meth. Phys. Res. A **485** (2002), p. 490. ↑ 28, 37.
- [98] L. PIILONEN *et al.* (The Belle collaboration). *BELLE Muon Identification*. Belle Note #338. KEK, 2000. http://belle.kek.jp/secured/belle_note/gn338/bellenote.ps.gz. ↑ 28, 29, 37.
- [99] F. FANG (The Belle collaboration). *Study of $K_S \rightarrow \pi^+\pi^-$ Selection*. Belle Note #323. KEK, 2000. http://belle.kek.jp/secured/belle_note/gn323/note323.ps.gz. ↑ 30.
- [100] K. HARA. *Flavor Tagging at KEK B-Factory*. Ph.D. Thesis. Osaka University, 2000. <http://osksn2.hep.sci.osaka-u.ac.jp/theses/master/1999/koji.ps>. ↑ 35.
- [101] H. KAKUNO *et al.* (The Belle collaboration). *Neutral flavor tagging for the measurement of mixing-induced violation at Belle*. Nucl. Instr. Meth. Phys. Res. A **533** (3: 2004), p. 516. ↑ 35, 36, 58.
- [102] O. BEHNKE *et al.* *Data Analysis in High Energy Physics: A Practical Guide to Statistical Methods*. Wiley, 2013. ISBN: 9783527653447. ↑ 38–42.
- [103] A. HOECKER *et al.* *TMVA — Toolkit for Multivariate Data Analysis*. 2007. arXiv: 0703039 [physics.data-an]. ↑ 40, 42.
- [104] *TMVA: Toolkit for Multivariate Data Analysis with ROOT*. <http://tmva.sourceforge.net/>. ↑ 40.
- [105] P. SPECKMAYER *et al.* *The toolkit for multivariate data analysis, TMVA 4*. J. Phys: Conf. Ser. **219**, 032057 (2010). ↑ 40.
- [106] A. P. BRADLEY. *The Use of the Area Under the ROC Curve in the Evaluation of Machine Learning Algorithms*. Pattern Recogn. **30** (7: 1997), p. 1145. ↑ 42.

- [107] G. C. FOX and S. WOLFRAM. *Event shapes in e^+e^- annihilation*. Nucl. Phys. B **149** (3: 1979), p. 413. ↑ 46.
- [108] G. C. FOX and S. WOLFRAM. *Observables for the Analysis of Event Shapes in e^+e^- Annihilation and Other Processes*. Phys. Rev. Lett. **41** (23: 1978), p. 1581. ↑ 46.
- [109] E. FARHI. *Quantum Chromodynamics Test for Jets*. Phys. Rev. Lett. **39** (25: 1977), p. 1587. ↑ 46.
- [110] H. TAJIMA *et al.* (The Belle collaboration). *Proper-time resolution function for measurement of time evolution of B mesons at the KEK B-Factory*. Nucl. Instr. Meth. Phys. Res. A **533** (3: 2004), p. 370. ↑ 49.
- [111] K.-F. CHEN *et al.* (The Belle collaboration). *Observation of time-dependent \mathcal{CP} violation in $B^0 \rightarrow \eta' K^0$ decays and improved measurements of \mathcal{CP} asymmetries in $B^0 \rightarrow \phi K^0$, $K_S^0 K_S^0 K_S^0$ and $B^0 \rightarrow J/\Psi K^0$ decays*. Phys. Rev. Lett. **98**, 031802 (2007). ↑ 49, 59.
- [112] H. SAHOO *et al.* (The Belle collaboration). *Measurements of time-dependent \mathcal{CP} violation in $B^0 \rightarrow \Psi(2S) K_S^0$ decays*. Phys. Rev. D **77**, 091103 (2008). ↑ 49, 59.
- [113] W. VERKERKE and D. KIRKBY. *The RooFit toolkit for data modeling*. CHEP-2003-MOLT007. 2003. arXiv: 0306116 [physics]. ↑ 49, 51.
- [114] L. MONETA *et al.* *The RooStats project*. 2010. arXiv: 1009.1003 [physics.data-an]. ↑ 49, 51.

B Comparison Pairs

Reference		\mathcal{T} -conjugate	
Transition	Final state	Transition	Final state
$\bar{B}^0 \rightarrow B_-$	$(l^+ X, J/\Psi K_S)$	$B_- \rightarrow \bar{B}^0$	$(J/\Psi K_L, l^- X)$
$B_+ \rightarrow B^0$	$(J/\Psi K_S, l^+ X)$	$B^0 \rightarrow B_+$	$(l^- X, J/\Psi K_L)$
$\bar{B}^0 \rightarrow B_+$	$(l^+ X, J/\Psi K_L)$	$B_+ \rightarrow \bar{B}^0$	$(J/\Psi K_S, l^- X)$
$B_- \rightarrow B^0$	$(J/\Psi K_L, l^+ X)$	$B^0 \rightarrow B_-$	$(l^- X, J/\Psi K_S)$

(a) Comparison pairs for \mathcal{T} -violation measurement.

Reference		\mathcal{T} -conjugate	
Transition	Final state	Transition	Final state
$\bar{B}^0 \rightarrow B_-$	$(l^+ X, J/\Psi K_S)$	$B^0 \rightarrow B_-$	$(l^- X, J/\Psi K_S)$
$B_+ \rightarrow B^0$	$(J/\Psi K_S, l^+ X)$	$B_+ \rightarrow \bar{B}^0$	$(J/\Psi K_S, l^- X)$
$\bar{B}^0 \rightarrow B_+$	$(l^+ X, J/\Psi K_L)$	$B^0 \rightarrow B_+$	$(l^- X, J/\Psi K_L)$
$B_- \rightarrow B^0$	$(J/\Psi K_L, l^+ X)$	$B_- \rightarrow \bar{B}^0$	$(J/\Psi K_L, l^- X)$

(b) Comparison pairs for \mathcal{CP} -violation measurement.

Reference		\mathcal{T} -conjugate	
Transition	Final state	Transition	Final state
$\bar{B}^0 \rightarrow B_-$	$(l^+ X, J/\Psi K_S)$	$B_- \rightarrow B^0$	$(J/\Psi K_L, l^+ X)$
$B_+ \rightarrow B^0$	$(J/\Psi K_S, l^+ X)$	$\bar{B}^0 \rightarrow B_+$	$(l^+ X, J/\Psi K_L)$
$B^0 \rightarrow B_-$	$(l^- X, J/\Psi K_S)$	$B_- \rightarrow \bar{B}^0$	$(J/\Psi K_L, l^- X)$
$B_+ \rightarrow \bar{B}^0$	$(J/\Psi K_S, l^- X)$	$B^0 \rightarrow B_+$	$(l^- X, J/\Psi K_L)$

(c) Comparison pairs for $\mathcal{CP}\mathcal{T}$ -violation measurement.

Tab. 12: Comparison pairs for use in analyses of fundamental discrete symmetries [47].

C Asymmetry Parameters

For \mathcal{T}		For \mathcal{CP}		For $\mathcal{CP}\mathcal{T}$	
$\Delta S_{\mathcal{T}}^+$	$:= S_{l^-, K_L}^- - S_{l^+, K_S}^+$	$\Delta S_{\mathcal{CP}}^+$	$:= S_{l^-, K_S}^+ - S_{l^+, K_S}^+$	$\Delta S_{\mathcal{CP}\mathcal{T}}^+$	$:= S_{l^+, K_L}^- - S_{l^+, K_S}^+$
$\Delta S_{\mathcal{T}}^-$	$:= S_{l^-, K_L}^+ - S_{l^+, K_S}^-$	$\Delta S_{\mathcal{CP}}^-$	$:= S_{l^-, K_S}^- - S_{l^+, K_S}^-$	$\Delta S_{\mathcal{CP}\mathcal{T}}^-$	$:= S_{l^+, K_L}^+ - S_{l^+, K_S}^-$
$\Delta C_{\mathcal{T}}^+$	$:= C_{l^-, K_L}^- - C_{l^+, K_S}^+$	$\Delta C_{\mathcal{CP}}^+$	$:= C_{l^-, K_S}^+ - C_{l^+, K_S}^+$	$\Delta C_{\mathcal{CP}\mathcal{T}}^+$	$:= C_{l^+, K_L}^- - C_{l^+, K_S}^+$
$\Delta C_{\mathcal{T}}^-$	$:= C_{l^-, K_L}^+ - C_{l^+, K_S}^-$	$\Delta C_{\mathcal{CP}}^-$	$:= C_{l^-, K_S}^- - C_{l^+, K_S}^-$	$\Delta C_{\mathcal{CP}\mathcal{T}}^-$	$:= C_{l^+, K_L}^+ - C_{l^+, K_S}^-$

Tab. 13: Asymmetry parameters to be used for evaluating \mathcal{T} -, \mathcal{CP} -, or $\mathcal{CP}\mathcal{T}$ -violation [47].

D List of Cuts

Variable	Cut Value	Description
π_e	> 0.6	The electron identification probability, i.e. the output of the electron identification algorithm <code>eid</code> for a given electron candidate using the algorithm parameters (3, -1, 5).
π_μ	> 0.6	The muon identification probability, i.e. the likelihood returned by the muon identification algorithm <code>muid</code> for a given muon candidate.
$ \vec{p}_{J/\Psi}^{[c.m.]} $	$< 2.0 \text{ GeV}$	The magnitude of the centre-of-mass three-momentum of the J/Ψ candidate.
$ \mathbf{P}_{J/\Psi} $		The magnitude of the four-momentum of the J/Ψ (i.e. the reconstructed invariant mass). The cut value depends on the reconstruction channel of the J/Ψ because electrons and muons have different photon radiation characteristics. The recommended cut values are $e\bar{e} : (2.94688, 3.13288) \text{ GeV} .$ $\mu\bar{\mu} : (3.03688, 3.13288) \text{ GeV} .$

Tab. 14: Cut variables used in the J/Ψ reconstruction.

Variable	Cut Value	Description
N_{KLM}		Number of associated hits in the KLM. If there is an associated ECL cluster the cut is performed for $N_{\text{KLM}} > 1$, otherwise for $N_{\text{KLM}} > 2$.
$\Delta\varphi_Q$	$> 5^\circ$	The angular separation between the measured K_L and the closest charged particle.
$\Delta\varphi_{\text{KE}}$	$< 15^\circ$	The angular separation between the three-momenta reconstructed from the KLM and ECL cluster separately. This cut is only applied if there is an associated ECL cluster.
E_{ECL}	$> 160 \text{ GeV}$	The ECL cluster energy; this cut is only applied if there is an associated ECL cluster.

Tab. 15: Cut variables used in the K_L reconstruction.

Variable	Cut Value	Description
$\Delta\varphi_{K_L}$	$< 35^\circ$	The angular separation between measured and expected K_L three-momentum. The expected three-momentum can be calculated by assuming that the B meson is at rest in the centre-of-mass system. Therefore, the J/Ψ and the K_L are approximately back-to-back in the centre-of-mass system. By combining the known centre-of-mass four-vector and the measured J/Ψ four-vector one can infer the expected K_L four-momentum.
$ \vec{p}_{J/\Psi}^{[c.m.]} $	$[1.42, 2.0] \text{ GeV}$	The magnitude of the centre-of-mass three-momentum of the J/Ψ candidate.
$ \mathbf{P}_{B^+} $	$\leq 1.2 \text{ GeV}$	The magnitude of the four-momentum of the B^+ (i.e. the reconstructed invariant mass).
M_{BC}	$\geq 5.14 \text{ GeV}$	The beam constrained mass as defined by: $M_{\text{BC}} := \sqrt{E_{\text{Beam}}^{[c.m.]^2} - p_{B\mathcal{CP}}^{[c.m.]^2}} ,$ where $E_{\text{Beam}}^{[c.m.]}$ is the beam energy in the centre-of-mass frame (i.e. half the centre-of-mass energy) and $p_{B\mathcal{CP}}^{[c.m.]}$ is the momentum of the \mathcal{CP} tagged B meson in the centre-of-mass frame.

Tab. 16: Cut variables used in the B^+ reconstruction.

Variable	Cut Value	Description
$ \vec{p}_{J/\Psi}^{[c.m.]} $	[1.42, 2.0] GeV	The magnitude of the centre-of-mass three-momentum of the J/Ψ candidate.
$ \vec{p}_{B^-}^{[c.m.]} $	[0.2, 0.45] GeV	The magnitude of the centre-of-mass three-momentum of the B^- candidate.

Tab. 17: Cut variables used in the B^- reconstruction.

Variable	Cut Value	Description
π_e	> 0.6	The electron identification probability, i.e. the output of the electron identification algorithm <code>eid</code> for a given electron candidate using the algorithm parameters (3, -1, 5).
π_μ	> 2	The muon identification probability, i.e. the likelihood returned by the muon identification algorithm <code>muid</code> for a given muon candidate.
π_K	> 0.5	The kaon identification probability, i.e. the likelihood returned by the <code>atc_pid</code> identification algorithm for a given kaon candidate using the algorithm parameters (0, 1, 0, 3, 2).

Tab. 18: Cut variables used by the rest charge reconstruction.

E MVA Parameters

Parameter	Setting	Parameter	Setting
SplitMode	Random	BoostType	AdaBoost
NormMode	NumEvents	AdaBoostBeta	0.5
NTrees	300	UseBaggedBoost	true
MinNodeSize	5 %	BaggedSampleFraction	0.6
MaxDepth	3	SeparationType	GiniIndex

Tab. 19: TMVA parameter settings used for training the BDTs. Parameters which are not listed retain their default value (for TMVA version 4.2.0).

F MVA Input Variables

F.1 List of Variables

Variable Symbol	Description
$\cos \theta_{P_{B_{CP}}}$	The cosine of the azimuthal angle of the reconstructed B_{CP} three-momentum in the lab frame.
N_{Charged}	Number of charged particles found by the Belle reconstruction scheme (i.e. number of entries in <code>Mdst_charged</code>).
π_{l_i}	Value of the electron/muon likelihood (i.e. L_{eid} respectively L_{muid}) obtained from the respective identification algorithm for a daughter particle of the J/Ψ . The value of the subscript i encodes the charge of the daughter particle: $i = 0$ corresponds to positive charge, $i = 1$ to negative charge.
$\cos \theta_{\text{Thrust}}$	Cosine Theta Thrust, one of the kinematic variables used for continuum suppression, cf. Section 8 in Chapter V for a more detailed explanation.
R_i	The i^{th} normalised FOX-WOLFRAM moment: one of the kinematic variables used for continuum suppression, cf. Section 8 in Chapter V for a more detailed explanation. The MVA uses the second ($i = 2$) and fourth ($i = 4$) normalised FOX-WOLFRAM moment.
ΔE_{CP}	The difference between the expected B_{CP} energy in the centre-of-mass frame $E_{CP, \text{Exp.}}^{[\text{c.m.}]} \equiv \frac{1}{2}E_{\text{Beam}}^{[\text{c.m.}]}$ and the reconstructed value $E_{CP, \text{Rec.}}^{[\text{c.m.}]}$.
$M_{BC, CP}$	The beam constrained mass as defined in Table 16 calculated for the B_{CP} .
E_{Rest}	Sum of the energy of the restcharge particles, cf. Section 6 in Chapter V for a discussion of the restcharges.
χ_{Flv}^2	Chisquare goodness-of-fit value obtained from the B_{Flv} vertex fit.
$p_{B, \text{Meas.}}^{[\text{c.m.}]}$	Reconstructed absolute value of the B_{CP} three-momentum in the centre-of-mass frame.
$p_{B, \text{Est.}}^{[\text{c.m.}]}$	Estimated absolute value of the B_{CP} three-momentum in the centre-of-mass frame. Uses the method explained in Section 4.1 in Chapter V in order to estimate the poorly measured K_L four-momentum.
$p_{J/\Psi}^{[\text{c.m.}]}$	Reconstructed absolute value of the J/Ψ three-momentum in the centre-of-mass frame.
E_{l_i}	Energy of a daughter particle of the J/Ψ ; the value of the subscript i encodes the charge of the daughter particle: $i = 0$ corresponds to positive charge, $i = 1$ to negative charge.
$\sphericalangle(l_0, l_1)$	Angle between the three-momenta of the J/Ψ daughter particles in the lab frame.
$\sphericalangle(J/\Psi, K_{S,L})$	Angle between the three-momenta of the J/Ψ and the $K_{S,L}$ in the lab frame.
$\Delta \varphi_{K_{S,L}}$	Angle between the expected and measured $K_{S,L}$ three-momentum in the lab frame. The calculation of the expected K_L three-momentum has been explained in Section 4 in Chapter V; calculation for the K_S follows the same approach.
$\Delta \varphi_Q$	Angle between the $K_{S,L}$ candidate three-momentum vector and the closest charge tracks. For the K_S this uses the three-momentum direction obtained from combining the four-momenta of the charged pions, for the K_L the direction obtained from the KLM cluster is used.
N_{KLM}	Number of hits in the KLM associated to the K_L candidate.
M_{ECL}	Mass of the ECL cluster associated with the K_L candidate.
W_{ECL}	Width of the ECL cluster that is attributed to the K_L candidate.
e_9/e_{25}	Ratio of the energy deposited in a 3×3 field (e_9) and the energy deposited in a 5×5 field (e_{25}) around the centre of the ECL cluster associated with the K_L candidate. This variable is typically used for electron/photon discrimination [66]. However, it can be equally used to test an ECL cluster's K_L -likeness.
E_{ECL}	The energy of the ECL cluster that is associated with the K_L candidate.
$\Delta \varphi_{\text{KE}}$	Angle between the three-momentum directions in the lab frame reconstructed from the KLM and ECL cluster separately for a given K_L candidate.

Tab. 20: List of input variables used by the different MVAs.

F.2 Distributions for minus MVA

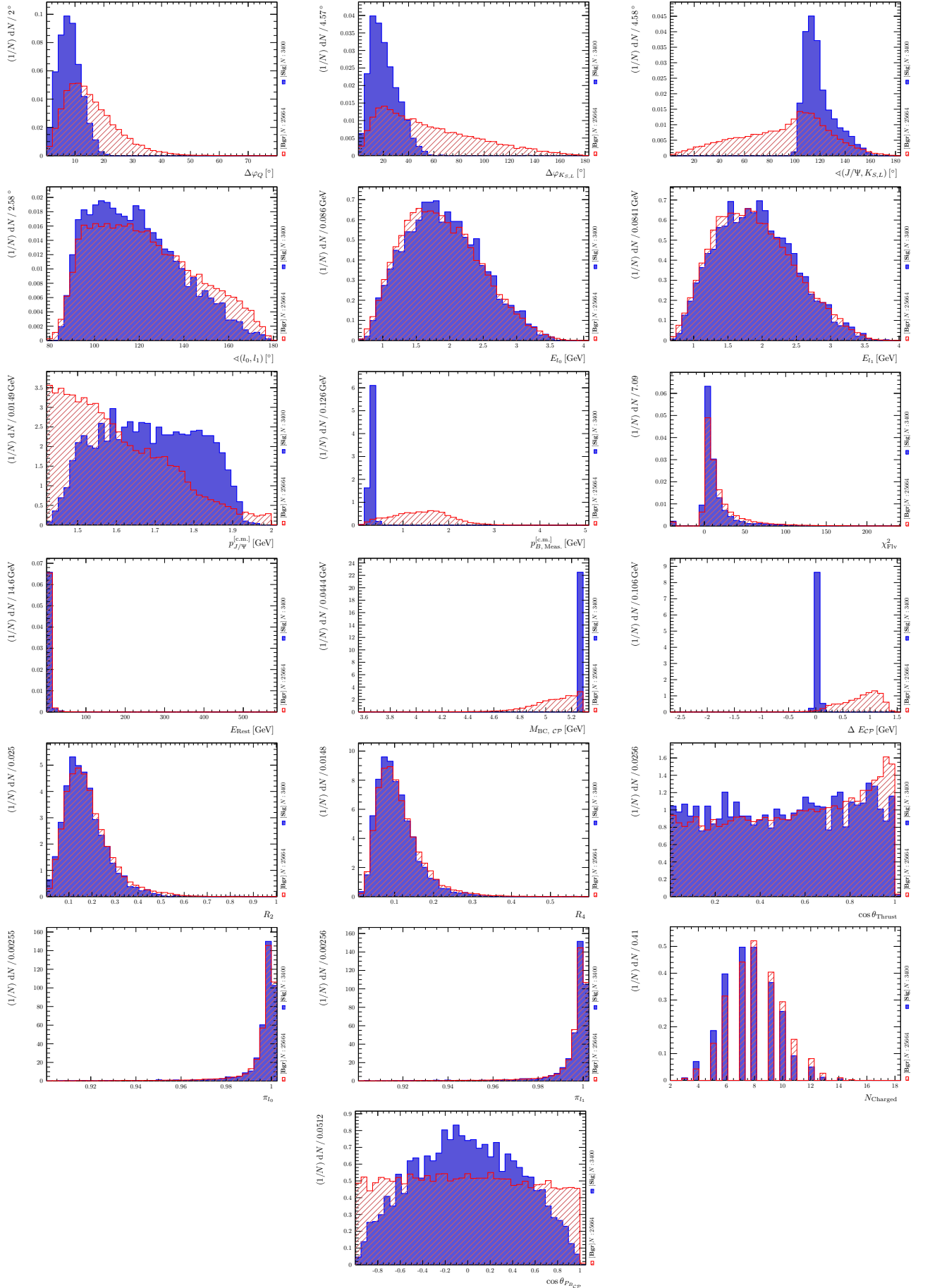


Fig. 38: Input distributions for the variables used by the minus MVA classifier. Shown are the signal and background distributions separately. Refer to Table 20 for a list of the different variables depicted in the above plots.

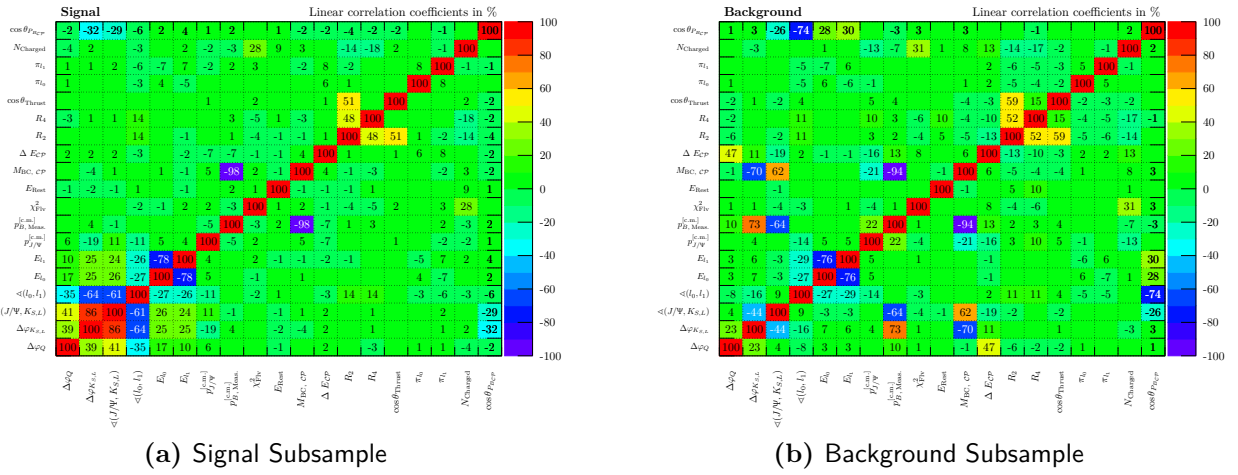


Fig. 39: Plot of linear correlation coefficients of input variables to the minus MVA classifier. Refer to Table 20 for a list of the different variables depicted in the above plots.

F.3 Distributions for plus_KLM MVA

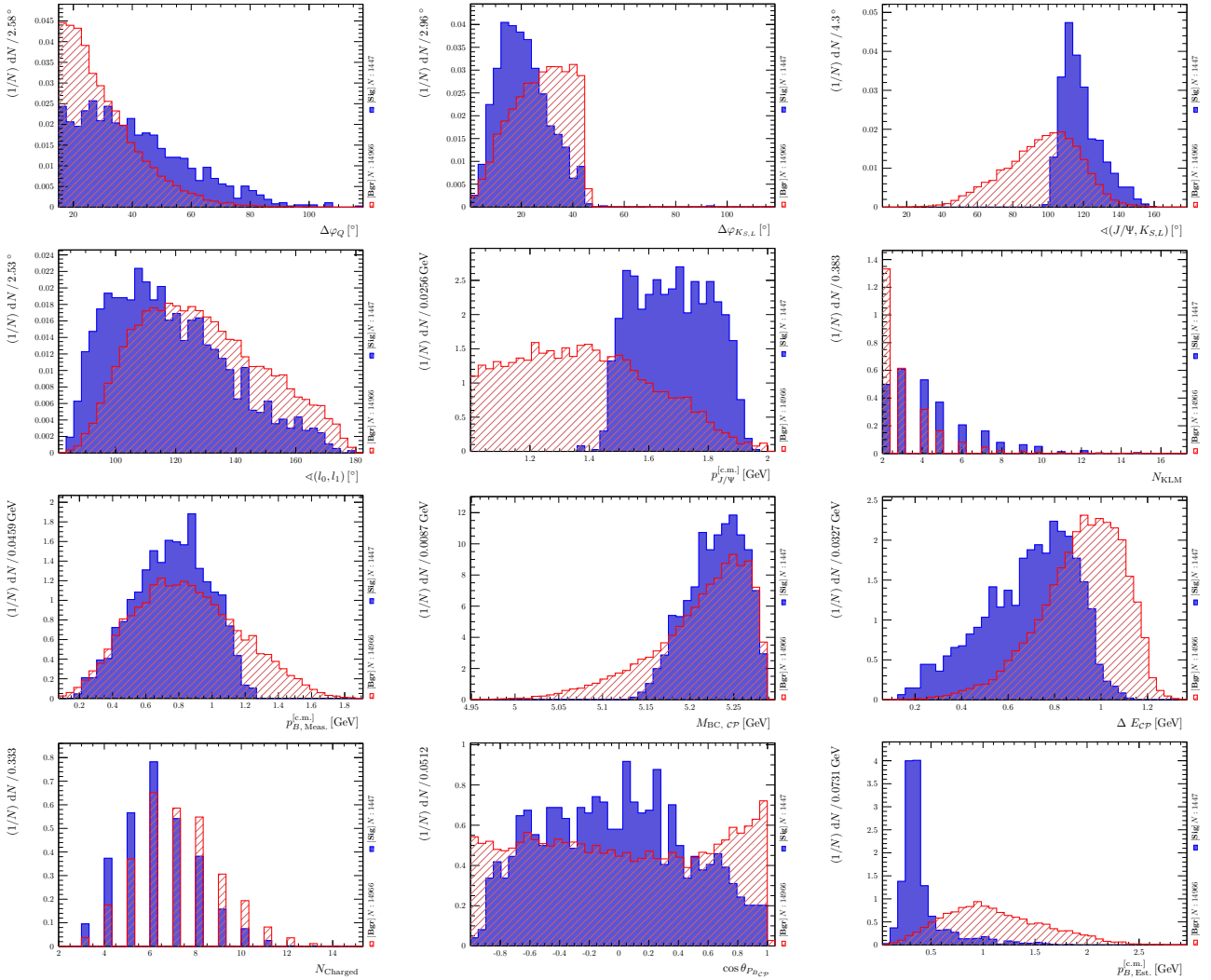


Fig. 40: Input distributions for the variables used by the minus_KLM MVA classifier. Shown are the signal and background distributions separately. Refer to Table 20 for a list of the different variables depicted in the above plots.

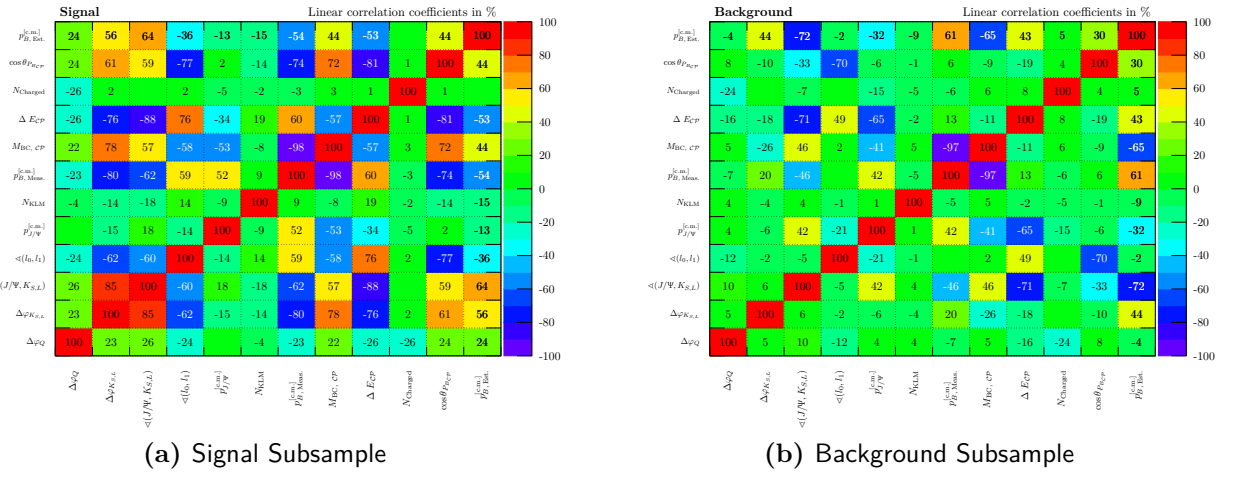
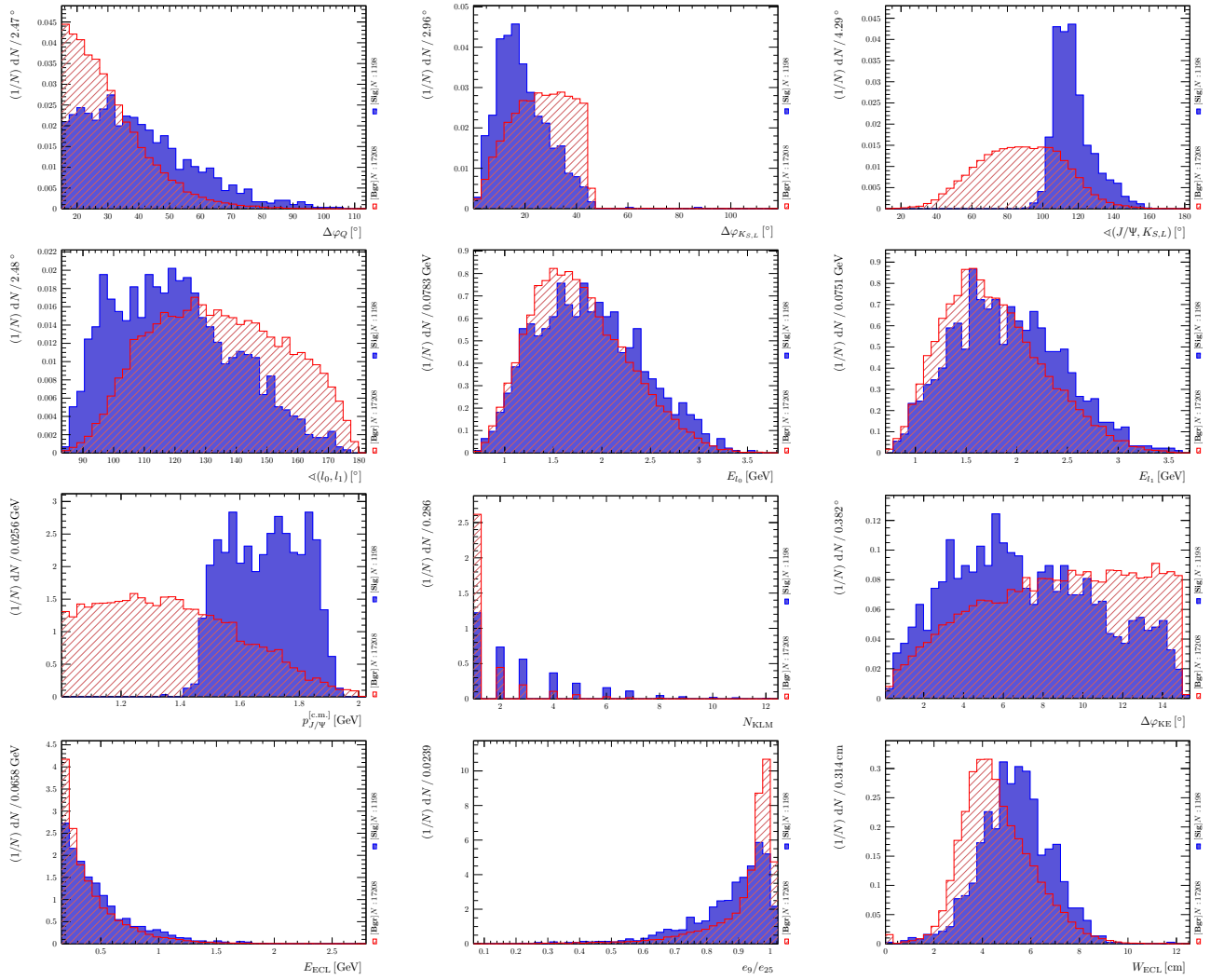


Fig. 41: Plot of linear correlation coefficients of input variables to the plus_KLM MVA classifier. Refer to Table 20 for a list of the different variables depicted in the above plots.

F.4 Distributions for plus_KLM_ECL MVA



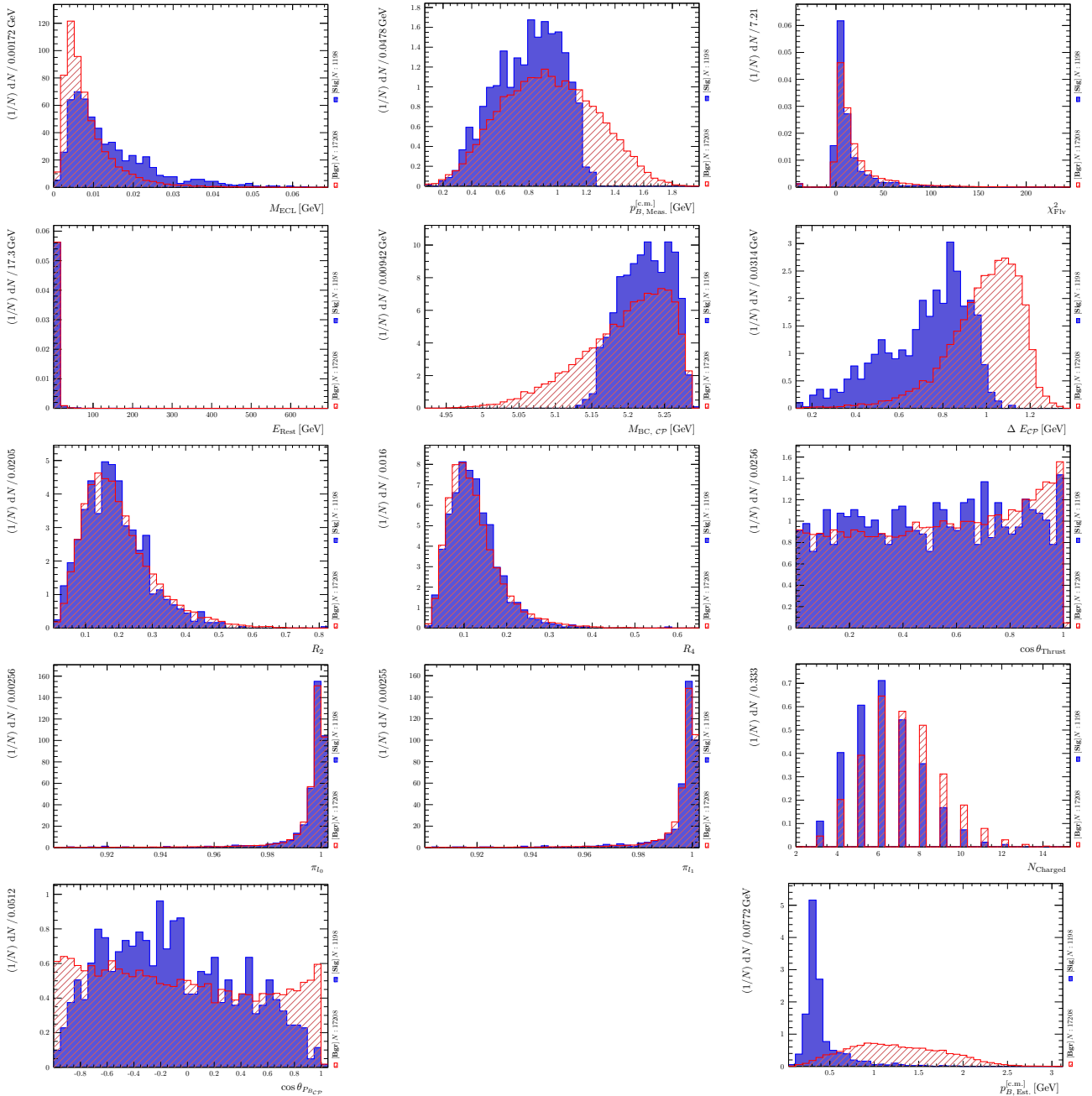
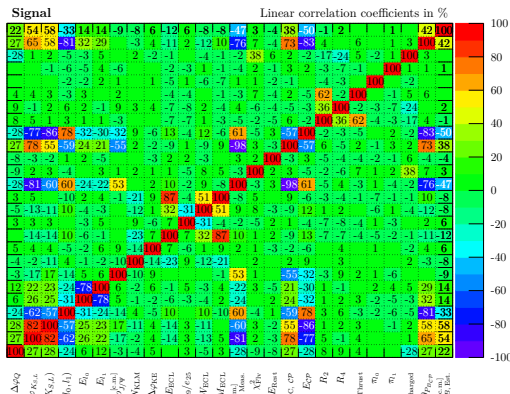
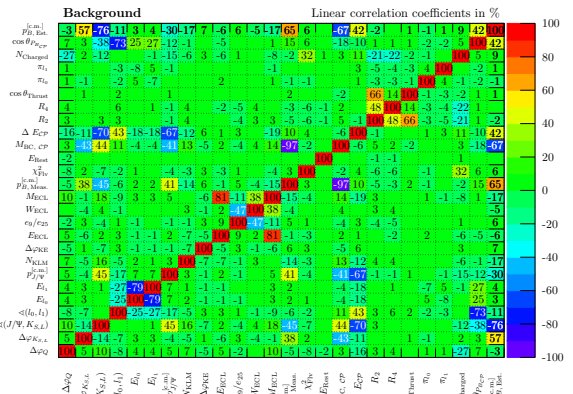


Fig. 42: Input distributions for the variables used by the plus_KLM_ECL MVA classifier. Shown are the signal and background distributions separately. Refer to Table 20 for a list of the different variables depicted in the above plots.



(a) Signal Subsample



(b) Background Subsample

Fig. 43: Plot of linear correlation coefficients of input variables to the plus_KLM_ECL MVA classifier. Refer to Table 20 for a list of the different variables depicted in the above plots.

Acknowledgements

First and foremost I would like to thank Prof. Dr. Ariane Frey for the opportunity to work once again in her group. Our shared interest in the previous \mathcal{T} -violation study, that was performed and published by *BABAR*, enabled me to work on this analysis. I am very grateful for her offer to work on a topic of my own choosing. Over the course of my master's period I have acquired many new skills and improved many old ones, a process in which she has been helpful and supportive.

Furthermore, I would like to thank Prof. Dr. Arnulf Quadt for being the second referee. I would like to use this opportunity to thank him for the support guidance he has offered throughout my master's studies, specifically I would like to thank him for helping with my application to the CERN summer school. Working at the very heart of particle physics has not only been a stimulating and enlightening experience, it was a lot of fun too.

I would like to offer a very special thank you to Dipl. Phys. Philipp Hamer who has, in a way, filled the role of a direct supervisor throughout my thesis. The discussions we engaged in and the input he provided have been prosperous and very helpful to this analysis. Furthermore, I would like to express my gratitude for his offer to proofread my thesis.

It is my pleasure to thank my friend Dominik Müller, who has helped in keeping me sane throughout my master's, especially during those times when "*the fit*" did not converge. I am ever grateful that he enabled me to focus on finalising my thesis by refurbishing our shared flat nearly single-handedly.

Moreover, I would like to thank Sara Gadeberg who has given me stability, sanity, and joy more than anyone else during my master's. Albeit particle physics is a topic which she finds utterly unappealing, she offered to proofread my thesis in all of its plentitude, for which I am very grateful.

Finally, I would like to thank everyone in the working group Frey; it has been a pleasure to work amongst each and everyone of you. I very much enjoyed the warm and welcoming working climate and cherished the input and help offered by my colleagues.

Declaration of Conduct

Erklärung

nach §17(9) der Prüfungs- und Studienordnung 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 nichtbestanden Prüfung an dieser oder einer anderen Hochschule eingereicht wurde.

Göttingen, den 28th October 2014

(Fabian Wilk)