DISSERTATION

Measurement of $|V_{us}|$ using hadronic τ decays at Belle II

ausgeführt an der
Fakultät für Physik der Technischen Universität Wien

in Zusammenarbeit mit dem
Institut für Hochenergiephysik (HEPHY) der Österreichischen Akademie der Wissenschaften (ÖAW)

zur Erlangung des akademischen Grades
Doktor der Technischen Wissenschaften

eingereicht von
Géraldine Räuber, MSc
Matrikelnummer: 12129471

Betreuung : Privatdoz. Dipl.-Ing. Dr. techn. Christoph Schwanda
Mitbetreuung : Univ. Lektor Dr. Gianluca Inguglia

Wien, am 10. Mai 2024

Unterschrift StudentIn

Unterschrift BetreuerIn

to the happy few

Abstract

This thesis presents a precise measurement of the CKM matrix element V_{us} using data collected at the Belle II experiment situated at the SuperKEKB electron-positron collider in Tsukuba, Japan. The measurement employs the ratio of branching fractions of tau lepton decays, denoted as $R_{K/\pi}$, with both pion and kaon mesons in the final state, given by $R_{K/\pi} = \mathscr{B}(\tau^- \to K^- \nu_\tau)/\mathscr{B}(\tau^- \to \pi^- \nu_\tau)$. Selection criteria involve a tag side consisting of three charged particles, defining an overall 3×1-prong topology. The preliminary obtained result, $R_{K/\pi} = 0.06438 \pm 0.00080$, corresponds to a $|V_{us}|$ value of 0.2229 ± 0.0017, and takes into account only the main systematic uncertainty sources. Recognizing the systematic nature of this measurement, a dedicated study addressing leading systematic uncertainties is undertaken. Specifically, a performance assessment of pion identification corrections is conducted to enhance the reliability and precision of the measurement.

Zusammenfassung

Diese Arbeit präsentiert eine präzise Messung des CKM-Matrixelements V_{us} unter Verwendung von Daten, die am Belle II-Experiment am SuperKEKB-Elektron-Positron-Kollider in Tsukuba, Japan, gesammelt wurden. Die Messung verwendet das Verhältnis der Zerfallsverzweigungen von Tau-Leptonen, bezeichnet als $R_{K/\pi}$, wobei sowohl Pionen als auch Kaonen im Endzustand vorhanden sind, gegeben durch $R_{K/\pi} = \mathcal{B}(\tau^- \to K^- \nu_\tau)/\mathcal{B}(\tau^- \to \pi^- \nu_\tau)$. Die Auswahlkriterien umfassen eine Tag-Seite mit drei geladenen Teilchen und definieren insgesamt eine 3×1-Zacken-Topologie. Das vorläufig erhaltene Ergebnis, $R_{K/\pi} = 0.06438 \pm 0.00080$, entspricht einem Wert von $|V_{us}| = 0.2229 \pm 0.0017$ und berücksichtigt nur die Hauptquellen systematischer Unsicherheit. Angesichts des systematischen Charakters dieser Messung wird eine spezielle Studie zur Behandlung der wichtigsten systematischen Unsicherheiten durchgeführt. Insbesondere wird eine Leistungsbeurteilung von Korrekturen zur Pionenidentifikation durchgeführt, um die Zuverlässigkeit und Präzision der Messung zu verbessern.

Contents

Introduction

"Three quarks for Muster Mark!"

- James Joyce, *Finnegans Wake, 1939*

The Standard Model of particle physics stands as one of the most successful frameworks in understanding the fundamental building blocks of the universe and their interactions. It elegantly categorises particles into fermions, which include quarks and leptons, and bosons, responsible for mediating interactions. Within this framework, the interplay of electromagnetism, weak force, and strong force is precisely described, leading to remarkable predictive power and experimental confirmations.

However, despite its successes, the Standard Model is not a complete theory. It leaves significant questions unanswered, such as the nature of dark matter, the origin of neutrino masses, and the unification of forces, including gravity. These limitations underscore the necessity of probing the model's boundaries with precision tests to uncover its deficiencies, and guide us toward more comprehensive theories.

One such precision test involves the Cabibbo-Kobayashi-Maskawa (CKM) matrix, which describes the flavor-changing weak interactions among quarks. The CKM matrix element V_{us} is of particular interest; it is key in understanding the mixing between down-type and strange-type quarks. Various methods to extract the magnitude of this parameter exist, but a particularly promising avenue lies within the ratios of branching fractions of tau decays with pion and kaon mesons in the final state.

This approach is particularly suited to experiments like Belle II, a high-energy physics experiment located in Tsukuba, Japan. With its asymmetric beam energy e^+e^- collider and detector located at the interaction point, it produces and detects a significant quantity of B mesons, aligning with its primary design objective. However, a unique and inherent feature of such "*B-factories*" lies in the fact that they also produce tau leptons in quantities that enable competitive precision measurements.

We embark on this thesis for an exploration of these key elements. Chapter 1 lays the groundwork for understanding fundamental concepts and models used throughout this dissertation. Subsequently, Chapter 2 shifts the subject for a few pages to the Belle II experiment, and details all the relevant aspects, from the SuperKEKB accelerator to tau physics at B-factories, including the sub-detectors composing the Belle II detector, particle identification, trigger systems, and data samples.

Following these two introductory chapters, Chapter 3 discusses in detail a performance study focusing on correction factors attributed to pion identification using tau decays. Finally, Chapter 4 is dedicated to the Belle II measurement of $|V_{us}|$. The work presented in these chapters was done under the supervision of *Gianluca Inguglia*. Additional guidance and assistance was provided by *Michel Hernández Villanueva*. More generally, collaboration within the *lfu-vus* working group contributed greatly to this work; its members are *Gianluca Inguglia*, *Armine Rostomyan*, *Michel Hernández Villanueva*, *Petar Rados*, *Alberto Martini*, *Paul Feichtinger* and *Zuzana Gruberová*.

1. Theoretical Motivation

It is essential to establish the theoretical framework within which the presented measurement is conducted. Although numerous theoretical aspects could be discussed, the focus here is directed towards elements relevant to the presented content, with a little touch of history to set the scene.

1.1 The Standard Model of Particle Physics

In the late 19th century, the landscape of physics underwent a transformation with the discovery of the electron. Joseph John Thomson's experiments [1] revealed the existence of these negatively charged particles, a revelation that challenged the notion of indivisible atoms that was prevalent at that time. The early 20th century brought another major discovery, as Ernest Rutherford's gold foil experiment [2] exposed the proton, the positively charged counterpart to the electron.

As the 20th century progressed, the research intensified, yielding an expanding zoo of particles: Photons [3, 4], neutrons [5], positrons [6], muons [7], and other entities were discovered, creating a disparate set of building blocks. The quest for order encouraged the development of various theoretical frameworks, including quantum electrodynamics (QED) and quantum chromodynamics (QCD). These theories, while significant, were but pieces of a larger puzzle, prompting physicists to seek a unified framework that could encapsulate the diverse particles and their interactions.

All the pieces came together in the 1970s, with the completion of the Standard Model (SM), a revolutionary comprehensive theoretical framework that unified previously disjointed theories into one [8]. The Standard Model can be more accurately described as a relativistic renormalisable quantum field theory based on the principle of gauge invariance that provides a structured representation of the subatomic universe, and describes their interactions through forces. It is often represented by a diagram like the one displayed in Figure 1.1.

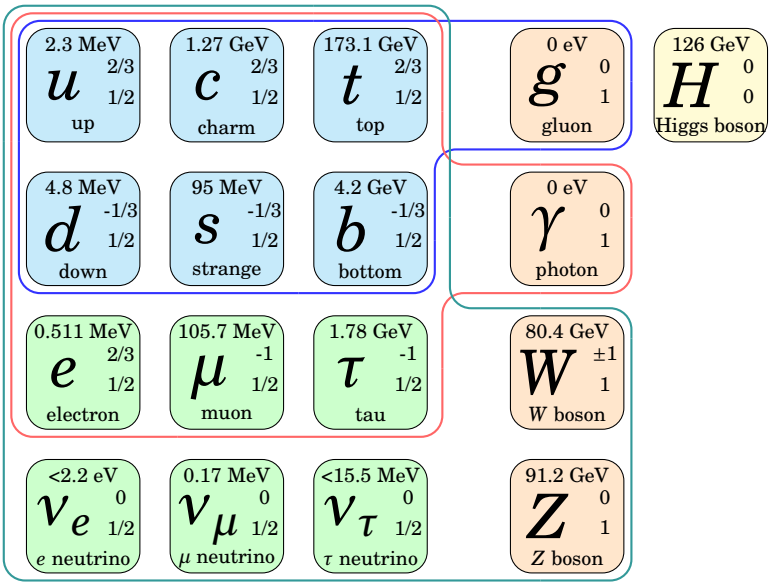


Figure 1.1: *Standard Model (SM) of particle physics.* Quarks are indicated in blue, leptons in blue, gauge bosons in orange, and the Higgs boson in yellow. Their respective mass, charge, spin and name are also mentioned. The additional lines represent the main interactions allowed between gauge bosons and fermions. Other interactions and antiparticles are omitted to simplify the representation and avoid redundancy.

The Standard Model stands as a remarkable achievement in theoretical physics, successfully predicting the existence of particles long before experimental confirmation. Among these predictions were, for example, the top quark and the Higgs boson. The top quark, the heaviest known elementary particle, was theorised by the Standard Model, but was only observed in 1995 at Fermilab's Tevatron [9]. Similarly, the Higgs boson, long sought-after for its role in imparting mass to other particles by the Higgs mechanism, was theorised by the Standard Model and eventually discovered in 2012 at CERN's Large Hadron Collider [10]. While these discoveries validate the Standard Model, others highlight its limitations. These unexplained phenomena, discussed briefly in Section 1.1.2, represent a challenge that physicists strive to solve.

1.1.1 Definition and Development of the Standard Model

A detailed examination of all the components depicted in Figure 1.1 is necessary to truly grasp the depth of the Standard Model, and appreciate its comprehensive nature.

The Standard Model classifies particles into two main categories: bosons, responsible for describing force carriers and characterised by integer spin, and fermions, accounting for the particles that constitute matter and characterised by half-integer spin. Spin, in this context, refers to the intrinsic quantised angular momentum of a particle. Another difference between bosons and fermions, satisfying the principle of causality, lies in their quantisation rules. Bosons are quantised with the commutator relation, while fermions are quantised with the anti-commutator relation.

For every particle in the Standard Model, there exists a corresponding antiparticle with identical mass but opposite charge. Antiparticles obey the same fundamental laws as their particle counterparts. The concept of antiparticles contributes to the overall balance of charge and other quantum numbers in particle interactions.

Fermions are divided into two categories of building blocks of matter: leptons and quarks.

Leptons are categorised into three generations, each featuring a charged lepton and a corresponding neutrino. The first generation includes the electron (e^-) and its associated neutrino (ν_e), the second consists of the muon (μ^-) and its neutrino (ν_μ), and the third includes the tau (τ^-) and its neutrino (ν_τ).

Quarks, on the other hand, come in six flavours, organised into three generations. The first generation includes up (u) and down (d) quarks, the second generation includes charm (c) and strange (s) quarks, and the third generation comprises top (t) and bottom (b) quarks. Quarks are never observed in isolation due to confinement, a phenomenon resulting from the strong force. Instead, they are contained within composite particles called hadrons, which are divided into two categories: mesons and baryons. Mesons consist of one quark and one antiquark, while baryons comprise three quarks. This arrangement stems from the fractional charges of quarks: the up, charm, and top quarks carry a charge of two thirds, while the down, strange, and bottom quarks have a charge of minus one third.

The fundamental forces that govern the interactions between particles in the Standard Model can be categorised into three main types: weak, electromagnetic, and strong forces. These forces, whose interactions with fermions are represented in Figure 1.1 with the green, red and blue lines, are carried by the W, Z, γ (photon) and g (gluon) gauge bosons respectively. Leptons can all engage in weak decays, and in electromagnetic decays in the case of charged leptons. Notably, the electron is stable, *i.e.*, it does not decay under normal circumstances, whereas the muon and tau particles decay weakly into lighter particles. Quarks interact in general via the strong, electromagnetic and weak forces.

These interaction are not limited to fermions, as demonstrated in Figure 1.2 where

they are more extensively depicted. In fact, some bosons also engage in interactions with each other, and even with themselves. The scalar Higgs boson is also included, and interacts with all massive particles, including itself.

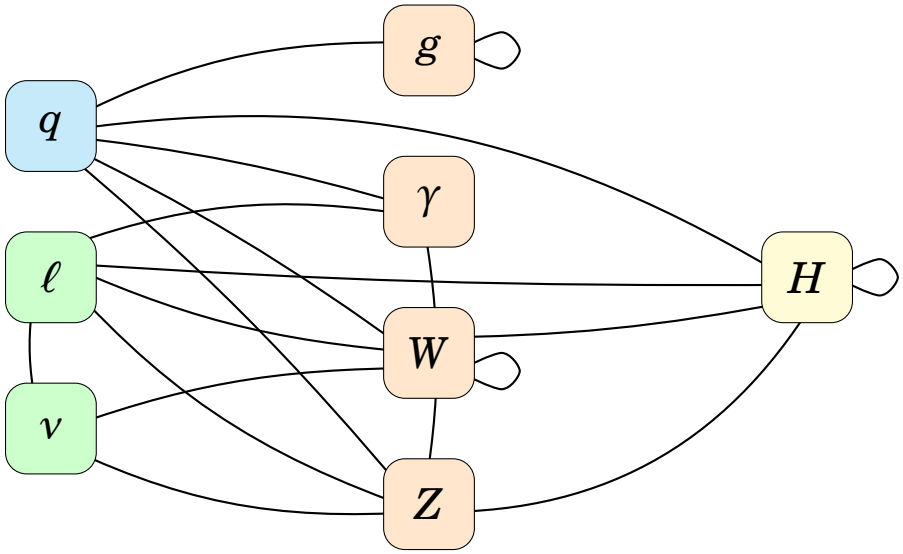


Figure 1.2: *Interactions within the Standard Model (SM) of particle physics.* The lines connecting the various particles of the Standard Model represent the possible interactions, including self-interactions. Quarks are denoted by the symbol q, charged leptons by ℓ, and neutral leptons by ν.

To understand these interactions, the fundamental theories that constitute the Standard Model need to be further examined. They are formulated using the quantum field theory (QFT). This theoretical framework, which aims at combining special relativity and quantum mechanics, is based on the fundamental hypothesis that the dynamics can be described by quantum fields, and that particles are manifestations of their excited states.

The mathematical description of these fields employs the Lagrangian formalism, which enables the derivation of equations of motion based on the principle of least action. In the context of the Standard Model, the Lagrangian consists of terms describing the dynamics of the fields associated with particles and their interactions. In fact, the sheer volume of these terms is such that when fully developed and written in standard font size, they scarcely fit onto an A4-sized paper. Nevertheless, it is possible to give a succinct form:

$$\mathscr{L} = \underbrace{-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}}_{1^{st}\text{ term}} + \underbrace{i\bar{\Psi}D\Psi + h.c.}_{2^{nd}\text{ term}} + \underbrace{\bar{\Psi}_i y_{ij}\Psi_j\phi + h.c.}_{3^{rd}\text{ term}} + \underbrace{|D_\mu\phi|^2 - V(\phi)}_{4^{th}\text{ term}}, \tag{1.1}$$

where the weak, electromagnetic, and strong forces are described by the first term,

their action on the elementary particles are translated by the second term, the Higgs boson is represented by the third term and the mechanism by which particles acquire their mass is expressed by the last term. In this context, $h.c.$ refers to an additional term which is the Hermitian conjugate of all of the preceding terms.

Furthermore, each interaction can be graphically represented by "*Feynman*" diagrams according to perturbation theory. These diagrams provide a simple, yet powerful tool for not only visualizing various particle interactions, but also calculating their probabilities in QFT.

Group theory is another and essential component of QFT, which is here only briefly discussed. A group, in the mathematical sense of the term, is a set of elements with operations, respecting associativity, and including an inverse for each element, as well as a single identity element. Groups that also observe commutativity are called "*Abelian*" groups.

There are numerous mathematical groups worth describing, yet two stand out as particularly noteworthy. The Unitary group $U(n)$ and the Special Unitary group $SU(n)$, both part of the General Linear group GL_n, can both be represented by $n \times n$ unitary matrices. The distinction lies in the fact that matrices representing $U(n)$ must possess a determinant with a norm of 1, whereas in the case of $SU(n)$, the determinant must be precisely equal to 1. The three groups are such that $SU(n) \subset U(n) \subset GL_n$.

The Unitary and Special Unitary groups are moreover Lie groups: they are both a group and a differentiable manifold. In this context, $U(n)$ is a real Lie group of dimension n^2, and $SU(n)$ is a strictly real Lie group of dimension $n^2 - 1$. This particularity implies above all that these groups are mathematically associated with a generator, a mathematical construct used to generate group elements by exponentiation. There is a deep connection between Lie group generators and the behaviors of fields and their excitations in QFT, as the symmetries of a physical system described by QFT often correspond to Lie groups.

Quantum electrodynamics (QED), one of the fundamental theories mentioned previously, first emerged from this established mathematical formulation. It is a relativistic quantum field theory of electrodynamics The Abelian $U(1)$ gauge group describes the gauge symmetry associated with the electromagnetic field, and the photon emerges as one of its generators. This force has an infinite range, and acts on all charged particles. The photon is a massless neutral particle which travels at the speed of light. The associated coupling strength α, called also fine-structure constant, is equal to 1/137.

This theory is unified with the weak interactions through a gauge group $SU(2) \times$

$U(1)$. This symmetry, known as electroweak symmetry, is preserved by the weak isospin fields $W1, W2$, and $W3$, and the weak hypercharge field B. The generators of $SU(2)$ and $U(1)$ are labeled as weak isospin (T) and weak hypercharge (Y) respectively. These generators give rise to the gauge bosons of the electroweak interactions: the W bosons (initially massless) and the B boson. However, these bosons are not yet physical fields until the process of spontaneous symmetry breaking occurs through the Higgs mechanism. This phenomenon leads to the observed particles in the Standard Model, the W and Z bosons, and the photon, acquiring mass and becoming the carriers of the weak and electromagnetic forces. And non-commutative nature of non-Abelian gauge groups allows the force carriers to interact with themselves too. In terms of QFT, the W^+ and W^- bosons are charged generators of $SU(2)$ and photon and Z are a linear combination of $SU(2)$ and $U(1)$. The weak force has a relatively small range of about 10^{-3} fm, and acts on flavour. The W and Z bosons have a mass of around 80 and 90 GeV respectively. The associated weak coupling strength α_W is equal to 10^{-6}.

Another and last fundamental theory to explore, within the framework of the Standard Model, is Quantum Chromodynamics (QCD). It is a quantum field theory described by the non-Abelian $SU(3)$ gauge group. QCD features eight generators, all of which manifest as gluons. These gluons are massless and neutral, yet they carry a distinct charge known as color, which is conserved. Each gluon possesses one color charge and one anti-color charge. Quarks, on the other hand, carry either a single color or a single anti-color. The color spectrum consists of three primary colors, red, blue, and green, and their corresponding complementary anti-colors, anti-red, anti-blue, and anti-green. Notably, only colorless entities have been observed in nature, leading to the concept of quark confinement. The strong force has a range of approximately 1 fm, and acts mainly on colour charge. The strong coupling constant α_S is equal to 1.

By combining the electroweak interactions, and the QCD, we obtain the composite gauge group $SU(3) \times SU(2) \times U(1)$, which represents the current mathematical formulation of the Standard Model.

1.1.2 Limitations and Consequences of the Standard Model

The Standard Model has been established as a comprehensive theoretical framework, accurately describing numerous physical phenomena. In recent years, however, several measurements have emerged, revealing its inherent limitations. Notably, phenomena such as neutrino oscillations [11], which confirm non-zero neutrino masses through their oscillation probabilities, persist without explanation

within the Standard Model framework. They have also raised additional challenges such as the hierarchy problem. Likewise, the existence of dark matter, inferred from observations such as the rotation curves of spiral galaxies [12], remains unaccounted for.

Several theoretical extensions have been proposed to address its limitations and answer its open questions. One such extension is Supersymmetry (SUSY), which postulates a symmetry between fermions and bosons, introducing superpartners for each particle in the Standard Model. These superpartners could potentially resolve issues such as the hierarchy problem and provide candidates for dark matter.

Other straightforward extensions to consider involve the addition of an extra $U(1)$ gauge group to the Standard Model. This addition results in several implications, notably the introduction of an additional gauge boson commonly denoted as Z'. There exists several extensions, but a particular one stands out by its simplicity. Known as the $L_\mu - L_\tau$ extension, it allows the Z' to couple to the Standard Model only through μ, τ, ν_μ and ν_τ with coupling g'. This extension offers explanations for anomalies such as the anomalous magnetic moment of the muon [13]. Two searches associated with this Z' have been performed at Belle II [14, 15], with the second one briefly discussed in Appendix C.1.

Furthermore, various other theoretical frameworks, such as Grand Unified Theories (GUTs), extra dimensions, and composite models, propose additional particles, interactions, and symmetries beyond those in the Standard Model. These extensions would likely entail modifications to the Standard Model Lagrangian, introducing new terms corresponding to the additional particles and interactions, as well as potentially altering the gauge group structure to accommodate the new symmetries.

To assess the viability of theories that complement the Standard Model, it is important to pursue ongoing investigations that would facilitate the emergence of new theories. But it is equally important to evaluate the precision of the Standard Model's predictions, identifying areas where it falls short.

An increasingly common test involves examining lepton flavour universality (LFU). This concept predicts that the interaction strength between gauge bosons and leptons remains consistent across different lepton flavours. Precision measurements of LFU parameters are essential for validating predictions within the Standard Model.

However, the scrutiny extends beyond the lepton sector, including precision measurements of the Cabibbo-Kobayashi-Maskawa (CKM) matrix elements within the quark sector. Through thorough experimental examination of these components, not only are the predictions of the Standard Model validated but also potential de-

viations that may hint at the existence of New Physics are probed.

1.1.2.1 Cabibbo–Kobayashi–Maskawa (CKM) Matrix and $|V_{us}|$

Weak currents are an inherent aspect of the electroweak interaction in the Standard Model. They are divided into neutral and charged currents, depending on the boson that is exchanged. The charged currents involve the exchange of W bosons, during which a change in the flavour of the participating particles occurs. These currents are further classified into hadronic, leptonic or even semileptonic, depending on the type of particle that is involved in the process.

Contrary to expectations, no principle of weak charged coupling universality between leptons and hadrons is respected, even though the same structure of generations is observed. Indeed, for leptons, coupling between different generations is forbidden, whereas for hadrons, coupling can occur across different generations, albeit with varying frequencies, as sketched in Figure 1.3.

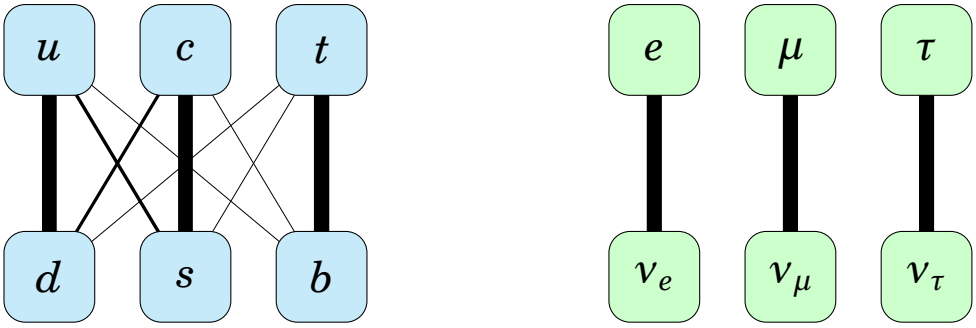


Figure 1.3: *Quark and leptons permitted mixings.* The lines connecting the various particles of the Standard Model represent the possible products resulting from an interaction with a W boson. Their thickness depict the interaction strength.

The observation of this difference of behaviour and unequal coupling strength for the quarks dates back to the time when only two generations of leptons and three quarks were known. It is in 1963 that Nicola Cabibbo suggested the existence of an extra state for weak interactions [16], for the quarks only, to explain the difference of decay rate observed between $K^- \to \mu^- \bar{\nu}_\mu$ and $\pi^- \to \mu^- \bar{\nu}_\mu$. This extra state would introduce two additional states d' and s', perpendicular to one another, and which would be expressed as:

$$\begin{cases} u &= u \\ d' &= d\,\cos(\theta_c) + s\,\sin(\theta_c) \\ s' &= -d\,\sin(\theta_c) + s\,\cos(\theta_c) \end{cases} \tag{1.2}$$

This definition uses the angle θ_c, an additional parameter that is needed to define these new states, derived for three quarks. This angle is nowadays also known as the "*Cabibbo angle*". After the introduction of the c quark in 1970 through the GIM mechanism [17], Kobayashi and Maskawa [18] postulated the existence of a third family of quarks, and generalised the states found by Cabibbo to these three generations:

$$\begin{pmatrix} u \\ d' \end{pmatrix}, \quad \begin{pmatrix} c \\ s' \end{pmatrix}, \quad \begin{pmatrix} t \\ b' \end{pmatrix}. \tag{1.3}$$

The motivation behind this generalisation stemmed from a desire to extend the Cabibbo-GIM model to accommodate Charge-Parity (CP) violation. In this context, charge symmetry (C) entails an operation that reverses particle charges to transform them into their corresponding antiparticles, while leaving other quantum numbers unaffected. Parity symmetry (P) refers to the transformation of spatial coordinates, and specifically involves reversing the direction of all spatial axes, effectively flipping the sign of each coordinate.

To describe these new states from the usual strong states, a 3×3 unitary matrix of rotation such as

$$\begin{aligned} \begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta_{13}} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta_{13}} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \\ &= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{13}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{13}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{13}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{13}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{13}} & c_{23}c_{13} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \end{aligned} \tag{1.4}$$

is needed, and includes not only the cosine and sine functions, abbreviated as c_{ij} and s_{ij}, but also a phase δ_{13}. This matrix, known as the so-called "*CKM matrix*", contains the interaction strength that were introduced earlier, and can be also expressed for simplicity as in the following expression:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \underbrace{\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}}_{V_{\mathrm{CKM}}} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{1.5}$$

It is worth mentioning that these interaction strengths and quark masses have a common origin in the Standard Model. They arise from the Yukawa interactions with the Higgs condensate,

$$\mathscr{L}_Y = -Y^d_{ij}\overline{Q^I_{Li}}\phi d^I_{Rj} - Y^u_{ij}\overline{Q^I_{Li}}\varepsilon\phi^* u^I_{Rj} + h.c. \tag{1.6}$$

where i, j are generation labels, $Y^{u,d}$ are 3×3 complex matrices, Q^I_L are left-handed quark doublets, ε is the 2×2 antisymmetric tensor, ϕ is the Higgs field, and d^I_R and u^I_R are right-handed down-type and up-type quark singlets, respectively, in the weak-eigenstate bases. The subsequent derivation of V_{CKM} is available in [19].

The CKM matrix possesses the property of unitarity [19], which yields:

$$V_{\text{CKM}}\,V^*_{\text{CKM}} \equiv V_{\text{CKM}}\,V^\dagger_{\text{CKM}} = V^\dagger_{\text{CKM}}\,V_{\text{CKM}} = I_3, \tag{1.7}$$

where $V^*_{\text{CKM}} \equiv V^\dagger_{\text{CKM}}$ is the conjugate transpose of V_{CKM}. From this property, nine equations involving the different elements of the matrix can be deduced. Three of them involve the magnitude of the elements:

$$\begin{aligned} |V_{ud}|^2 + |V_{us}|^2 + |V_{ub}|^2 &= 1 \\ |V_{cd}|^2 + |V_{cs}|^2 + |V_{cb}|^2 &= 1 \\ |V_{td}|^2 + |V_{ts}|^2 + |V_{tb}|^2 &= 1 \end{aligned} \tag{1.8}$$

And the remaining six equations are expressed as follows:

$$\begin{aligned} V_{ud}V^*_{us} + V_{cd}V^*_{cs} + V_{td}V^*_{ts} &= 0 \qquad & V_{ud}V^*_{cd} + V_{us}V^*_{cs} + V_{ub}V^*_{cb} &= 0 \\ V_{us}V^*_{ub} + V_{cs}V^*_{cb} + V_{ts}V^*_{tb} &= 0 \qquad & V_{cd}V^*_{td} + V_{cs}V^*_{ts} + V_{cb}V^*_{tb} &= 0 \\ V_{ud}V^*_{ub} + V_{cd}V^*_{cb} + V_{td}V^*_{tb} &= 0 \qquad & V_{ud}V^*_{td} + V_{us}V^*_{ts} + V_{ub}V^*_{tb} &= 0 \end{aligned} \tag{1.9}$$

For simplicity, Equations (1.8) and (1.9) can be reduced using Einstein summations to

$$V_{ij}V^*_{ik} = \delta_{jk} \quad \text{and} \quad V_{ij}V^*_{kj} = \delta_{ik}. \tag{1.10}$$

In particular, one of the element within the CKM matrix, denoted as V_{us}, encapsulates the coupling strength between the up and the strange quarks. This coupling is associated to interactions involving two different generations of quarks, also called

"*Cabibbo suppressed*". There are several ways of determining this element, whose value is usually expressed as a magnitude $|V_{us}|$, as the matrix includes imaginary numbers. These determinations are all summarised in Figure 1.4, but before commenting these results, the different methods are being discussed.

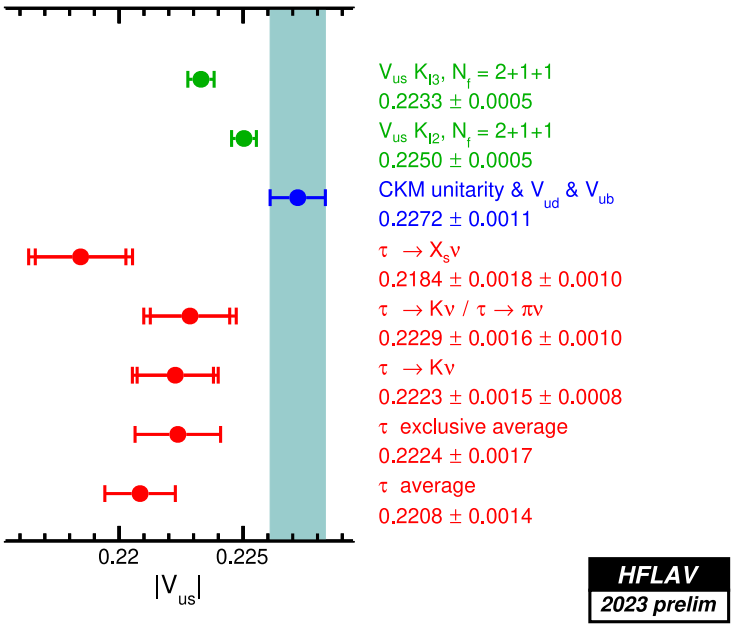


Figure 1.4: *Measurements of* $|V_{us}|$. Several methods are employed to measure its value, such as unitarity and measurement from kaon and tau decays. All values, coming from [20], represent a preliminary update of values available in [21].

The first method employs the property of unitarity of the CKM matrix. It follows from Equation (1.8) that:

$$|V_{us}|^2 = 1 - |V_{ud}|^2 - |V_{ub}|^2, \tag{1.11}$$

where $|V_{ud}|$ is determined from superallowed nuclear beta decays [22] and $|V_{ub}|$ can be measured via B meson decays. One advantage of that method is that the associated theory uncertainties are expected to be of the order of 0.025%. Another important aspect is that the term $|V_{ub}|^2$ can often be neglected, as $|V_{ub}| \sim 10^{-5}$ and the experimental uncertainty of $|V_{ud}| \sim 10^{-4}$, which gives

$$|V_{us}|^2 \simeq 1 - |V_{ud}|^2. \tag{1.12}$$

Independent measurements of $|V_{us}|$ are performed via other methods, one of which uses kaon decays. These decays are either $K^- \to \ell^- \bar{\nu}_\ell$ or $K^- \to \pi \ell^- \bar{\nu}_\ell$ labeled as $K_{\ell 2}$ or $K_{\ell 3}$ respectively [23], with $\ell = e, \mu,$.

In the first case of kaon decays, a value of $|V_{us}|$ can be extracted by the comparison of the radiative inclusive decay rates [19] for $K \to \ell\nu(\gamma)$ and $\pi \to \ell\nu(\gamma)$, such as

$$\frac{\Gamma_{K_{l2}}}{\Gamma_{\pi_{l2}}} = C\,\frac{|V_{us}|^2}{|V_{ud}|^2}\,\frac{f_K^2}{f_\pi^2}, \tag{1.13}$$

where C is a constant translating a small error coming from electroweak radiative corrections, and f_K^2/f_π^2 are form factors calculated from lattice gauge theory.

In the second case, a value for $|V_{us}|$ can be derived as follows [19]:

$$\Gamma_{K_{l3}} = \frac{G_F^2\, m_K^5}{192\,\pi^3}\, S_{EW}\,(1+\delta_K^\ell+\delta_{SU2})\, C^2\,|V_{us}|^2\, f_+^2(0)\, I_K^\ell, \tag{1.14}$$

where G_F is the Fermi constant, m_K is the kaon mass, S_{EW} is the short-distance radiative correction, δ_K^ℓ is the mode-dependent long-distance radiative correction, δ_{SU2} is the deviation from one for the ratio of $f_+(0)$ for the charged to neutral kaon decays, which is equal to zero for the neutral kaon, C^2 is a constant equal to 1/2 for charged kaon decays, and to 1 for neutral kaon decays, $f_+(0)$ is the calculated form factor at zero momentum transfer for the $\ell\nu$ system, and I_K^ℓ is the phase-space integral, which depends on measured semileptonic form factors.

It is worth mentioning that most early determinations of $|V_{us}|$ using this method were based solely on $K \to \pi e\nu$ decays as the $K \to \pi\mu\nu$ decays had a larger uncertainties in I_K^ℓ. Indeed, theory uncertainties are also playing a major role in the determination of $|V_{us}|$, as it is a precision measurement. In the case of kaon decays, its precision is limited by uncertainties of the lattice QCD estimates of the meson form factor $f_+^2(0)$ and decay constant in f_K/f_π. This leads to an expected theory uncertainty of the order of 0.50% for $K_{\ell 2}$ and 0.58% for $K_{\ell 3}$.

While $|V_{us}|$ is most precisely determined from kaon decays [24, 25], it is possible to use an alternative method that does not depend on lattice QCD, has small theory uncertainties [26, 27, 28, 29, 30, 31, 32], and which uses another observable: tau decays branching fractions.

Various types of tau branching fractions can be considered, one of which is attributed to inclusive decays, denoted as $\mathscr{B}(\tau \to X_s\nu)$, as it includes all strange modes. This first case relies on the fact that the tau hadronic partial width is the sum of both tau partial widths of strange and non-strange hadronic final states,

$$\Gamma_{\tau,\text{had}} = \Gamma_{\tau,\text{strange}} + \Gamma_{\tau,\text{non-strange}}. \tag{1.15}$$

Ratios $R_{\tau,X}^w$ can be constructed by diving any partial width Γ_X by the electronic

partial width Γ_e, translating the decay $\tau^- \to e^- \bar{\nu}_e \nu_\tau$, which gives

$$R^w_{\tau,\text{had}} = R^w_{\tau,\text{strange}} + R_{\tau,\text{non-strange}}. \tag{1.16}$$

In terms of such ratios, $|V_{us}|$ can be measured as

$$|V_{us}|^2 = \frac{R^w_{\tau,\,\text{strange}}}{\frac{R^w_{\tau,\,\text{non-strange}}}{|V_{ud}|^2} - \delta R^w_\tau}, \tag{1.17}$$

where δR^w_τ can be determined using perturbative QCD and partly relying on experimental low energy scattering data. The associated theory uncertainty are expected to be of the order of 0.23% up to 0.47%.

$|V_{us}|$ can also be determined using the absolute branching fraction $\mathscr{B}(\tau^- \to K^- \nu_\tau)$, as in the following:

$$\mathscr{B}(\tau^- \to K^- \nu_\tau) = \frac{G_F^2 S_{EW}}{16\pi\hbar}\,\tau_\tau\, m_\tau^3\, f_K^2\, |V_{us}|^2 \left(1 - \frac{m_K^2}{m_\tau^2}\right)^2 \left(1 + \delta_{R_{\tau,K}}\right), \tag{1.18}$$

where τ_τ is the lifetime of the tau lepton, m_τ is its mass, m_K is the mass of the kaon, and $\delta_{R_{\tau,K}}$ is a radiative correction to $\tau \to K\nu_\tau$. The other parameters have been introduced in the previous equations. In that case, the theory uncertainties are one of the highest, as they are expected to be at the order of 1.27%.

One last case uses a ratio of the branching fraction $\mathscr{B}(\tau^- \to K^- \nu_\tau)$ over $\mathscr{B}(\tau^- \to \pi^- \nu_\tau)$, and using Equation (1.18), this yields

$$R_{K/\pi} = \frac{\mathscr{B}(\tau^- \to K^- \nu_\tau)}{\mathscr{B}(\tau^- \to \pi^- \nu_\tau)} = \frac{f_K^2}{f_\pi^2}\,\frac{|V_{us}|^2}{|V_{ud}|^2}\,\frac{\left(1 - (m_K^2/m_\tau^2)\right)^2}{\left(1 - (m_\pi^2/m_\tau^2)\right)^2}\left(1 + \delta_{R_{\tau,K/\tau,\pi}}\right), \tag{1.19}$$

where m_π is the mass of the pion, and the rest of the parameters have been previously introduced. The current most precise measurement using this determination is coming from the BaBar collaboration [33], and gives $R_{K/\pi} = 0.06531 \pm 0.00056\,(\text{stat.}) \pm 0.00093\,(\text{syst.})$, which yields $|V_{us}| = 0.2255 \pm 0.0024$. Another important aspect to highlight is that they achieved a purity of 78.7% and an efficiency of 0.324% in the pion channel, and a purity of 76.6% with an efficiency of 0.330% in the kaon channel for this measurement.

All these methods led to results that are represented in Figure 1.4. It compares the determination of $|V_{us}|$ assuming unitarity (Equation (1.11)) versus the measurement from kaon (Equations (1.13) and (1.14)) and tau decays (Equations (1.17), (1.18) and (1.19)).

A tension is observed between the different types of methods, but what raises more pressing questions is the tension observed within tau decays methods. Another observation is that all determinations of $|V_{us}|$ from branching fraction measurements are lower than the CKM-unitarity assumption. This motivates an updated measurement of channels accessible with Belle II data.

2. The Belle II Experiment

The Belle II experiment is a high-energy physics experiment hosted at the High Energy Accelerator Research Organisation (KEK) in Tsukuba, Japan. Belle II began construction in 2010 and took approximately seven years to complete, with its commissioning phase starting in 2018. It was constructed as an upgrade to its predecessor, the Belle experiment, which operated from 1999 to 2010.

The experiment operates at the SuperKEKB accelerator, an asymmetric beam energy electron-positron collider designed to produce large amounts of B mesons, hence the frequent use of the term "*B-factory*" to qualify such experiments. The collision data is collected at the interaction point (IP) by the Belle II detector. Belle II's primary goals are to search for new physics (NP) and to improve the measurements of Standard Model parameters. But, like its predecessor, it also focuses on exploring Charge-Parity (CP) violation and rare decays in B mesons to better understand the universe's matter-antimatter asymmetry.

In fact, notable discoveries from Belle include the observation of rare B meson decays [34], contributions to our understanding of CP violation [35] and precision measurements of Standard Model parameters [36]. Belle II's physics program encompasses a wide range of topics, including, among others, B, D and τ physics, and investigations into dark matter and dark energy phenomena [37]. As it continues its operation, the Belle II experiment holds the potential to make further discoveries.

2.1 SuperKEKB Accelerator

SuperKEKB is a circular electron-positron accelerator. As depicted in Figure 2.1, it consists first of a linear injection system (linac) that accelerates the electrons e^- and positrons e^+. As positrons are produced with higher emittance than electrons, they have to go through an additional stage, the damping ring. This dedicated positron damping ring prepares the positron beam by reducing its emittance and

stabilizing its trajectory. Once the beam is sufficiently conditioned, it is injected, like the electron beam is, into the main accelerator rings. This last part includes two rings, one for each particle, and has a circumference of 3 km. The Belle II detector, that will be detailed in Section 2.2 is situated at the interaction point of both beams.

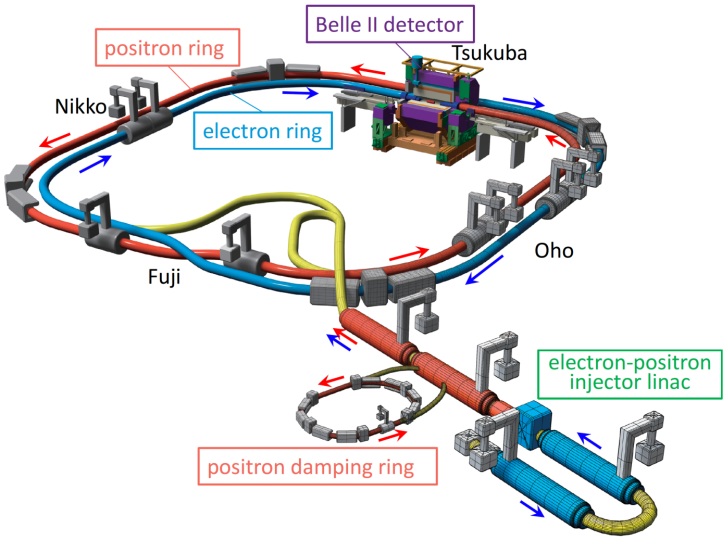


Figure 2.1: *SuperKEKB Accelerator.* Schematic view showing the injection system, the positron damping ring, the electron ring (HER), positron ring (LER) and the Belle II detector, from [38].

As mentioned earlier, SuperKEKB is an asymmetric beam energy accelerator, which means that electrons and positrons are not accelerated with the same energy. The electron ring is a high energy ring (HER) of 7 GeV, and the positron ring is a low energy ring (LER) of 4 GeV. The asymmetry is a key element to allow for time dependent charge parity measurements. In comparison, Belle had a LER of 3.5 GeV and a HER of 8 GeV. This change of Lorentz boost from $\beta\gamma = 0.42$ to $\beta\gamma = 0.28$ enables measurement improvements. Even if it reduces the beam loss due to Touschek scattering in the lower energy beam, detailed in Section 2.1.1, which, in turns, reduces the spatial separation between B mesons, it leads to improvements in solid angle acceptance for missing energy decays, and therefore is advantageous for analysis with neutrinos in the final state.

The beam energies are chosen not only to get boosted collision products, but also such that the resulting center of mass system (CMS) energy is 10.58 GeV. This energy corresponds to the $\Upsilon(4S)$ resonance, depicted in Figure 2.2. This resonance maximises the production of the B mesons through the decay $e^+e^- \rightarrow (\Upsilon(4S) \rightarrow B\bar{B})$, as the branching fraction of $\Upsilon(4S) \rightarrow B\bar{B}$ itself is higher than 96% [19]. Even though

the majority of the data is collected at the $\Upsilon(4S)$ resonance, the beam energies are flexible enough to cover other resonances as well, from $\Upsilon(1S)$ at 9.46 GeV to $\Upsilon(6S)$ at 11.24 GeV. These resonances do not necessarily decay into the same final products, allowing further analysis to be undertaken. The CMS energy is confined within this specific resonance range; the upper limit is set by the constrained power of the linear injector, and even if this limit could be circumvented, the beam transport limit due to magnet strength would only allow a restricted increase. The lower limit is due to the current lattice design not being sufficient to accommodate lower energies.

Υ Res.	M [GeV/c^2]
$(1S)$	9.46
$(2S)$	10.02
$(3S)$	10.36
$(4S)$	10.58
$(5S)$	10.86
$(6S)$	11.02

Table 2.1: Υ *resonances.* Mass values from [19].

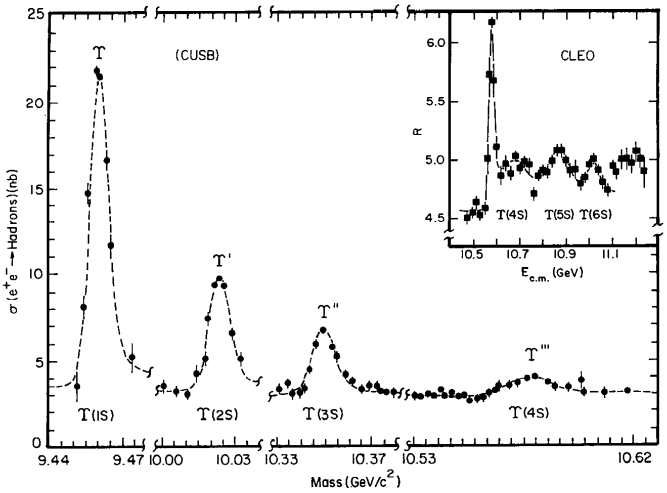


Figure 2.2: Υ *spectroscopy.* Cross section for inclusive production of hadrons vs m or E_{CMS}, from [39].

The performance of an accelerator can be expressed by the luminosity it is capable of delivering. The luminosity, denoted as $\mathscr{L}$, is a measure of the number of events N that are produced in a certain amount of time t, which takes into account their cross-section σ, as:

$$\mathscr{L} = \frac{1}{\sigma}\frac{dN}{dt}.$$

It is expressed in the units of $\mathrm{cm}^{-2}\mathrm{s}^{-1}$ or $\mathrm{b}^{-1}\mathrm{s}^{-1}$, where $1\ \mathrm{b} = 10^{-24}\ \mathrm{cm}^2$. A useful related quantity is the integrated luminosity $\mathscr{L}_{\mathrm{int}}$, which is:

$$\mathscr{L}_{\mathrm{int}} = \int \mathscr{L}\,\mathrm{d}t.$$

The luminosity can be estimated by different methods that mainly relies on counting the number of events of a very specific type of controlled events produced in great quantities.

During its first run period, called Run 1, SuperKEKB collected a data set corresponding to an integrated luminosity of 426.478 ± 0.022 (stat.) ± 2.750 (syst.) fb^{-1}, and achieved a luminosity world record of 4.71×10^{34} $\text{cm}^{-2}\text{s}^{-1}$ (which corresponds to 47.1 $\text{nb}^{-1}\text{s}^{-1}$). These remarkable numbers have been reached thanks to a number of upgrades or modifications that have been made in comparison to KEKB. This includes a change of the beam current, which has almost doubled, a reduction of the beam size by a factor 20 at the interaction point, going from 1 μm to 50 nm, an increase of the crossing angle from 22 mrad to 83 mrad and the use of a new final focus superconducting magnet system (QCS) at the interaction point, which consists of quadrupole magnets, corrector magnets and compensations solenoid coils cancelling the effect of the solenoid field of the Belle II detector on the beams. This is known as the so-called "*nano beam*" scheme [40].

As of the writing of this thesis in spring 2024, the Belle II run period is divided into three main parts. Already mentioned earlier, the first run period Run 1, comprising three different phases, started in February 2016 and lasted until July 2022. A long shutdown (LS1) then took place until January 2024 when the second run period Run 2 started, and is meant to continue until 2026 or 2027.

The first phase of Run 1 corresponds to the commissioning of SuperKEKB from February to June 2016, without the Belle II detector and aforementioned superconducting magnet system. The second phase then took place, during four months of which the first data was taken, corresponding to an integrated luminosity of 0.5 fb^{-1}. As the full vertex sub-detector, detailed in Section 2.2.1, was not yet completed during these first runs, different background sensors were instead installed [41]. The third and last phase of Run 1 started in March 2019, when the full detector, including the vertex sub-detector, was installed. This last phase corresponds to the data-taking period detailed in Figure 2.3.

As highlighted in the timeline on the right-hand side, this data-taking period is divided into several parts, called "*Experiment*". There is a total of 14 different Experiments, all categorised as "*physics runs*", *i.e.*, runs during which physical data is acquired. The Experiments correspond to specific run conditions; while the majority of these Experiments translate a production of data at the $\Upsilon(4S)$ resonance, there are two Experiments which solely translate either another resonance, or off-resonance data: Experiment 21 and 25. As for the Experiments 8, 12, 18 and 22, they not only correspond to a production of data at the $\Upsilon(4S)$ resonance, but also a production of data at either another resonance, or by a scan of a certain resonance.

LS1 was necessary to make modifications and improvements to the accelerator. It includes, among other things, the replacement of the beam pipe, the reinforcement

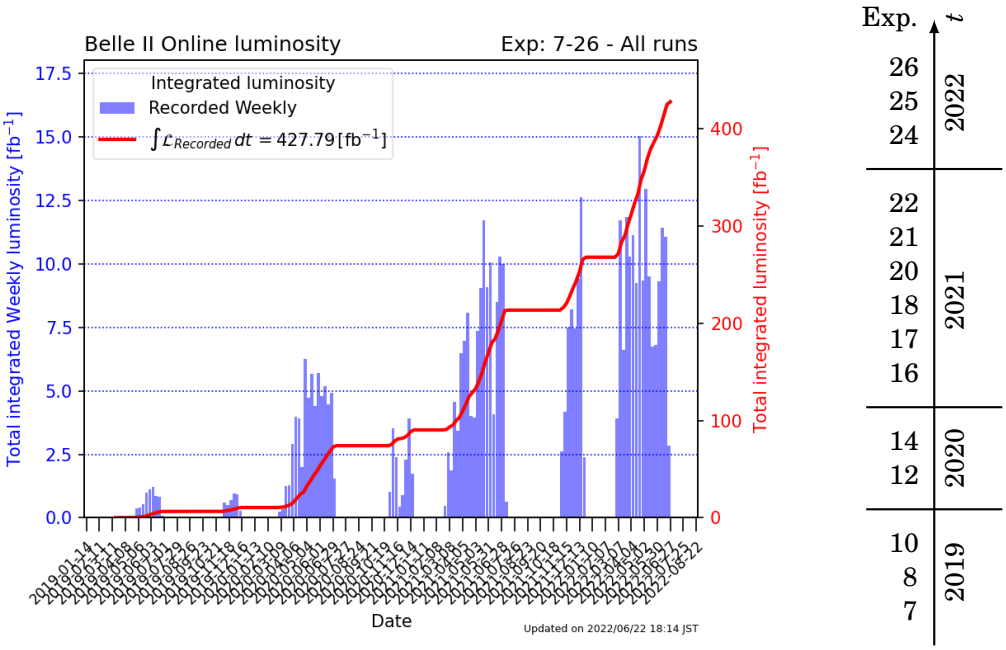


Figure 2.3: *Belle II data-taking and luminosity.* Total recorded integrated luminosity before Long Shutdown 1 (LS1), from [42]. Experiments numbers are detailed on the right hand side according to the year in which they were measured.

of radiation shielding around the interaction point, the replacement of the cap at the head of the QCS, the installation of a new collimator in a straight section of the accelerator, and the chamber modification of the HER injection section. But also other upgrades of the Belle II detector, which will be detailed later, in Section 2.2.

After these upgrades, SuperKEKB was ready to resume operations in January 2024, when Run 2 started. After a few tests, the first collisions were observed in February 20, continuing the data-taking at experiment 30. Three main goals were set for the year 2024: run stably at 10^{35} cm^{-2}s^{-1}, reach 150 fb^{-1} per month, and exceed 1 ab^{-1} of data. It is hard to predict now whether these goals will be achieved, but all the means are in place to achieve them.

2.1.1 Beam Backgrounds

There exists several sources of beam background that are expected from SuperKEKB. The five most dominant in the Belle II detector environment and which should impact performances are the following:

- ⋆ *Touschek effect*: It describes the phenomenon when electrons in a particle bunch undergo scattering, resulting in energy deviations and potential particle loss within the beam pipe. When such losses occur in proximity to the

interaction point (IP), the resulting shower of particles may reach the active detector material. At SuperKEKB, this effect is magnified due to the nano-beam scheme in place; the scattering rate is influenced by factors such as the beam bunch current and the number of bunches, while inversely proportional to the beam size and the third power of the beam energy. To counteract this, horizontal and vertical collimators are strategically positioned around the ring and immediately preceding the interaction region.

⋆ *Beam-gas scattering*: Residual gas molecules present in the beam pipe may scatter beam particles. This effect can manifest through either Bremsstrahlung scattering, leading to a reduction in particle energy, or Coulomb scattering, affecting particle direction. The expected impact of this background is contingent upon factors such as the beam current, vacuum pressure within the rings, and the material surrounding the magnets. Similar mitigation strategies to those employed for the Touschek effect are implemented to address this phenomenon.

⋆ *Radiative Bhabha processes*: Photons generated by the radiative Bhabha process $e^+e^- \rightarrow e^+e^-(\gamma)$ may propagate along the beam axis before interacting with the iron components of the magnets. This interaction results in a significant production of neutrons via the giant resonance process within the nuclei of the magnets. The rate of neutron generation correlates directly with the luminosity, necessitating additional shielding within the accelerator tunnel to mitigate potential risks.

⋆ *Two-photon processes*: The occurrence of luminosity-dependent background attributed to low momentum electron-positron pairs generated through the two-photon process $e^+e^- \rightarrow e^+e^-e^+e^-$ have the potential to spiral along the solenoid field lines. This can result in multiple hits within the inner detectors.

⋆ *Synchrotron radiation*: Synchrotron radiation (SR), primarily emitted by electrons from the HER beam, is proportional to the square of the beam energy and the square of the magnetic field strength. To prevent SR photons from reaching the inner detectors, the inner surface of the beryllium beam pipe is coated with a gold layer.

2.2 Belle II Detector and Sub-Detector Systems

The Belle II detector surrounds the interaction point (IP) of the HER and LER [43], and is used to detect particles resulting from ee collisions. The whole detector is

approximately 8 m high × 8 m wide × 8 m long, and weighs about 1400 tons.

As depicted on Figure 2.4, the Belle II detector is composed of several sub-detector systems, each responsible for detecting a specific property of a given group of particles, thereby defining their location within the detection system. These sub-detectors undertake various tasks together, encompassing the measurement of particle passage, position, time, momentum, energy, and identity. To this end, they capitalise on diverse physical effects such as ionisation in gases, liquids, and solids, deflection in magnetic fields, electromagnetic showers, and Cherenkov radiations. Despite their fundamental differences, all sub-detector systems have one thing in common: their detection principle is based on the particle's passage through matter, hence their strong dependence on the type of particle they have to detect.

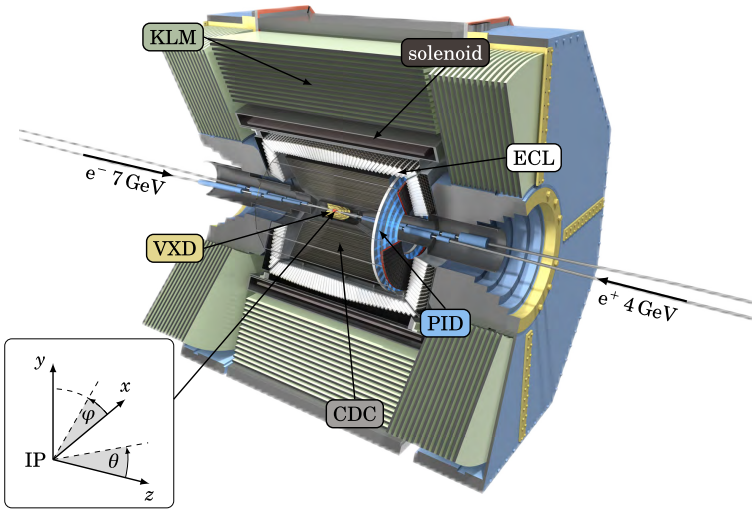


Figure 2.4: *3D rendering of the Belle II detector and sub-detector systems.* The different sub-detector systems are indicated and the interaction point (IP) is detailed with its coordinate system. The rendering is coming from [44].

The seven sub-detectors comprise an inner part that reconstructs vertices and tracks (PXD, SVD and CDC), a hadron identification system (TOP and ARICH), an electromagnetic calorimeter (ECL), a solenoid to deflect charged particles trajectories, and an outer K_L^0 and μ chamber (KLM), all mentioned from the nearest to the farthest from the *ee* interaction point. The position of all sub-detectors are detailed in Figure 2.5, on which the Belle II detector is shown from the side.

Before introducing all sub-detectors mentioned previously in greater detail, it is worth mentioning the upgrades that have been introduced by the change from the Belle to the Belle II detector, as well as the improvements undertaken during LS1. The changes from the Belle detector include the readjustment of the PXD detector

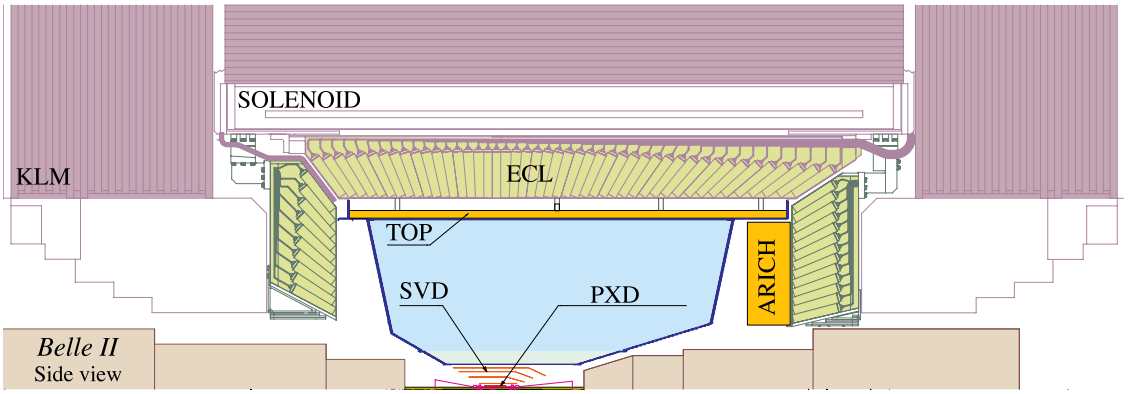


Figure 2.5: *Side view of the Belle II detector and sub-detector systems.* The different sub-detector systems are projected on the YZ-plan. The sketch is coming from [45].

at a smaller distance from the IP, a larger SVD acceptance, an improved K/π separation from the TOP and ARICH detectors, improved trigger systems adapted to a higher event rate, an improved coverage of almost the full solid angle, a higher reconstruction efficiencies for charged particles, a better momentum resolution and a more precise measurement of photons energy and direction [37, 43]. During LS1, the full two-layer pixel detector was installed and the photo-multiplier tubes (PMTs) of the time-of-propagation (TOP) detector were replaced.

2.2.1 Decay Vertices Reconstruction and Tracking

The reconstruction of decay vertices and particle tracking are two tasks that are undertaken by two detectors close to the interaction point: the vertex detector (VXD) and the Central Drift Chamber (CDC). These two detectors are the closest to the interaction point as they measure the impact parameters of charged tracks, reconstruct the primary and secondary decay vertices of particles decaying close to the IP, *e.g.*, B or D mesons, and perform high-precision tracking of charged particles through magnetic fields. The latter is crucial to provide information on the momentum of charged particles and hence on their mass.

The VXD detector is more precisely divided into two parts: the Pixel Silicon Sensor (PXD), positioned as close as possible to the IP, and the Silicon Strip Sensor (SVD) that is situated right after, as detailed in Figure 2.6.

Figure 2.6: *PXD and SVD.* Illustration of the innermost tracking sub-detectors, from [46]. The behaviour of four tracks in this particular sub-detector system is highlighted.

2.2.1.1 Pixel Silicon Sensors (PXD)

The challenging criteria for detectors functioning near the IP include requirements for timing precision, spatial resolution, and long-term operational performance under radiation exposure. Currently, these standards are fulfilled by pixel detectors, which consist of two distinct components: the particle sensing element, known as the "*sensor*", and the integrated electronics circuitry, referred to as the "*read-out chip*".

The underlying principle of the sensor lies in its use of Depleted P-channel Field Effect Transistor (DEPFET) [47] active pixels embedded within fully depleted silicon bulk. When the sensor is traversed by incoming charged particles, it generates a charge that is then collected by the electric field. Subsequently, the charge is brought to contacts where the read-out chip collects the charge for further processing.

The detector comprises two nearly cylindrical layers positioned just outside the beam pipe, with radii of 14 mm and 22 mm, respectively. Each layer consists of mirrored modules bonded together to form ladders. The inner layer accommodates eight ladders, while the outer layer holds twelve, totaling 20 ladders and 40 sensors. Overall, the detector comprises nearly 8 million pixels. Due to delays in ladder assembly, only the two ladders of the outer layer and the complete inner layer were installed for Run 1. The installation of the full detector was done for Run 2.

2.2.1.2 Silicon Strip Sensors (SVD)

The SVD is designed to enhance the functionality of the PXD detector, tasked with reconstructing decay vertices beyond the IP. As is it positioned at a greater distance, the SVD adopts a slightly modified design, utilizing double-sided strip technology for particle detection, while maintaining the fundamental operational principle associated with semiconductors.

Comprising four layers positioned at radii of 38, 80, 115, and 140 mm, the SVD spans the entire Belle II angular acceptance of $17° < \theta < 150°$. Each layer features a barrel-shaped region housing rectangular sensors, with the three outermost layers adopting a lamp-shade geometry hosting trapezoidal sensors in the forward direction. With 8 to 17 ladders per layer, each containing 2 to 5 sensors, the SVD boasts a total of 49 ladders and 187 sensors [43]. Achieving an impressive resolution of approximately 50 μm, the SVD ensures precise particle tracking across its detection layers.

2.2.1.3 Central Drift Chamber (CDC)

The CDC, depicted in Figure 2.7, complements the tracking module by being inserted around the VXD. Unlike the first two sub-detectors, the CDC operates on the fundamental principle of drift chamber technology. As charged particles traverse the CDC, they ionise the He-C_2H_6 gas mixture within. Guided by an electric field generated by the field wires, the resulting charges drift towards the nearest sense wire, among a total of 14336 arranged in layers, at an average velocity of 3.3 cmμs^{-1}, with a maximum drift time of approximately 350 ns. Subsequently, the read-out is is performed once charges have been accumulated at the sense wires.

The CDC consists of adjacent layers that group together to form superlayers, totaling six or eight in number. These superlayers alternate between axial (A) and stereo (U, V) orientations. Axial wires align with the beam axis, while stereo wires are oriented at an angle of 45.4 to 74 mrad relative to the beam axis in both positive and negative directions. Hits recorded in both axial and stereo wires can be combined to achieve complete 3D track reconstruction.

With its extensive coverage spanning up to 2.3 m long × 2.2 m of radius in the tracking volume, the CDC enables high-resolution reconstruction of charged particle tracks; its acceptance range extends from 17° to 150° in polar angle, with a resolution of approximately 100 μm. Additionally, it facilitates particle identification through the measurement of energy loss within the gas volume, particularly benefiting particles that do not interact with the sub-detectors outside of the CDC.

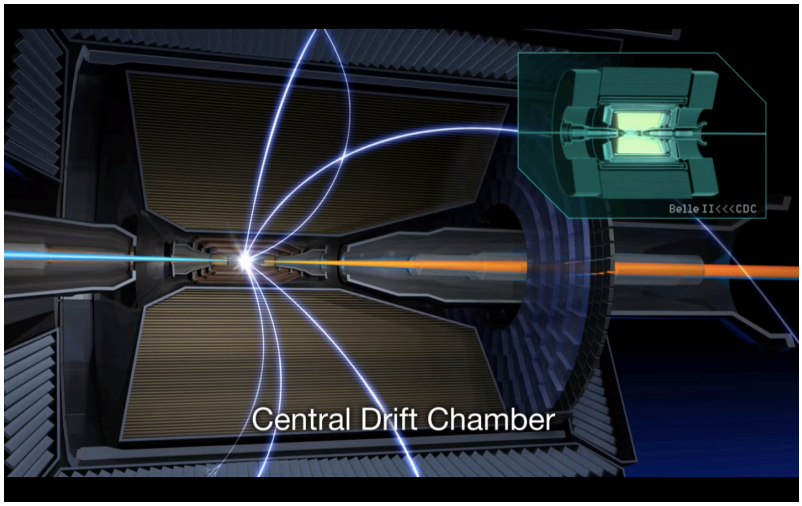


Figure 2.7: *CDC.* Illustration of the outermost tracking sub-detector, from [46]. The behaviour of six tracks in this particular sub-detector system is highlighted.

2.2.2 Hadron Identification System

Akin to the BaBar's Detector of Internally Reflected Cherenkov Light (DIRC), Belle II features a dedicated sub-detector systems aimed at hadron identification, specifically for distinguishing K from π; the identification of other particles occurs through the integration of multiple outputs from all detectors, resulting in likelihood estimates, as detailed in Section 2.3.

In contrast, the LHCb detector uses sub-detectors called Ring Imaging Cherenkov Detectors (RICH1 and RICH2) to identify all particles, supported by the Preshower (PS) and Scintillating Pad Detector (SPD) to further differentiate between electrons and photons.

In all three cases of experiments, the detectors rely on "*Cherenkov*" radiation processes for identification. The underlying principle is the following: when a charged particle with a velocity v travels faster than the speed of light c in a medium, also known as "*radiator*", with a reflective index n, it emits light. The angle θ_c at which the light is emitted is given by:

$$\cos\theta_c = \frac{1}{n\frac{v}{c}} = \frac{1}{n\beta}. \tag{2.1}$$

It is worth nothing that here the subscript c is used to attribute this angle to Pavel Cherenkov, and not to Nicola Cabibbo. The precise handling of hadron identification is shared by two distinct sub-detectors: the Time-Of-Propagation Counter (TOP) and the Aerogel Ring-Imaging Cherenkov Counter (ARICH). Their location

is also different. While the TOP is located in the region surrounding the innermost detectors, called "*barrel*", the ARICH is positioned in the region of the detector that covers the direction from the IP with small polar angles, called "*forward endcap*". Their shape and location are shown in Figure 2.8.

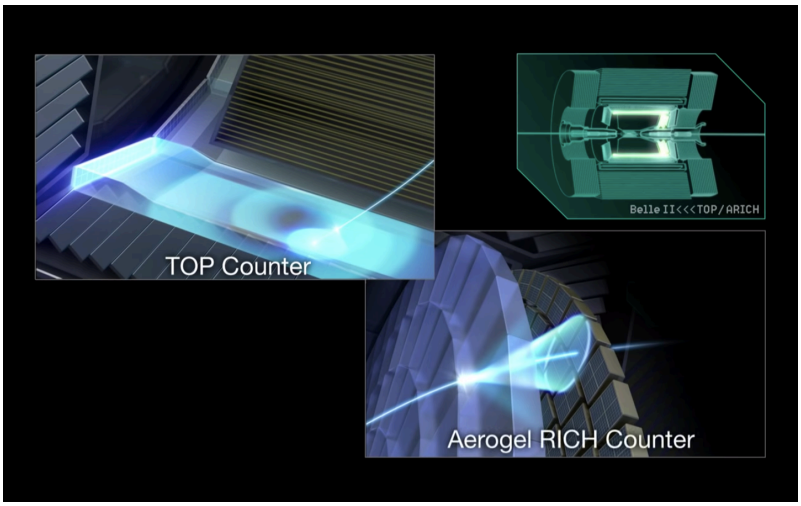


Figure 2.8: *TOP and ARICH.* Illustration of the sub-detectors responsible for the hadron identification, from [46]. The behaviour of a track in the barrel region and in the forward endcap is highlighted in blue shades.

The need for two separate detectors originates from several requirements within the Belle II experiment. While the ARICH effectively handles the identification of charged hadrons, the TOP detector complements it by providing precise timing measurements and time-of-flight-based particle identification. This division ensures comprehensive coverage and improves the overall particle identification. On top of that, having two separate detectors adds redundancy, enhancing its robustness against various experimental uncertainties and potential detector malfunctions; if one detector fails or provides unreliable data, the other detector can still provide useful information.

2.2.2.1 Time-Of-Propagation Counter (TOP)

The TOP is a unique type of Cherenkov detector that utilises two-dimensional information from a Cherenkov ring image, provided by the time of arrival and impact position of Cherenkov photons at the photo-detector situated at one end of a quartz bar. Its coverage spans from 31° to 128° in polar angle θ and consists of 16 modules arranged along the outer wall of the CDC. The modules feature quartz bars of 2.6 m length, each measuring 45 cm in width and 2 cm in thickness. Each module comprises four distinct components: two quartz bars serving as Cherenkov radiators,

a mirror positioned at the front end of the bar, and a prism coupling the bar to an array of microchannel plate photomultiplier tubes (MCP-PMTs). Due to the high refractive index of quartz ($n = 1.44$ at $\lambda = 405$ nm), some Cherenkov radiation emitted by particles traveling through the bar becomes trapped by total internal reflection and propagates to the MCP-PMT. The Cherenkov image is reconstructed by amalgamating the information on photon position coordinates and the measurement of their detection time.

2.2.2.2 Aerogel Ring-Imaging Cherenkov Counter (ARICH)

The key components of ARICH consist of aerogel tiles serving as the radiator, an array of position-sensitive photon detectors, and a readout system. The aerogel radiation was deliberately designed to be inhomogeneous, comprising multiple layers with varying refractive indices, aimed at maximizing the number of detected Cherenkov photons while simultaneously preserving the resolution of Cherenkov angle [48].

The chosen refractive indices ensure that Cherenkov rings from consecutive layers overlap at the photo detection plane, with values of $n = 1.045$ for upstream tiles and $n = 1.055$ for downstream tiles, respectively. For photon detection, the Hybrid Avalanche Photo-Detector (HAPD) is employed, which is developed jointly with Hamamatsu. It boasts high sensitivity to single-photon detection, immunity to magnetic fields, and the ability to provide position information. The HAPD sensor has dimensions of 73 mm long × 73 mm wide with 144 channels, where photoelectrons undergo acceleration over a potential difference of 8 kV and are subsequently detected in avalanche photodiodes (APD). The detector's performance encompasses an acceptance range of $14^\circ < \theta < 30^\circ$, and is especially capable of distinguishing pions and kaons within a momentum region spanning from 0.4 to 4 GeV/c.

2.2.3 Electromagnetic Calorimeter (ECL)

The Electromagnetic Calorimeter (ECL), as its name suggests, is specifically designed to measure the energy of incoming particles, and therefore must effectively stop them as much as possible to have an accurate measurement, as suggested in Figure 2.9. Consequently, it is strategically positioned further from the IP, allowing the measure of other properties before reaching the calorimeter.

The primary function of the ECL is to detect photons, accurately measure their energy and position, identify electrons, and generate trigger signals. Its objective is to completely halt electrons, minimise photon attenuation, and induce showers.

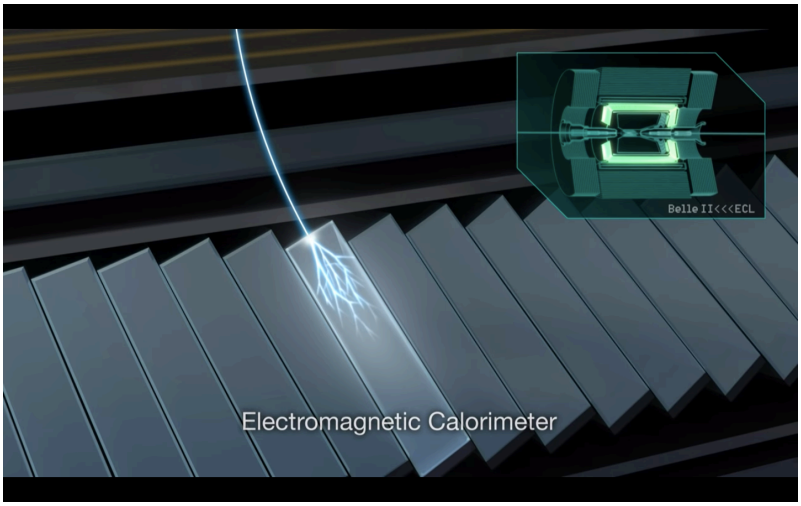


Figure 2.9: *ECL.* Illustration of the sub-detector responsible for the energy measurement, from [46]. The behaviour of a track entering this sub-detector system is highlighted.

The ECL is divided into three distinct detector regions: the barrel section, forward annular endcap, and backward annular endcap. These regions are equipped with a total of 8736 thallium-doped caesium iodide (CsI(Tl)) crystals, covering approximately 90% of the solid angle in the CMS. Each crystal is paired with two photodiodes affixed to its rear end, with signals emitted by both photodiodes processed by field programmable gate arrays (FPGAs). These devices measure the magnitude and timing of energy deposits in the crystal and extract additional waveform information to enable pulse shape discrimination (PSD). PSD serves as an effective means for distinguishing between hadron and photon showers and marks the first application of this technique at an electron-positron collider [49]. The ECL provides an effective coverage within the range of $12.4° < \theta < 155.1°$, with two small gaps situated between the barrel and the endcaps, spanning from 31.4° to 32.2° and 128.7° to 130.7°.

2.2.4 Detector Solenoid

While not traditionally considered a sub-detector component, the solenoid holds significance in the detection process due to its role in generating the magnetic field necessary for deflecting charged particles, and therefore allows for measurements of the mass of the particle through other parameters.

The solenoid generates a magnetic field of 1.5 T within a cylindrical volume measuring 3.4 m in diameter and 4.4 m in length, surrounding the ECL sub-detector [50].

It is integrated into a multi-layer structure comprising iron plates and calorimeters. Over the past decade, upgrades to the solenoid's power supply have significantly improved stability and reduced ripples, maintaining its performance while enhancing efficiency.

Additionally, the cryogenic system supporting the solenoid benefits from the reuse of much of its hardware, ensuring consistent heat load levels. Optimisation of operating conditions presents an opportunity for substantial cost savings through improved efficiency, aligning with successful strategies employed in other cryogenic systems [51].

2.2.5 K_L^0 and μ Chamber (KLM)

The last sub-detector to conclude the list of all sub-detectors defining the Belle II detector, and which is furthest from the IP is the K_L^0 and μ Chamber (KLM), which has two distinct behaviours that are depicted in Figure 2.10.

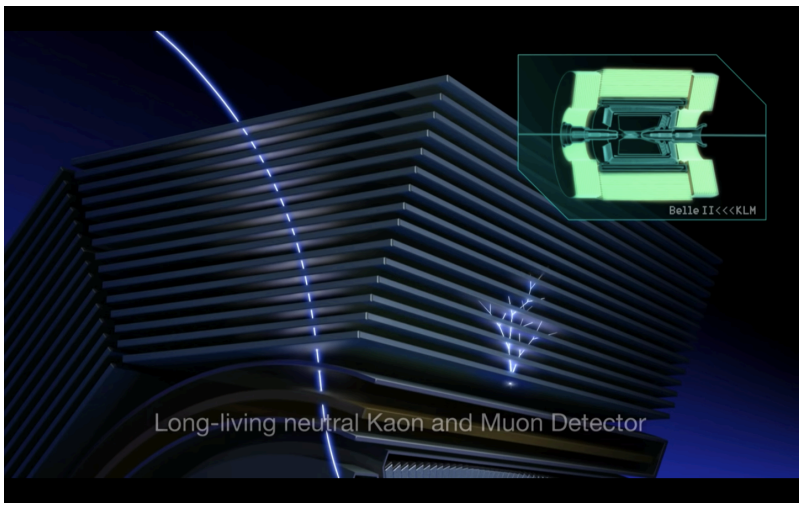


Figure 2.10: *KLM*. Illustration of the sub-detector responsible for the identification of K_L^0 and μ, from [46]. The behaviour two different tracks entering this sub-detector system is highlighted.

The KLM detector comprises alternating layers of iron plates and active detector elements, forming its structure in both the barrel and two endcaps, enveloping the superconducting solenoid. The detector elements are primarily glass-electrode resistive plate chambers (RPCs), where gas-filled gaps between glass electrodes are subjected to distributed high voltage. As charged particles traverse the chambers, they ionise gas molecules along their path, inducing an accelerated motion of electrons and ions towards the anode and cathode, respectively. This phenomenon,

intensified by a strong electric field (4.3 kVmm^{-1}) in the gap, triggers more ionisations and results in a streamer discharge between the electrodes, detectable by the RPCs. Hadronic showers generated by kaons and lambdas interacting in the iron plates are detected via this process.

To mitigate the long dead time of RPCs during recovery after a discharge, particularly under high background rates, RPC layers have been replaced with scintillator strips and silicon photomultipliers for read-out in the endcaps and inner barrel layers. The alternating design features 4.7 cm thick iron plates and active detector elements, resulting in a total of 15 detector layers and 14 iron plates in the barrel, and 14 detector layers and 14 iron plates in each endcap.

The detector's coverage spans five distinct polar angle regions: the forward endcap from 18° to 37°, the forward overlap from 37° to 47°, the barrel from 47° to 122°, the backward overlap from 122° to 130°, and the backward endcap from 130° to 155°.

2.3 Particle Identification (PID)

All Belle II sub-detectors contribute to the particle identification (PID) determination, with the exception of the PXD. Their contribution is more precisely defined by the computation of likelihoods $\mathscr{L}_x^d$, used to assess the identity of a particle. These likelihoods depend on several parameters that vary given the sub-detector $d \in \{$ SVD, CDC, TOP, ARICH, ECL, KLM$\} \equiv D$. And is also generated for each long-lived particle $x \in \{e, \mu, \pi, K, p, d\} \equiv X$. They are generally further combined into ratios of log likelihoods to consolidate the hypothetical identity of the particle and mitigate floating-point errors.

In the case of the SVD, the likelihood relies on the measurement of the energy loss ($\mathrm{d}E/\mathrm{d}x$) of charged particles traversing its sensors. A track typically provides eight measurements of $\mathrm{d}E/\mathrm{d}x$ (two for each sensor), with their average exhibiting a Landau distribution. The determination of the peak position is further improved by truncation, *i.e.*, eliminating the two highest energy measurements. The likelihoods is measured as

$$\mathscr{L}_x^{\mathrm{SVD}} = \prod_i \mathscr{P}_x\big((\mathrm{d}E/\mathrm{d}x)_i, p\big), \tag{2.2}$$

where the product runs on all the SVD remaining hits, and $\mathscr{P}_x$ is the two-dimensional probability density function (pdf) for particle hypothesis x.

In the CDC, measurements of $(\mathrm{d}E/\mathrm{d}x)$ are obtained from individual drift cells, where

the pulse signal is digitised, and values surpassing a certain threshold are integrated to create the raw ADC readout. Prior to utilizing these values in an average, various corrections are necessary, including geometrical path length adjustments and disparities among individual electronic channels. Variations in CDC gas gain, arising from fluctuations in pressure, temperature, and gas mixture composition, are monitored, determining a unified scale factor for each data-taking run. To approximate Gaussian behavior, we exclude the lower 5% and higher 25% of the corrected ADC measurements when calculating the average $(\mathrm{d}E/\mathrm{d}x)_m$. The likelihoods is expressed as

$$\mathscr{L}_x^{\mathrm{CDC}} = \exp\left[-\frac{1}{2}\left(\frac{(\mathrm{d}E/\mathrm{d}x)_m - (\mathrm{d}E/\mathrm{d}x)_p}{\sigma_p}\right)^2\right], \tag{2.3}$$

where $(\mathrm{d}E/\mathrm{d}x)_p$ is the predicted truncated average and σ_p is the expected resolution on the $\mathrm{d}E/\mathrm{d}x$ measurement, which depends on the polar angle, the number of hits on the track and the value $\mathrm{d}E/\mathrm{d}x$ itself.

As for the TOP, it is the Cherenkov photons, produced by the passage of relativistic charged particles and internally reflected through the bar until their detection at one extremity of each bar, that will be used. More precisely, their number, their arrival times and impact positions will serve to compute the likelihood, as

$$\mathscr{L}_x^{\mathrm{TOP}} = \exp\left[\sum_i \log\left(\frac{N_x S_x(c_i,t_i) + N_B B(c_i,t_i)}{N_x + N_B}\right) + \log P_N(N_x + N_B)\right], \tag{2.4}$$

where N_x and N_B are the number of expected signal and background photons, $S_x(c,t)$ and $B(c,t)$ are the signal and background pdf's, which depend on the position of the pixel defined by the index c and on the time of arrival of the photon t, and P is the Poisson probability to detect N photons while expecting $N_x + N_B$.

For the ARICH, the likelihood assessment relies on comparing the observed spatial distribution of photons, *i.e.*, hits, on the photo-detector plane with the pdf describing the anticipated distribution based on the parameters of a track traversing the ARICH and the assumed charged particle hypothesis. Typically, approximately 13 photons per relativistic charged particle are detected on average. The likelihood function for each hypothesis x is formulated by multiplying the probabilities of individual pixels being in the observed state, whether on or off, with no distinction made for pixels hit by single or multiple photons. As the probability of a pixel being hit by photons follows a Poisson distribution, it is given by $(1 - e^{-n_{x,k}})$, where $n_{x,k}$ is the expected average number of photons on pixel k for hypothesis x. The likelihoods is determined by

$$\mathscr{L}_x^{\mathrm{ARICH}} = \exp\left[-N_x + \sum_k (n_{x,k} + \log(1 - e^{-n_{x,k}}))\right], \tag{2.5}$$

where N_x is the expected total number of hits and the sum runs only over the pixels that were hit in an event.

In the case of the ECL, the likelihood is currently based only on the E/p ratio, which is the ratio between the measured energy E_{cluster} of a calorimeter cluster associated to the track connected to the particle hypothesis, and the momentum $p_{\mathrm{lab}} \equiv p$ determined by the tracking system. The likelihood is there given by

$$\mathscr{L}_x^{\mathrm{ECL}} = \mathscr{P}_x(E/p), \tag{2.6}$$

where $\mathscr{P}_x(E/p)$ is the pdf for particle hypothesis x.

Finally, for the KLM, the likelihood determination is based on the distance traveled by a charged particle within the detector; muons are able to traverse most or all of it, while hadrons are typically stopped within the first layers. The computation of the likelihood involves two main stages. First, the track is extrapolated from the outermost layer of the CDC that registered a hit. Then, the likelihood is computed for each particle hypothesis, but if the extrapolated track fails to intersect with the KLM, the algorithm assigns an identical likelihood to all particle hypotheses. The KLM likelihood comprises two components: the "*longitudinal*" term evaluates the observed penetration depth against the expected depth, while the "*transverse*" term is based on transverse shower characteristics. The total likelihood is expressed as:

$$\mathscr{L}_x^{\mathrm{KLM}} = \mathscr{L}_x^{\mathrm{KLM,L}} \cdot \mathscr{L}_x^{\mathrm{KLM,T}} = \left(\prod_k \mathscr{L}_{x,k}^L\right) \cdot \mathscr{L}_x^{\mathrm{KLM,T}} \tag{2.7}$$

where $\mathscr{L}_x^{\mathrm{KLM,L}}$ corresponds to the longitudinal term, $\mathscr{L}_x^{\mathrm{KLM,T}}$ is the transverse term, $\mathscr{L}_{x,k}^L$ represents the probability of hits for a particle of type x in the kth layer of the KLM. $\mathscr{L}_{x,k}^L = p_{x,k}$ if the hit has been recorded, and $\mathscr{L}_{x,k}^L = (1 - p_{x,k}\varepsilon_k)$ otherwise, with ε_k the measured detection efficiency of the kth layer.

As already mentioned, these likelihoods, detailed now for all sub-detectors participating in the PID determination, are basic ingredients of all PID variables that are employed. In the form of ratios, they define all high-level hID variables today in use at Belle II. There are actually three main families of variables that have been developed. The default approach relies on the assumption that sub-detectors likelihoods are independent, and thus can be all combined to obtain the full likelihood for a particle hypothesis x as:

$$\log \mathscr{L}_x = \log\left(\prod_{d\in D} \mathscr{L}_x^d\right) = \sum_{d\in D} \log \mathscr{L}_x^d. \tag{2.8}$$

In order to discriminate between the hypothesis x and all the other stable particle hypotheses, a ratio considering them all is built, and is given as:

$$P_i = \frac{\exp\left(\sum_{d\in D} \log \mathscr{L}_i^d\right)}{\sum_{x\in X} \exp\left(\sum_{d\in D} \log \mathscr{L}_x^d\right)} = \frac{\prod_{d\in D} \mathscr{L}_i^d}{\sum_{x\in X}\left(\prod_{d\in D} \mathscr{L}_x^d\right)} = \frac{\exp\left(\log \mathscr{L}_i\right)}{\sum_{x\in X} \exp\left(\log \mathscr{L}_x\right)}. \tag{2.9}$$

The ratio presented in Equation (2.9) is known as the "*Global*" PID variable. It is worth noting that if no information is available for a specific particle x and a detector d, i.e. $\mathscr{L}_x^d = \text{NaN}$, this particular likelihood is set to 1 and does not contribute. Another ratio, defined on the same principle, but relying on only two hypotheses, can be also expressed as:

$$P_{i \text{ vs } j} = \frac{\exp\left(\log \mathscr{L}_i\right)}{\exp\left(\log \mathscr{L}_i\right) + \exp\left(\log \mathscr{L}_j\right)}. \tag{2.10}$$

As two hypotheses are used, the variable built with the ratio computed in Equation (2.10) is known as the "*Binary*" PID variable.

The advantage of this approach, which directly uses the likelihoods, is its simplicity, since combining the information from all the sub-detectors is straightforward. It is also possible to further customise these variables by discarding a sub-detector that was not optimally calibrated in some data taking period, or by adding more particle hypothesis in the likelihood ratio, in case we want to discriminate among additional particle hypotheses.

There are nevertheless significant drawbacks, such as the fact that potential correlations between the sub-detectors likelihoods are not taken into account, or that some detectors might spoil the information provided by the other sub-detectors, or that the discrimination power of a sub-detector between two particle hypotheses x or y may not guarantee equal probabilities favoring x over y. The latter would indeed result in a slightly biased binary or global likelihood ratio as, in the default combination, one of the two sub-detectors might be unfairly weighted compared to the other. These reasons motivated the development of more advanced PID variables, such as the following two approaches.

In this second approach, calibration weights are applied to the individual likelihoods to overcome some of the limitations of the default approach. These weights $w_{x,d}$ are defined for each detector d and particle hypothesis x. Using Equation (2.8), a weighted likelihood can be computed as:

$$\log \mathscr{L}'_x = \log\left(\prod_{d\in D}\left(\mathscr{L}_x^d\right)^{w_{x,d}}\right) = \sum_{d\in D} w_{x,d}\log \mathscr{L}_x^d, \tag{2.11}$$

which gives the following ratio:

$$P'_i = \frac{\exp(\sum_{d\in D} w_{i,d}\log \mathscr{L}_i^d)}{\sum_{x\in X}\exp(\sum_{d\in D} w_{x,d}\log \mathscr{L}_x^d)} = \frac{\exp\left(\log \mathscr{L}'_i\right)}{\sum_{x\in X}\exp\left(\log \mathscr{L}'_x\right)}. \tag{2.12}$$

This variable is referred to as the "*Reweighted*" PID variable. The calibration weights used in Equation (2.11) and Equation (2.12) are organised in the form of a 6×6 matrix, expressed as:

$$w = \begin{array}{c} \begin{matrix} \text{svd} & \text{cdc} & \text{top} & \text{arich} & \text{ecl} & \text{klm} \end{matrix} \\ \begin{pmatrix} 0.80897 & 2.27021 & 0.43437 & 0.59494 & 2.46420 & 0.15824 \\ 1.36293 & 1.95849 & 0.42866 & 0.61061 & 1.96239 & 0.48289 \\ 1.06375 & 1.43302 & 0.42645 & 0.59816 & 2.01812 & 0.18273 \\ 1.79259 & 1.90825 & 0.42230 & 0.60463 & 1.66284 & 0.21942 \\ 1.71904 & 1.86573 & 0.41388 & 0.64555 & 1.96317 & 0.22059 \\ 1.26471 & 2.02610 & 0.38130 & 0.69305 & 4.08499 & 0.14939 \end{pmatrix} \end{array} \begin{matrix} e \\ \mu \\ \pi \\ K \\ p \\ d \end{matrix}$$

These weights are obtained after training a neural network (NN) on a sample of simulated single-particle events in which all six categories are equally represented. They are more precisely implemented as in a usual neural network, where the procedure aims at minimizing the cross entropy loss function. An additional binary cross entropy term is added to the loss function, in order to increase the separation between pions and the other particles, as they are the most abundantly produced particle and have a significant probability to be misidentified as leptons, kaons and protons. One observation is that setting all weights to 1 would lead to the default likelihood ratio.

A third and last approach that is employed to improve the default implementation relies on multivariate methods, such as neural networks or boosted decision trees (BDT).

The variable known as the "*BDT*" or also "*MVA*" PID variable was developed in a view of improving the separation between electrons, muons and pions. For that, the likelihood of the ECL, depending on the E/p ratio as detailed in Equation (2.6), is modified. Even if this variable is in general very powerful in discriminating electrons and muons against hadrons, such as pions, it shows a reduction of discriminating power towards low momentum ranges. Indeed, in that specific scenario,

electrons have a stronger bending of their trajectory, which leads to longer paths through material and therefore increasing energy losses from Bremsstrahlung, affecting the separation power between muons and electrons. The hadronic inelastic interactions rate is also higher in that case, which impacts the separation power between muons and pions.

To counter these effects, several variables related to the lateral and longitudinal shower shape development, and track penetration depth into the ECL are combined into a BDT to allow further discrimination. As these variables are expected to be highly correlated, the use of a BDT is preferable to provide a better handle at exploiting non-trivial dependencies across inputs to improve classification performance. The BDT is trained in multi-class mode to separate leptons from all other particle hypotheses, but it is also possible to train it in binary lepton vs pion mode. In order to exploit the particle ID capability in its entirety, these inputs are combined with the high-level likelihoods for the electron, muon and pion hypotheses from the other sub-detectors, and gives the following ratio of likelihood:

$$\tilde{P}_i = \frac{\exp(\sum_{d\in D} \log \tilde{\mathscr{L}}_i^d)}{\sum_{x\in X} \exp(\sum_{d\in D} \log \tilde{\mathscr{L}}_x^d)} = \frac{\exp(\log \tilde{\mathscr{L}}_i)}{\sum_{x\in X} \exp(\log \tilde{\mathscr{L}}_x)}, \tag{2.13}$$

where $\log \tilde{\mathscr{L}}_x$ is defined using Equation (2.8). Another emerging variable, called "*NN*" PID variable, uses neural networks to discriminate pions from kaons. Ongoing advancements aim to extend its capability to discriminate among all particle hypotheses. However, since this feature is still under development, additional details are currently lacking.

2.4 Trigger System and Data Acquisition

During Belle II operations, each sub-detector transmits its readout data upon receiving an external trigger signal, and the combined data from all sub-detectors triggered by a specific event constitutes what is known as "*one event*", that can be subsequently stored as data.

The trigger system operates on two levels: a hardware-based low-level trigger, the "*L1*" trigger, and a software-based high-level trigger, the "*HLT*" trigger.

The L1 trigger, responsible for generating trigger signals for significant collisions, processes a live stream of data from the CDC, ECL, and KLM sub-detectors, with the TOP sub-detector also contributing data, as shown in Figure 2.11. This streamed data undergoes near real-time processing using specialised fast electronics such as

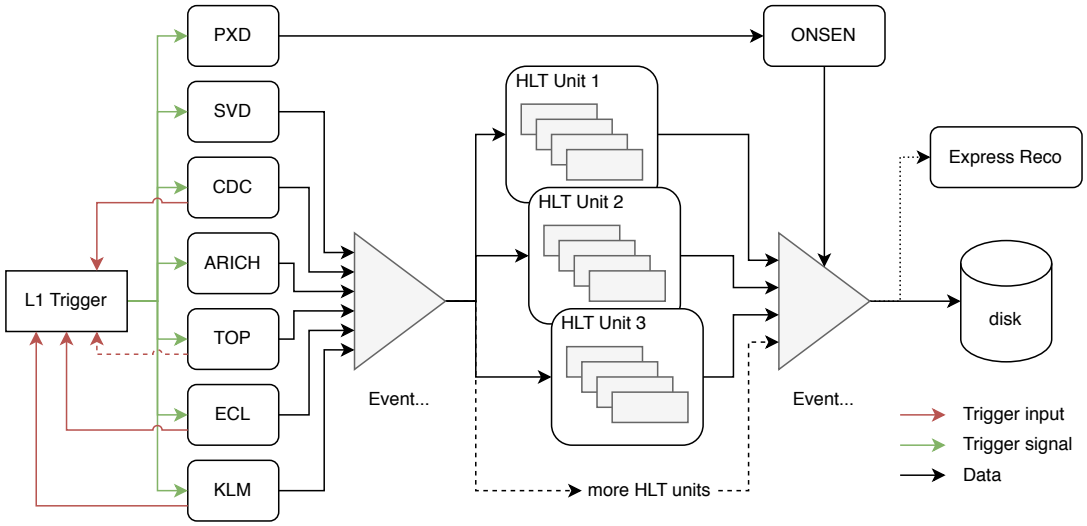


Figure 2.11: *Belle II data flow.* Simplified diagram that explains the relationship between the different component involved in the data acquisition, from [52].

Field Programmable Gate Arrays (FPGAs), and is continuously evaluated against predefined trigger conditions. Upon detection of an interesting collision event, the L1 trigger issues a signals to all sub-detectors, capable of generating up to 30 kHz of triggers at full SuperKEKB luminosity, with a latency of 5 microseconds. The DAQ system ensures synchronised delivery of trigger signals to all sub-detectors and facilitates high-speed data transfer to the HLT system, responsible for reading out detector L1 trigger signals. Additionally, the Online Selection Nodes (ONSEN) system reduces background data [53], and the particle collision events are quickly reconstructed by a system, referred to as Express Reco, while the data is seamlessly transferred to storage systems.

Efficient operation of these triggers poses a significant challenge, particularly given the substantially high background rates, detailed in Section 2.1.1, that are expected from SuperKEKB. To prevent data storage overload, it is sometimes necessary to introduce a pre-scale factor, which will further decrease the rate at which the events are acquired. These pre-scale factors can be associated to the triggers bits for a certain run period, and are not necessarily the same for all of them.

Further trigger combinations, called also "*trigger bits*", have been developed throughout the Belle II operation; there exists now a wide trigger assortment that covers the needs of almost all analyses. In the context of this thesis, ECL- and CDC-based trigger bits are deemed preferable, as they are able to trigger low-multiplicity events and decay topologies with several tracks within their detection range. It is worth noting that the triggger bits respect a naming convention, that is summarised in Table 2.2.

Name	Brief description
f	There is a CDC full 2D tracks, which requires 4 clustered hits in 5 axial super layers (0, 2, 4, 6, 8).
s	There is a CDC short track, which requires 5 clustered hits in 5 inner most super layers (0, 1, 2, 3, 4).
y	There is a CDC full 3D track, which requires -40 cm $< dz <$ 40 cm, measured by the Neural Network 3D module.
z	There is a CDC full 3D track measured by conventional 3D module.
o	There is an opening angle of CDC tracks ϕ bigger than 90°.
b	There is an opening angle of CDC tracks ϕ bigger than 150°.
hie, *lowe*	The energy sum of ECL clusters in barrel and part of forward endcap passes the high or low energy threshold.
$lmlx$, $x \in 0,...,13$	There are ECL low energy multi clusters correspond to the x combination.
cx, $x \in 1,...,5$	The number of ECL cluster is more than or equal to x.

Table 2.2: *Trigger naming convention.* These letters or group of letters are used to name the different trigger bits that are available. Not all possible names are shown. The number of instances represents the number of tracks. For instance, *ff* signifies two CDC full 2D tracks.

2.5 Real Data and Monte Carlo Simulation

Belle II, just as any other high-energy physics experiment, involve a complex process from data collection to publication. The analysis itself takes months or years, and could not be completed without real data processing, usually referred to simply as "*data*", and Monte Carlo (MC) simulations. These two key points result from the simplified process described by the diagram shown in Figure 2.12.

Real data are obtained from the detector through inputs from various sub-detectors, described in Section 2.2, processed by the data acquisition (DAQ) system to convert signals into usable computer values, as mentioned in Section 2.4. Monte Carlo simulation involves event generation and simulation to mimic detector behavior. Reconstruction is then performed on both real data and Monte Carlo simulation to minimise differences, with an additional calibration step needed for real data.

Analysts use the Belle II analysis software framework, called "*basf2*" [54, 55] to

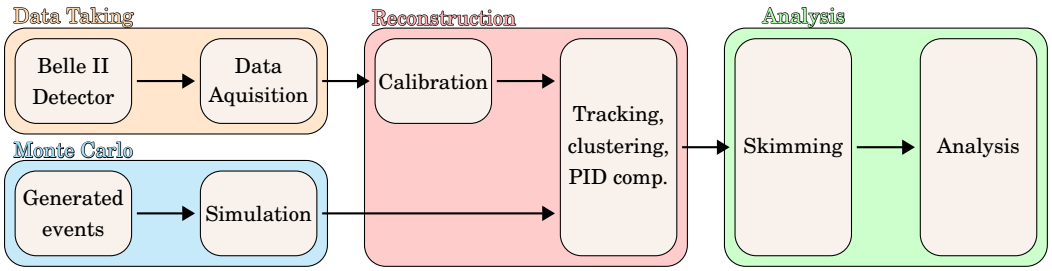


Figure 2.12: *Belle II Data scheme.* Making of data and Monte Carlo samples.

reconstruct events. It is an open source framework based on C++ and Python modules, which evolves over time through releases, each identified by a keyword and the corresponding month and year of accessibility. It also contains useful tools for data analysis. The latter is performed on reconstructed events, on top of which a skim can be applied to reduce further the data set size. This helps streamlining and accelerating analysis processes.

Having provided a rather general overview regarding Monte Carlo simulations and real data, the following two sections are dedicated to providing more detailed insights about the different steps happening before event reconstruction, in the context of Belle II.

2.5.1 Real Data

Real data undergo two primary processing stages: prompt processing and reprocessing; the distinction between them lies in the extent of calibrations applied during processing.

"*Prompt*" processing occurs immediately after data collection: Data is gathered over a fixed period, typically two weeks, then calibrated and processed iteratively. The dataset resulting from this operation is referred to as a "*Bucket*". Prompt data always utilises the latest major software release and undergoes thorough calibration, with quality comparable to reprocessed data. Consequently, most calibrations performed during prompt processing are retained for reprocessing.

Reprocessing involves recalibrating and reprocessing all data collected up to a specific time. Reprocessing, labeled as "*Proc*" and representing a total of five "*Chunks*", occurs annually (or biennially from 2023). Recalibration starts from prompt calibration and is refined with improved or new calibration algorithms. Full reprocessing utilises a major software release, ensuring a coherent processing of the entire data set. Both prompt and reprocessed data sets are accessible on the grid via corresponding collections, stored in the micro data summary tapes (mDST) format.

2.5.2 Monte Carlo Simulation

As previously discussed, event generation and simulation are two distinct processes that take place before the event reconstruction.

Event generation involves simulating particle collisions based on the initial conditions of electron and positron collisions, according to specific physics models like Standard Model, SUSY or dark matter. Based on that choice, the particles position and four-vectors are generated, which takes a few milliseconds per event at Belle II. Various generators like KKMC [56, 57], PYTHIA [58], BabaYaga@NLO [59, 60, 61, 62, 63], TAUOLA [64, 65, 66, 67, 68], PHOTOS [69], AAFH [70, 71, 72], TREPS [73] or even EvtGen [74] cater to different physics processes and analyses needs. Table 2.3 summarises the generators of the main e^+e^- processes observed at Belle II.

These processes can be described by their Feynman diagrams, which are illustrated in Figure 2.13. The order of the diagram gives a sense of the order of magnitude of the contribution of the process.

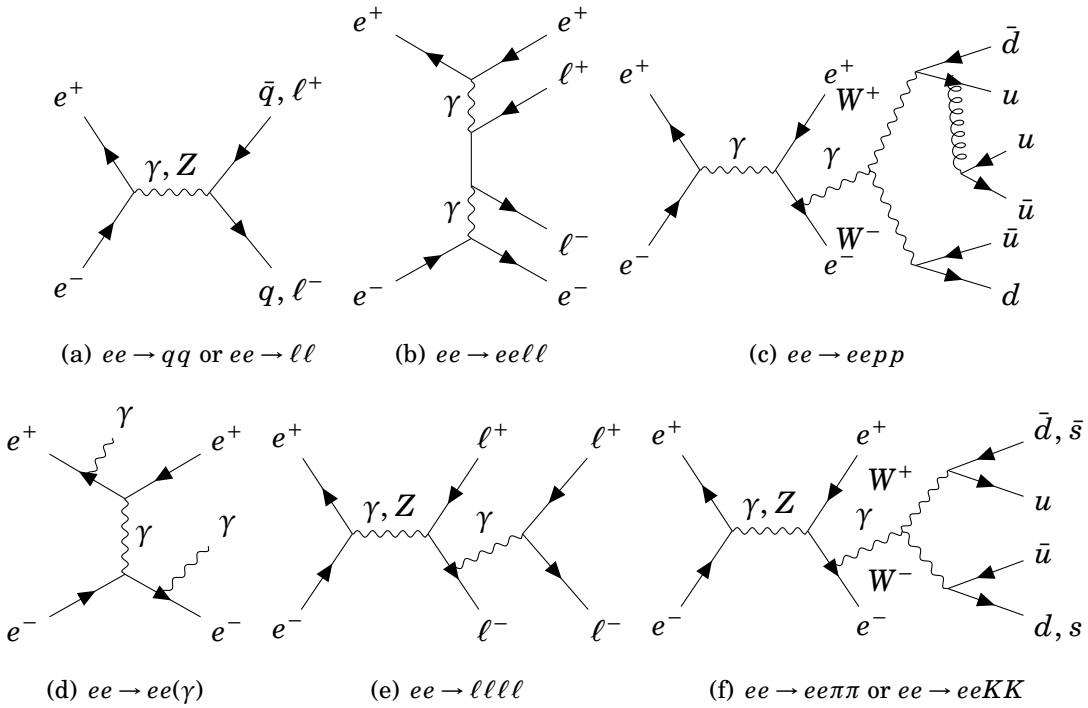


Figure 2.13: *Feynman diagram of the main ee processes observed at Belle II.* Only one representation per process is given.

A generator called Madgraph [75] is also available, and is particularly useful for dark sector analysis. It generates and simulates specific signals as defined by the user.

Process	σ [nb]	Selection Criteria	Generator
$c\bar{c}$	1.30	-	KKMC, PYTHIA
$d\bar{d}$	0.40	-	KKMC, PYTHIA
$s\bar{s}$	0.38	-	KKMC, PYTHIA
$u\bar{u}$	1.61	-	KKMC, PYTHIA
$e^+e^-(\gamma)$	300 ± 3	$10° < \theta_e^* < 170°$, $E_e^* > 0.15$ GeV	BABAYAGA.NLO
$e^+e^-(\gamma)$	74.4	$p_e > 0.5$ GeV/c and e in ECL	-
$\mu^+\mu^-(\gamma)$	1.148	-	KKMC
$\mu^+\mu^-(\gamma)$	0.831	$p_\mu > 0.5$ GeV/c in CDC	-
$\mu^+\mu^-\gamma(\gamma)$	0.242	$p_\mu > 0.5$ GeV/c in CDC, $\geq 1\gamma$ ($E_\gamma > 0.5$ GeV) in ECL	-
$\tau^+\tau^-(\gamma)$	0.919	-	KKMC, TAUOLA, PHOTOS
$e^+e^-e^+e^-$	39.7 ± 0.1	$W_{\ell\ell} > 0.5$ GeV/c^2	AAFH
$e^+e^-\mu^+\mu^-$	18.9 ± 0.1	$W_{\ell\ell} > 0.5$ GeV/c^2	AAFH
$e^+e^-\tau^+\tau^-$	0.01836	-	KKMC
$e^+e^-p^+p^-$	0.0117	-	TREPS
$e^+e^-K^+K^-$	0.0798	-	TREPS
$e^+e^-\pi^+\pi^-$	1.895	-	TREPS
$B^0\bar{B}^0$	0.51	-	EvtGen
B^+B^-	0.54	-	EvtGen

Table 2.3: *Generators of the main ee processes observed at Belle II.* The cross section as well as selection criteria, if any, are indicated. $W_{\ell\ell}$ is the minimum invariant secondary fermion pair mass, as detailed in [37].

Once the four-vectors are generated, the subsequent stage involves replicating the real detector's output, a process known as simulation. The real detector captures various interactions of particles with its materials, encompassing ionisation, scintillation, Bremsstrahlung, pair production, and Cherenkov radiation. Simulation software has been developed to emulate these processes, with GEANT4 [76] emerging as the foremost choice. GEANT4 models how particles engage with a virtual Belle II detector, resulting in energy deposition and particle generation in each sub-detector. Simulating the entire detector is computationally intensive, typically requiring around a second for Belle II.

The simulation process can be further categorised into two main types: "*run-inde-*

pendent" and "*run-dependent*" Monte Carlo simulations. The distinction between these lies in the fact that run-dependent simulations aim to replicate the background conditions corresponding to various run periods in the data. Following the simulation, an mDST file is generated, akin to those produced for real data.

2.6 τ Physics at Belle II

Even though Belle II is mostly known as a B-factory, it is also *de facto* a τ-factory; the production cross section of tau lepton pairs $\sigma(e^+e^- \to \tau^+\tau^-)$, highlighted in Figure 2.14, is equal to 0.919 nb, which is approximately the same as B meson pairs $\sigma(e^+e^- \to B\bar{B})$, which is equal to 1.05 nb.

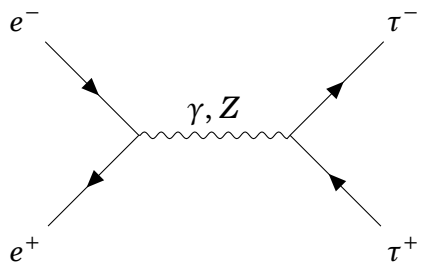


Figure 2.14: *Feynman diagram of the process* $e^+e^- \to \tau^+\tau^-$. An annihilation diagram (s-channel) is required to get two tau leptons in the final state.

Therefore, and not just in the interest of this thesis, it is pertinent to provide an overview of τ physics and detail the fundamental properties of the tau lepton.

The tau lepton distinguishes itself from the other leptons by its mass [19], as reported in Equation (2.14). Additionally, it emphasises the comparatively greater uncertainty in the determination of the tau lepton mass, which exhibits an uncertainty magnitude several orders larger than that of other leptons.

$$\begin{aligned} m_e &= (0.51099895000 \pm 0.00000000015)\ \text{MeV}/c^2 \\ m_\mu &= (105.6583755 \pm 0.0000023)\ \text{MeV}/c^2 \\ m_\tau &= (1776.86 \pm 0.12)\ \text{MeV}/c^2 \end{aligned} \tag{2.14}$$

Because its mass is so large, the tau lepton can decay not just into the two other leptons, but also into hadrons, as depicted by the Feynman diagrams in Figure 2.15.

These two examples are part of a set of many other decays. The resulting final states are well-documented and understood. Most of them have a branching fraction measurement, which is shown in Figure 2.16.

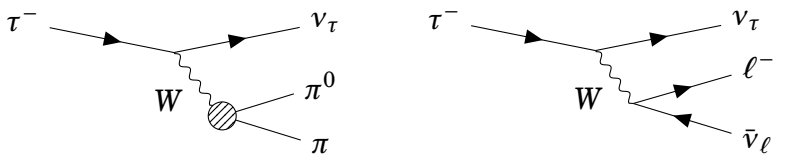


Figure 2.15: *Feynman diagram of the processes* $\tau^- \to X^-$. Here, X is either a hadronic system, *e.g.*, $\pi\pi^0$, or a leptonic system, *e.g.*, $\ell\nu_\ell$.

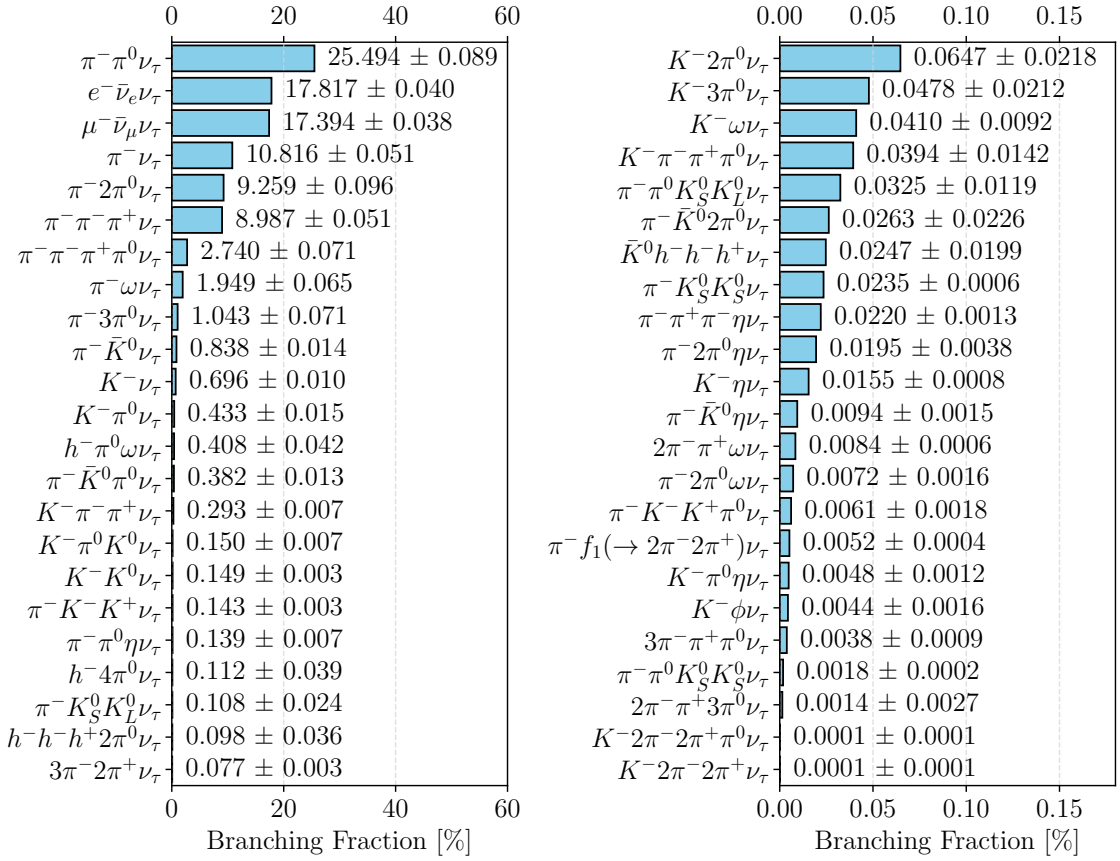


Figure 2.16: *Tau decay modes.* All observed tau decays final states, sorted according to their respective values, from top to bottom and left to right. The representation is divided into two parts to ease the readability. The values are taken from [19].

As previously shown in Figure 2.14, the tau lepton is produced in pairs in e^+e^- collisions. In the CMS, the tau lepton pairs are produced back-to-back. This particularity allows the definition of two hemispheres in space, which are separated by the so-called "*thrust*" axis, depicted in Figure 2.17. The latter corresponds to the maximal momentum projection of all visible particles, from which the thrust value

can be further derived as:

$$\text{thrust} \overset{\text{max}}{=} \sum_i \frac{|\vec{p}_i^{\,\text{CMS}} \cdot \hat{n}_{\text{thrust}}|}{\sum |\vec{p}_i^{\,\text{CMS}}|}, \tag{2.15}$$

where $\vec{p}_i^{\,\text{CMS}}$ is the momentum in the CMS of the ith particle and $\hat{n}_{\text{thrust}}$ is the unit vector in the direction of the thrust axis. The thrust value, by definition, is dimensionless. Hence, according to Equation (2.15), tau lepton pairs are expected to exhibit a thrust value nearing 1.

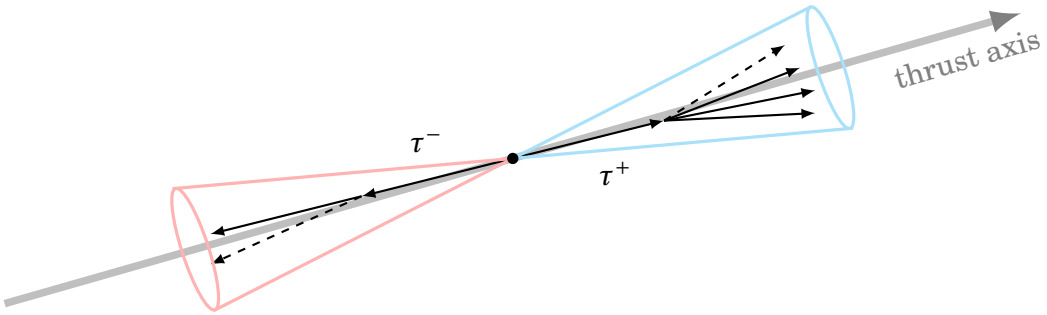


Figure 2.17: *Thrust axis sketch.* Illustration of a case where the tau lepton pair decays into one and three charged particles.

In general, due to its well-established production mechanisms and decay properties governed by the weak force, τ physics is convenient to conduct precision measurements within the Standard Model framework.

3. Physics Performance

As explained in Section 2.5, simulations are used to study real data collected by detectors. Even if their accuracy is generally acceptable, it is impossible to simulate data identically. Some features will remain imperfect, and will affect variables such as the ones related to particle identification (PID). To correct the simulations, a correction factor is established by comparing the PID efficiency between data and simulations of a relatively pure particle sample.

This chapter presents a performance study which aims at computing correction factors for pion identification (πID) using hadronic tau decays. A specific emphasis is put on a low multiplicity channel as the already available corrections were solely determined on high multiplicity channels, such as $D^{*+} \to D^0(K^-\pi^+)\pi^-$, sometimes abbreviated as "D^*" and $K_s \to \pi^+\pi^-$, also referred to as "K_s". And the use of a tau decay satisfies a topology closer to that studied in Section 4. Charge conjugation is always implied, except where explicitly stated otherwise.

3.1 Methodology Overview

A particular tau decay topology resulting from the electron-positron collision is chosen: given the two tau leptons produced in the collision, one is required to decay into one charged track (labeled as "*1-prong*"), while the other tau decays into three charged pions (labeled as "*3-prong*").

This "*3×1-prong*" topology, along with a corresponding Feynman diagram, is illustrated in Figure 3.1. With this distinctive event shape, it is possible to select tau lepton pair events on data with a low level of background from other processes. The signal side is defined as the 3-prong side, and corresponds to the decay $\tau^+ \to \pi^+\pi^-\pi^+\nu_\tau$. The tag side, on the other hand, is the 1-prong side, and corresponds to $\tau^- \to X^-\nu_\tau$, with X being any charged particle system.

A tag-and-probe approach is used to get a pure sample of pions. It consists of applying tight requirements on the identification variable of two of the pion tracks to

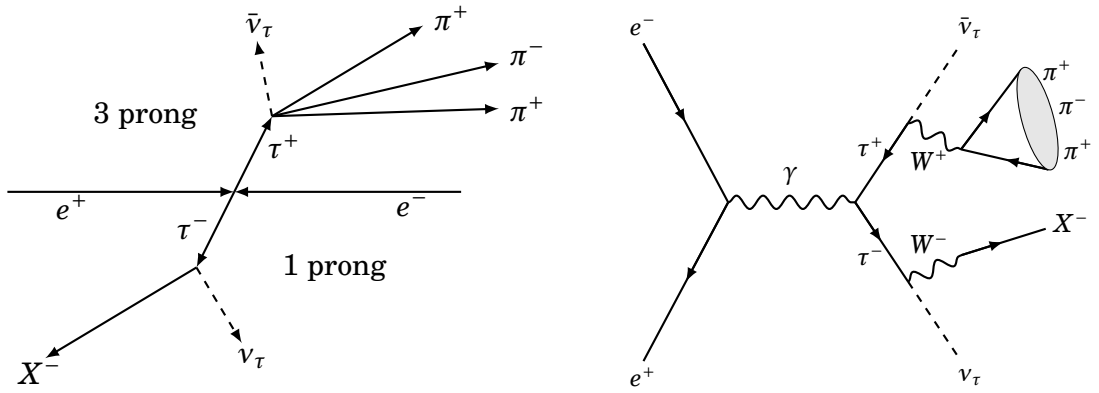


Figure 3.1: *Diagrams of targeted processes.* Illustration of the 3×1 prong topology and Feynman diagram for the $e^+e^- \to (\tau^+ \to \pi^+\pi^-\pi^+\bar{\nu}_\tau)(\tau^- \to X^-\nu_\tau)$ process targeted by this study.

tag a third charged pion used as a probe track. This method is motivated by the fact that the decay $\tau^+ \to \pi^+\pi^+K^-$ is highly suppressed, as its branching fraction is of the order of 10^{-11}. Hence, the probability of mistagging a kaon as a pion in the probe track gets negligible up to $\pi \to K$ fake rates in the tag tracks.

This study had to undergo several iterations and adjustments to accommodate evolving data sets. Multiple revisions and corrections were thus needed, and are promptly detailed here to give a context to this study.

It began by examining Monte Carlo samples, corresponding to the Belle II 14^{th} official campaign of run independent Monte Carlo, referred to as "*MC14ri*", to establish a selection procedure. This procedure was explored through two distinct approaches. The first method involved employing a series of rectangular cuts applied sequentially, with additional optimisation steps implemented. And the second approach utilised a Boosted Decision Tree (BDT), a multivariate method for event selection. Given that the BDT selection demonstrated only slight improvements of selection performance in comparison to the rectangular cut one, only the latter has been retained as the preferred selection procedure. At this stage of the study, tables of correction factors, binned in momentum p and cosine of polar angle θ were produced. These factors correspond to the ratio of efficiency in data over Monte Carlo, computed with the reprocessed data matching the same release as Monte Carlo, called "*Proc12*", and Prompt data sets.

These factors did not contain systematic uncertainties yet, only statistical ones. The systematic uncertainties were evaluated only with the next and currently latest iteration of Monte Carlo samples, which corresponded to the Belle II 15^{th} official campaign of run-independent and run-dependent Monte Carlo, called "*MC15ri*" and

"*MC15rd*" respectively, and the corresponding reprocessed "*Proc13*" and Prompt data sets.

The use of run-dependent Monte Carlo samples prompted a reevaluation of certain selection choices done for this study, emphasizing its integration into the *Systematic Corrections Framework*. This framework is a tool that can provide analysts the uncertainties associated to the particle identification variables they use. It is still under development, but already provides corrections, *e.g.*, hadron identification (ID) and hadronic misidentification (misID) rates. One advantage of having corrections already available is that it is possible to assess the consistency of the new corrections by making comparisons, which is done in Section 3.5.2.

The integration in the *Systematic Corrections Framework* consists of several steps, the first of which is to reconstruct the events according to the optimised selection. This step is, at the time of the thesis, about to be completed. The second step, which has been delayed for the time being, is to calculate the efficiencies. As it happens, the efficiency calculation based on the number of events is not yet implemented in the framework, and only calculations based on a peak fit are integrated. Which is why full implementation has been put on hold.

3.2 Input Samples

For this study, the data collected by Belle II from 2019 to 2022 was used. Called also "*LS1*" data set, it corresponds to a total integrated luminosity of 364.093 fb^{-1}. LS1 can be further separated into Proc13 and Prompt data, and into Chunks and Experiments as detailed in Table 3.1, along their corresponding integrated luminosity.

The simulated samples are taken from the official Belle II 15^{th} Monte Carlo campaign of run dependent samples. They correspond to generic and low multiplicity samples, and are detailed in Table 3.2, along their cross section, integrated luminosity and number of events. The generators used to get these samples are described in Section 2.5.2.

3.3 Event Selection

In order to accurately determine the correction factors linked to a specific particle species, it is essential to establish a sample with high purity. For this reason, the decay $\tau^+ \to \pi^+\pi^-\pi^+\nu_\tau$ was not chosen at random, since it represents the largest

Sample		Experiments	$\int \mathscr{L} \mathrm{d}t$ [fb^{-1}]
Proc13	Chunk 1	7	$0.508 \pm 0.002 \pm 0.004$
		8	$1.673 \pm 0.003 \pm 0.011$
		10	$3.703 \pm 0.002 \pm 0.024$
	Chunk 2	12	$55.277 \pm 0.005 \pm 0.348$
	Chunk 3	14	$16.657 \pm 0.006 \pm 0.105$
	Chunk 4	16	$10.449 \pm 0.004 \pm 0.066$
		17	$10.866 \pm 0.005 \pm 0.069$
	Chunk 5	18	$91.296 \pm 0.011 \pm 0.575$
Prompt		20	$3.854 \pm 0.003 \pm 0.024$
		22	$32.578 \pm 0.007 \pm 0.205$
		24	$86.805 \pm 0.011 \pm 0.547$
		26	$55.646 \pm 0.009 \pm 0.351$
Total			$364.093 \pm 0.021 \pm 2.331$

Table 3.1: *Proc13 and Prompt data set.* Samples are defined with Chunk and Experiment numbers. Their luminosity is indicated in the last column, with the central value, and the statistical and systematic uncertainties.

proportion of 3-charged particles decays, as reported in Figure 2.16, and amounts to 8.987% in total. The second highest 3-charged particles decay is $\tau^+ \to \pi^+\pi^-\pi^+\pi^0\nu_\tau$ with 2.740% and the third is $\tau^+ \to \pi^+(\omega \to \pi^-\pi^+\pi^0)\nu_\tau$ with 1.949%, containing also pions.

The selection of events has to not only yield a high purity, but also an efficiency high enough to result in relatively low statistical uncertainties. As these correction factors are computed using the whole LS1 data set and corresponding Monte Carlo simulations, representing respectively 364.093 fb^{-1} and 1455.052 fb^{-1} for generic samples, the efficiency does not represent a major challenge, and allows selection optimisation to focus more on purity.

The cuts employed for the selection of the signal are applied during the reconstruction of the decay, but also after this first selection to adjust more precisely the cuts. These two steps are often called "*online*" and "*offline*" selections, but in what follows they are defined as "*preselection*" and "*selection cuts*".

Process	σ [nb]	$\int \mathscr{L} \mathrm{d}t$ [fb^{-1}]	N [10^6]
$e^+e^- \to \tau^+\tau^-$	0.919	1455.052	1337.19
$e^+e^- \to c\bar{c}$	1.329	1455.052	1933.76
$e^+e^- \to d\bar{d}$	0.401	1455.052	583.48
$e^+e^- \to s\bar{s}$	0.383	1455.052	557.28
$e^+e^- \to u\bar{u}$	1.605	1455.052	2335.36
$e^+e^- \to B^+B^-$	0.54	1455.052	785.73
$e^+e^- \to B^0\bar{B}^0$	0.51	1455.052	742.08
$e^+e^- \to e^+e^-(\gamma)$	295.8	36.3731	10759.16
$e^+e^- \to \mu^+\mu^-(\gamma)$	1.148	1455.052	1670.40
$e^+e^- \to e^+e^-e^+e^-$	39.55	363.767	14386.98
$e^+e^- \to e^+e^-\mu^+\mu^-$	18.83	363.767	6849.73
$e^+e^- \to e^+e^-\tau^+\tau^-$	0.01836	363.767	6.68
$e^+e^- \to e^+e^-K^+K^-$	0.0798	363.767	29.03
$e^+e^- \to e^+e^-\pi^+\pi^-$	1.895	363.767	689.34
$e^+e^- \to e^+e^-p^+p^-$	0.0117	363.767	4.26
$e^+e^- \to \mu^+\mu^-\tau^+\tau^-$	1.441×10^{-4}	363.767	5.24×10^{-2}
$e^+e^- \to \tau^+\tau^-\tau^+\tau^-$	2.114×10^{-7}	363.767	7.69×10^{-5}
$e^+e^- \to K^+K^-(\gamma)$	0.0163	363.767	5.93
$e^+e^- \to K^0\bar{K}^0(\gamma)$	0.008864	363.767	3.22
$e^+e^- \to \pi^+\pi^-(\gamma)$	0.1667	363.767	60.64
$e^+e^- \to \pi^+\pi^-\pi^0(\gamma)$	0.02378	363.767	8.65

Table 3.2: *Simulated processes with their cross section, corresponding integrated luminosity and number of events.* The generators used to get these samples are described in Section 2.5.2.

3.3.1 Selection Assessment

The performance of the events selection is assessed by two factors: the purity p and the efficiency ε, which are expressed as:

$$\mathrm{p} = \frac{N_{\mathrm{sig}}}{N_{\mathrm{sig}} + N_{\mathrm{bkg}}} \tag{3.1}$$

and

$$\varepsilon = \frac{N_{\text{sig}}}{2 \cdot N^{\text{gen}} \cdot \mathscr{B}_{3p} \cdot \mathscr{B}_{1p}}, \tag{3.2}$$

where N_{sig} is the number of signal events, N_{bkg} the number of background events, $N^{\text{gen}} = \sigma_{\tau\tau} \cdot \int \mathscr{L} \mathrm{d}t = N_{\tau\tau}$ the number of generated tau lepton pairs events indicated in Table 3.2, and $\mathscr{B}_{3p} = 9.31\%$ and $\mathscr{B}_{1p} = 84.58\%$ the branching fraction [19] of the 3-prong and 1-prong side respectively.

Combined, they evaluate how pure the sample is, and how large it remains after the selection. A figure of merit (FOM) is needed to optimise the event selection according to these two criteria, and is expressed in this case as:

$$\text{FOM} = \frac{N_{\text{sig}}}{\sqrt{N_{\text{sig}} + x \cdot N_{\text{bkg}}}}, \tag{3.3}$$

where x is a penalty coefficient that encourages the optimisation to favour purity over efficiency. If not specifically mentioned, this coefficient it set to 100. This value was chosen to enhance the purity of the sample.

3.3.2 Selection Variables

Before delving into the selection process, it is essential to establish clear definitions for the variables involved. These variables must satisfy two criteria: first, they need to effectively isolate the signal while discriminating against background, and second, they should be accurately modeled to minimise discrepancies between data and Monte Carlo simulations.

One first variable that fulfills these criteria is the thrust, which was previously introduced in Section 2.6, Equation (2.15), and which is in particular useful to distinguish tau lepton pair, $q\bar{q}$ and Bhabha events. A variable related to the thrust is the cosine of the polar angle component of the thrust axis ($\cos\theta_{\text{thrust}}$). To enhance the differentiation between various processes, it is necessary to incorporate events associated with both shape and kinematics. For that, the polar angle component of the missing momentum of the event in the CMS frame ($\theta_{p_{\text{miss}}^{\text{CMS}}}$), the missing mass squared of the event (m^2_{miss}) and the visible energy of the event in the CMS frame ($E^{\text{CMS}}_{\text{vis}}$) particularly suit well this condition. Another type of variables, constructed with event and shape parameters, are related to the Fox-Wolfram moments [77, 78]. These moments are rotationally-invariant parametrisations of the distribution of particles in an event. Their normalised moments, noted as R_ℓ, with $\ell = 1,...,4$, are more usually used, and are defined as

$$R_\ell = \frac{H_i}{H_0}, \tag{3.4}$$

with

$$H_\ell = \sum_{i,j} p_i p_j P_\ell(\cos\theta_{i,j}), \tag{3.5}$$

where p_k are the momentum vectors, P_ℓ is the Legendre polynomial of ℓ-th order and $\theta_{i,j}$ is the angle between p_i and p_j.

Concerning the tag side variables, only the transverse momentum, denoted as $p_{T,\tau,\text{tag}} \equiv p_{T,\text{tag}}$, is considered. It is computed as

$$p_{T,\text{tag}} = \sqrt{p_{x,\text{tag}}^2 + p_{y,\text{tag}}^2}, \tag{3.6}$$

where $p_{x,\text{tag}}$ is the tag side momentum projection on the X-axis and $p_{y,\text{tag}}$ is the tag side momentum projection on the Y-axis. As for the signal side, the momentum in the CMS frame ($p_{\tau,\text{sig}}^{\text{CMS}}$), energy ($E_{\tau,\text{sig}}$) and mass ($m_{\tau,\text{sig}}$) are providing direct information for selecting pions on the signal side. The result of the chi-square statistical test of the track fit, ($\chi^2_{\tau,\text{sig}}$) is, in turn, useful for further constraints.

The number of neutrals are useful variables, which are notably also used to build the events during the reconstruction. More precisely, the number of neutral pions in the signal and tag sides ($n_{\pi^0,\text{sig}}$ and $n_{\pi^0,\text{tag}}$) and the number of photons in the signal and tag sides ($n_{\gamma,\text{sig}}$ and $n_{\gamma,\text{tag}}$) are used.

3.3.3 Particle Identification (PID)

As the particle identification performance is the main focus of this study, it is necessary to first select a particle identification variable for which the correction factors are computed. Belle II offers several types of particle identification variable, as detailed in Section 2.3. Three types of variables are selected: The Global PID, corresponding to Equation (2.9), the Reweighted PID, detailed in Equation (2.12) and the BDT PID, defined by Equation (2.13). To facilitate readability, previously introduced equations are reiterated here for the reader's convenience. Considering $D \equiv \{$ SVD, CDC, TOP, ARICH, ECL, KLM$\}$ a set of sub-detectors and $X \equiv \{e, \mu, \pi, K, p, d\}$ a set of long-lived particles, we get:

$$\begin{cases} \text{Glob.} \equiv P_i = \dfrac{\exp\left(\log \mathscr{L}_i\right)}{\sum\limits_{x\in X}\exp\left(\log \mathscr{L}_x\right)} = \dfrac{\exp(\sum_{d\in D}\log \mathscr{L}_i^d)}{\sum_{x\in X}\exp(\sum_{d\in D}\log \mathscr{L}_x^d)} \\ \text{Rew.} \equiv P'_i = \dfrac{\exp\left(\log \mathscr{L}'_i\right)}{\sum\limits_{x\in X}\exp\left(\log \mathscr{L}'_x\right)} = \dfrac{\exp(\sum_{d\in D} w_{i,d}\log \mathscr{L}_i^d)}{\sum_{x\in X}\exp(\sum_{d\in D} w_{x,d}\log \mathscr{L}_x^d)} \\ \text{MVA} \equiv \tilde{P}_i = \dfrac{\exp(\log \tilde{\mathscr{L}}_i)}{\sum\limits_{x\in X}\exp(\log \tilde{\mathscr{L}}_x)} = \dfrac{\exp(\sum_{d\in D}\log \tilde{\mathscr{L}}_i^d)}{\sum_{x\in X}\exp(\sum_{d\in D}\log \tilde{\mathscr{L}}_x^d)} \end{cases} \tag{3.7}$$

The correction associated to particle identification are usually binned in momentum p and polar angle θ. This choice is motivated by the unequal population of particle species within the phase space. Implementing a global factor could potentially lead to larger uncertainties, as it would need to accommodate variations across both favorable and unfavorable regions. In addition, the polar angle bins used for hadrons, in particular for pions and kaons, take into account the different sub-detector regions, while remaining more or less equal in size. Their coverage spans from 17° to 150°. The details regarding the momentum and polar angle bins can be found in Table 3.3. A sketch of the Belle II detector from the side with the regions corresponding to the polar angle bins is shown in Figure 3.2. If not otherwise specified in the following sections of this chapter, this binning is the one used by default for all aspects related to particle identification corrections.

momentum	polar angle			
p [GeV/c]	θ [rad]	θ [deg]	$\cos\theta$ [-]	number
[0.5, 1]	[0.297, 0.489]	[17, 28]	[0.883, 0.956]	8
[1, 1.5]	[0.489, 0.698]	[28, 40]	[0.766, 0.883]	7
[1.5, 2]	[0.698, 1.047]	[40, 60]	[0.500, 0.766]	6
[2, 2.5]	[1.047, 1.344]	[60, 77]	[0.225, 0.500]	5
[2.5, 3]	[1.344, 1.676]	[77, 96]	[-0.104, 0.225]	4
[3, 3.5]	[1.676, 2.007]	[96, 115]	[-0.423, -0.104]	3
[3.5, 4.5]	[2.007, 2.321]	[115, 133]	[-0.682, -0.423]	2
	[2.321, 2.618]	[133, 150]	[-0.866, -0.682]	1

Table 3.3: *Momentum and polar angle binning.* This binning is used for both pion and kaon corrections. The polar angle bins are given in radian, degrees and as their cosine value.

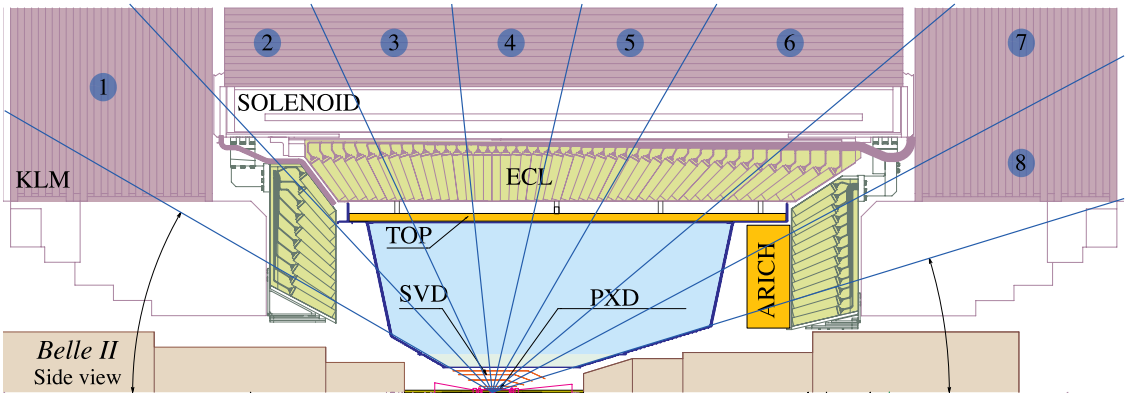


Figure 3.2: *Side view of the Belle II detector and sub-detector systems.* The different sub-detector systems are projected on the YZ-plan. The sketch is coming from [45]. The blue markers correspond to the polar angle bins defined in Table 3.3.

3.3.4 Preselection

Events are reconstructed using charged tracks originating from the interaction region, which is delimited by restricting the radial and axial impact parameters as $dr < 1.0$ cm and $|dz| < 3.0$ cm. Events are required to contain exactly four charged tracks, and their sum is required to be equal to zero.

Neutral particles are divided into neutral pion candidates and photons not coming from a neutral pion. Neutral particles are used to compute event shape and kinematics variables for the above-mentioned charged tracks.

Four different corrections are applied on the events. The correction applied on data includes a tracking momentum scale factor and a photon energy bias, and the correction applied on Monte Carlo is composed of a correction of the photon efficiency and a correction of the neutral pion efficiency. In particular, the two corrections applied on Monte Carlo samples are intended for application within the same Monte Carlo campaign, but on run independent samples. Despite this difference, they provide a highly accurate approximation.

A skim is applied on the event to reduce the data set size and accelerate the reconstruction of the events, as explained in Section 2.5.

3.3.5 Selection Cuts

To fulfill the requirements of the tag-and-probe method, a tight particle identification cut of 0.9 is applied on the tag tracks composing the signal. The PID variable

used in this case corresponds to the Global PID detailed in Equation (3.7). Following the recommendation for any performance study using hadrons, the number of hits recorded by the CDC is required to be higher than 20 for both tag tracks. The trigger selection corresponds to the use of any trigger bits available.

An optimised cut selection has been chosen, and the underlying principle is the following: first, a set of discriminating variables is defined. Variables are used one by one, and for each of them, the range of the cut applied on it is defined by the maximum of the figure of merit computed in Equation (3.3). In some cases, the cut is adjusted in the pursuit of higher purity. The cut is then applied, before moving on to the next variable in the list. These steps are repeated until all variables have been used.

An algorithm to refine these cuts is put in place to enhance the purity in a systematic way. It is based on the assumption that a better figure of merit can be attained in the close neighbourhood of the initial range. And so by shifting one or both endpoints to a lower or upper value of a width that is decreasing up until a limited value, this new range provides a better overall performance. The final set of cuts is reported in Table 3.4

The composition of the remaining Monte Carlo samples after the selection is reported in Table 3.5, along the result of the truth-matching of the probe track in all samples. The pion fraction amounts to 98.749% in the probe track, which represents the purity of the sample. The largest contamination of other sample is coming from $q\bar{q}$ events, which represents 0.904% of the events. And the two largest truth-matched particles after the pions are attributed to muons with 0.284% and kaons with 0.254%.

3.4 Measurement of Correction Factors

The corrections for both πID efficiency and $\pi \to K$ misidentification rates are calculated across several working points ranging from 0.05 to 0.95. Each measurement includes a central value along with statistical and total systematic uncertainties. The systematic uncertainties are assessed with a conservative approach to facilitate the integration of this study in the *Systematic Corrections Framework*. Four sources have been identified, taking into account tag side modeling, trigger selection, dominant fake rates and non-tau lepton pair backgrounds. A particular emphasis will be placed on the πID efficiency in Section 3.5, and in particular on its correction tables.

	Feature	Cut	Unit
1	$p_{t,\text{tag}}$	$\in [0.3101, 5.31]$	GeV/c
2	$E_{\text{vis}}^{\text{CMS}}$	$\in [2.4101, 8.3001]$	GeV
3	$\theta_{p_{\text{miss}}^{\text{CMS}}}$	$\in [0.55101, 2.6]$	rad
4	$E_{\tau,\text{sig}}$	$\in [2.481, 6.511]$	GeV
5	$p_{\tau,\text{sig}}^{\text{CMS}}$	$\in [2.489, 5.101]$	GeV/c
6	$\cos\theta_{\text{thrust}}$	$\in [-0.47, 0.881001]$	-
7	R_3	$\in [0, 0.9]$	-
8	R_4	$\in [0.308999, 0.98]$	-
9	m_{miss}^2	$\in [-47.8, 53.0]$	GeV$^2/c^4$
10	$n_{\pi^0,\text{tag}}$	≤ 1	-
11	$n_{\gamma,\text{tag}}$	$= 0$	-
12	$n_{\pi^0,\text{sig}}$	≤ 1	-
13	$n_{\gamma,\text{sig}}$	≤ 2	-
14	$m_{\tau,\text{sig}}$	$\in [0, 1.5901]$	GeV/c^2
15	$\chi^2_{\tau,\text{sig}}$	$\in [0.00999, 0.99999]$	-
16	thrust	$\in [0.90901, 1]$	-

Table 3.4: *Selection cuts.* Features and ranges defining the cuts applied to select signal. The numbers on the left designate the order in which the cuts have been applied.

3.4.1 Correction Factors Computation

The performance of the PID variables is evaluated through the computation of the pion efficiency, but also the rate at which pions are faking kaons. The latter is conventionally denoted as "$\pi \to K$" misidentification rate. These two quantities are computed using ratios, and are computed differently for Monte Carlo and data samples, such as:

$$\varepsilon_{MC} = \frac{N_{MC}^{\text{probe}} - N_{MC}^{\text{probe,bkg}}}{N_{MC}^{\text{tag}} - N_{MC}^{\text{tag,bkg}}} = \frac{N_{MC}^{\text{probe,sig}}}{N_{MC}^{\text{tag,sig}}} \quad \text{and} \quad \varepsilon_{data} = \frac{N_{data}^{\text{probe}} - N_{MC}^{\text{probe,bkg}}}{N_{data}^{\text{tag}} - N_{MC}^{\text{tag,bkg}}}, \tag{3.8}$$

where N^{probe} is the number of truth-matched events remaining after a hID cut applied, N^{tag} the total number of truth-matched events and $N^{\text{x,y}}$, x = probe, tag,

Process	Comp. [%]	π [%]	K [%]	p [%]	μ [%]	e [%]	none [%]	other [%]
$\tau^+\tau^-$	98.967	98.908	0.129	0.004	0.284	0.045	0.630	0.000
$q\bar{q}$	0.904	81.822	13.957	2.940	0.295	0.144	0.767	0.074
$\ell^+\ell^-X^+X^-$	0.124	98.193	0.261	0.022	0.283	0.479	0.762	0.000
BB	<0.001	37.705	18.033	0.000	0.000	29.508	3.279	0.000
e^+e^-	0.004	0.000	0.000	7.692	0.000	0.000	0.000	0.000
$e^+e^-\ell^+\ell^-$	<0.001	0.000	0.000	0.000	0.000	0.000	100.000	0.000
Total	100.000	98.749	0.254	0.031	0.284	0.047	0.631	0.001

Table 3.5: *Composition of the remaining samples.* The particle listed in the table header correspond to the probe track Monte Carlo truth-matching results. The category "*none*" corresponds to no identification of the particle, whereas "*other*" corresponds to other types of particles not listed in the table. The $\ell^+\ell^-X^+X^-$ sample groups decays with $e^+e^-K^+K^-$, $e^+e^-\pi^+\pi^-$, $e^+e^-p^+p^-$, $e^+e^-\tau^+\tau^-$, $\mu^+\mu^-\tau^+\tau^-$, and $\tau^+\tau^-\tau^+\tau^-$ final states. And $e^+e^-\ell^+\ell^-$ groups decays with $e^+e^-e^+e^-$ and $e^+e^-\mu^+\mu^-$ final states.

y = sig, bkg the number of signal and background events truth-matched in Monte Carlo samples with and without the application of hID cut respectively.

The efficiency and misidentification rate are first computed on the integrated momentum and polar angle spectrum, *i.e.*, without any further requirement applied on these two quantities. In total, 19 values of hID cut are used, starting at 0.05 and repeated every 0.05 until 0.95. To further refine the correction factor, bins of momentum and polar angle are used, and correspond to the ones presented in Table 3.3 in Section 3.3.3. The associated correction factor are defined as:

$$R_\varepsilon = \frac{\varepsilon_{data}}{\varepsilon_{MC}}. \tag{3.9}$$

The uncertainty associated to the correction factors are divided into statistical and systematic components. The statistical uncertainties are determined by assuming a Poisson distribution associated to the number of events, and then propagated in the ratio assuming that they are independent. As for the systematic uncertainties, several sources have been identified, and are detailed in Section 3.4.2.

3.4.2 Systematic Uncertainties

The systematic uncertainties were deduced from the difference of efficiency ratio with and without the corresponding variation applied. They are computed differ-

ently for Monte Carlo and data samples, as follows:

$$\sigma_{\text{syst}} = \frac{\varepsilon_{data}}{\varepsilon_{MC}} - \frac{\varepsilon_{data}}{\varepsilon'_{MC}} \quad \text{and} \quad \sigma_{\text{syst}} = \frac{1}{2}\left(\frac{\varepsilon'_{data}}{\varepsilon_{MC}} - \frac{\varepsilon_{data}}{\varepsilon_{MC}}\right), \tag{3.10}$$

where ε'_{MC} and ε'_{data} are the efficiencies computed with the corresponding variation. This assumption is very conservative. It is motivated by the fact that relying on the central values rather that the uncertainties would benefit for the implementation in the *Systematic Corrections Framework*, as they are assumed to be stable over time, unlike the uncertainties. A list of possible sources has been established, and each of them are detailed in the next sections.

3.4.2.1 Signal Side Modeling

The current generator that is employed for the tau lepton pair sample is TAUOLA, as detailed in Table 2.3. The systematic uncertainty associated to the model generator is evaluated through a variation on the Monte Carlo reconstructed samples, which results from the use of two other models. These two models are the Resonance Chiral Lagrangian initialisation [65] and the CLEO form factor [79]. For each configuration of charge, a sample corresponding to an integrated luminosity of 100 fb^{-1} has been produced. It is possible to compute the variation ε'_{MC} for each sample by exchanging the nominal tau lepton pair sample with each differently generated ones. The biggest difference between the two configurations is translated as a systematic uncertainty.

3.4.2.2 Trigger

In order to determine the variation associated to the trigger selection, the trigger efficiency has to be evaluated in both data and Monte Carlo samples. It is computed as follows

$$\varepsilon_t = \frac{N_{t_m \wedge t_r}}{N_{t_r}}, \tag{3.11}$$

where $N_{t_m \wedge t_r}$ is the number of events after the use of a main and a reference trigger, whereas N_{t_r} is the number of events after the use of a reference trigger only.

A reference and main trigger have to be selected. For this study, a CDC-based selection is opted for the main trigger, and an ECL-based combination is used as a reference trigger. Their definition is discussed in Section 2.4. These two kind of triggers are assumed to be independent, as they are coming from two different

detectors. The CDC triggers that are used correspond to three tracks triggers *fff*, *ffo*, *ffy* and *fyo*. The ECL reference trigger is *hie*. Their definition is given in Table 3.6.

Name	Base	Definition
hie	ECL	total ECL energy > 1 GeV and no ECL Bhabha veto signal
ffy	CDC	≥ 2 full tracks and one track reconstructed using a Neural Network
fyo	CDC	1 full track and 1 Neural Network reconstructed track, with $\Delta\phi > 90°$
fff	CDC	≥ 3 full tracks
ffo	CDC	≥ 2 full tracks, with $\Delta\phi > 90°$

Table 3.6: *ECL and CDC-based trigger bits.* A short definition is provided for each trigger bit.

The trigger selection is corrected by the ratio of efficiency of data over Monte Carlo. It is applied on Monte Carlo to correct $N_{MC}^{\text{probe,bkg}}$ and $N_{MC}^{\text{tag,bkg}}$ in Equation (3.8) to compute the data sample variation ε'_{data}.

3.4.2.3 Misidentification Rates

The dominant identified particle sources after pions are muons and kaons, as demonstrated in Table 3.5. As $\mu \to \pi$ misidentification rates are not available from official production, only the $K \to \pi$ misidentification rates are taken into account. The misidentification rates were produced using the same binning used for the correction factors. The variation is performed on Monte Carlo samples by correcting the nominal sample for the $K \to \pi$ misidentification rates, and results in ε'_{MC}.

3.4.2.4 Background

This systematic source involves mostly $q\bar{q}$ events, as it is the dominant non-tau lepton pair background sample as reported in Table 3.5. A correction factor accounting for these events is computed using a side-band region. This region is defined by changing the cut on the signal mass, such that $m_{\tau,\text{sig}} > 2$ GeV/c^2, and by selecting the events that have a thrust value below 0.99 as to avoid Bhabha events. The trigger selection employed in Section 3.4.2.2 is as well applied.

In order to understand more the differences between data and Monte Carlo, the side-bands are divided into Chunks (Chunk1 to Chunk5) and Prompt data sets.

Their corresponding luminosity can be found in Table 3.1. The correction factors are computed as follows:

$$\varepsilon_{q\bar{q}}^{\text{corr}} = \frac{N_{data} - N_{MC}^{\text{other}}}{N_{MC}^{q\bar{q}}}, \tag{3.12}$$

where N_{data} is the number of data events, N_{MC}^{other} is the number of events that are coming from all sources except $q\bar{q}$, and $N_{MC}^{q\bar{q}}$ the number of $q\bar{q}$ events. The correction factors are displayed in Figure 3.3 for the different data sets.

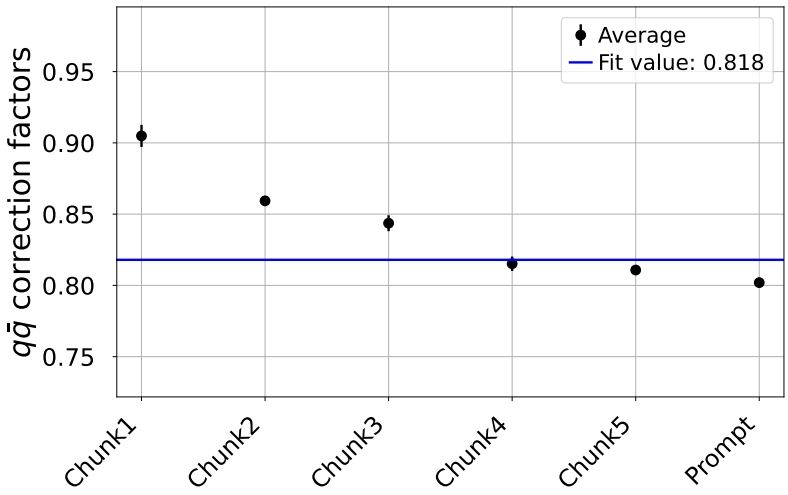


Figure 3.3: *$q\bar{q}$ correction factors.* The factors are computed for Chunks and Prompt data sets.

The result of the fit, which is $\varepsilon_{q\bar{q}}^{\text{corr,fit}}$ and is equal to 0.818, is applied to correct the $q\bar{q}$ background events present in $N_{MC}^{\text{probe,bkg}}$ and $N_{MC}^{\text{tag,bkg}}$ from Equation (3.8), to compute the data sample variation ε'_{data}.

3.5 Selected Results

A few selected results highlight the performance of these correction factors. The efficiencies as well as a comparison with already available channels are provided in what follows. More results are available in Appendix A.

3.5.1 Efficiencies

The integrated efficiencies computed for the Global, Reweighted and MVA PID variables, as well as for both Monte Carlo and data samples are shown in Figure 3.4.

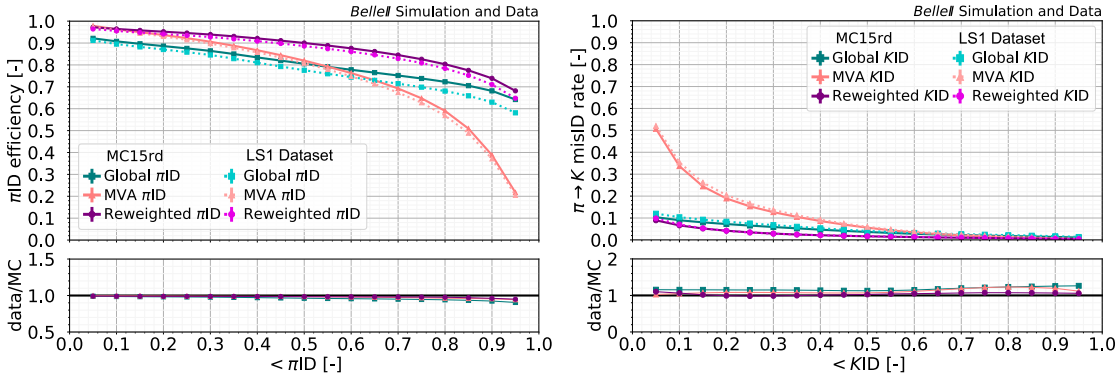


Figure 3.4: *πID efficiency and $\pi \to K$ misID rate.* πID efficiency (left) and $\pi \to K$ misID rate (right) computed for the Global, MVA and Reweighted hID variables once the cut-based selection is put in place. The solid lines correspond to Monte Carlo simulations, whereas dashed lighter lines represent data. The bottom part of the plot shows the ratio of data over Monte Carlo. The statistical uncertainties are represented by vertical lines, which are in this case too small to be visible.

Several observations can be deduced from this result. First, for the πID efficiency, the best performance is attributed to the Reweighted hID variable, and not only in terms of efficiency value, but also because it yields reduced discrepancies between Monte Carlo and data. In comparison, the Global hID variable has a lower efficiency and larger discrepancies. The MVA hID variable performance manifests an interesting behaviour, as its efficiency is rather low, but the difference observed between data and Monte Carlo is almost negligible, as opposed to the other variables. The same conclusion applies in the case of the $\pi \to K$ misID rate.

Tables providing the efficiency computed for Monte Carlo, along their propagated statistical and systematic uncertainties obtained for each hID variable with a working point of 0.5 are displayed in Figure 3.5.

The region determined by the first bin of the cosine of the polar angle highlights a different behaviour than that shown in other bins, in all hID variable cases. Indeed, the TOP detector, as depicted in Figure 3.2 and which is responsible for the detection of hadrons such as pions and kaons, is not covering this region, which might explain the poorer performance that is obtained. In the case of the Reweighed hID variable, the efficiency seems to suffer less from this effect, probably because it is compensated by the weights applied on its likelihoods. The second bin of cosine polar angle seems to also suffer from this effect, although to a reduced extent. As to avoid redundancy, the results for the data is not shown. However, the corresponding efficiency corrections are displayed in Figure 3.6.

The same problematic bins are emerging from the computation of the ratio R_ε, and

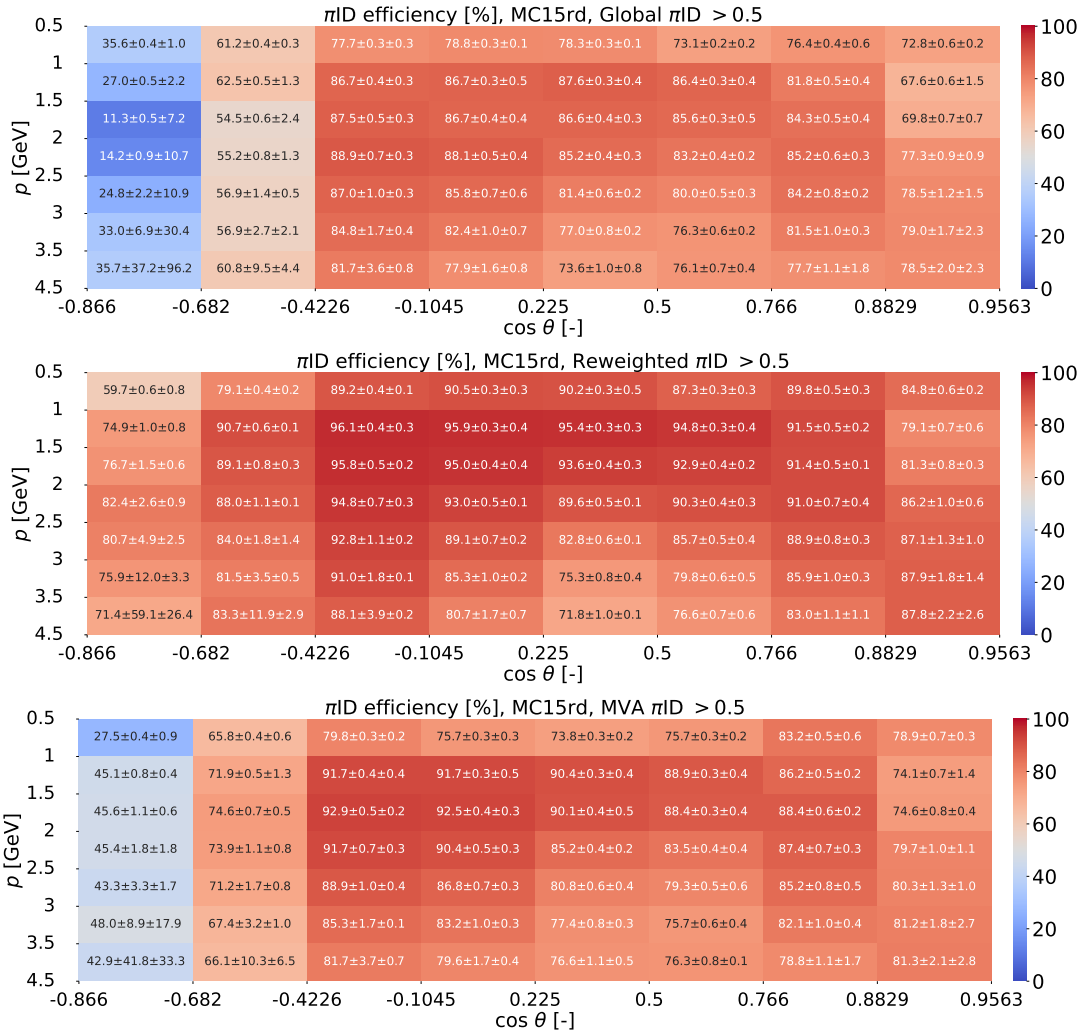


Figure 3.5: *Monte Carlo efficiencies.* The PID variables used to get the correction factors binned in momentum and polar angle correspond to the Global (top), the Reweighted (middle) and the MVA (bottom) variables.

are this time more striking in the case of the Global hID variable.

3.5.2 Comparison with Other Studies

To compare this study with the already available ones, the Global πID efficiency is used, as it provides a fair basis. Its performance is compared to the one resulting from the use of $D^{*+} \to D^0(K^-\pi^+)\pi^+$ decays. The comparison is done by computing the ratio of the correction factors, with their uncertainties propagated assuming that they are independent. This assumption might be revisited, as this study is correlated to the misidentification rate corrections used in the estimation of the systematic uncertainties associated to the misidentification.

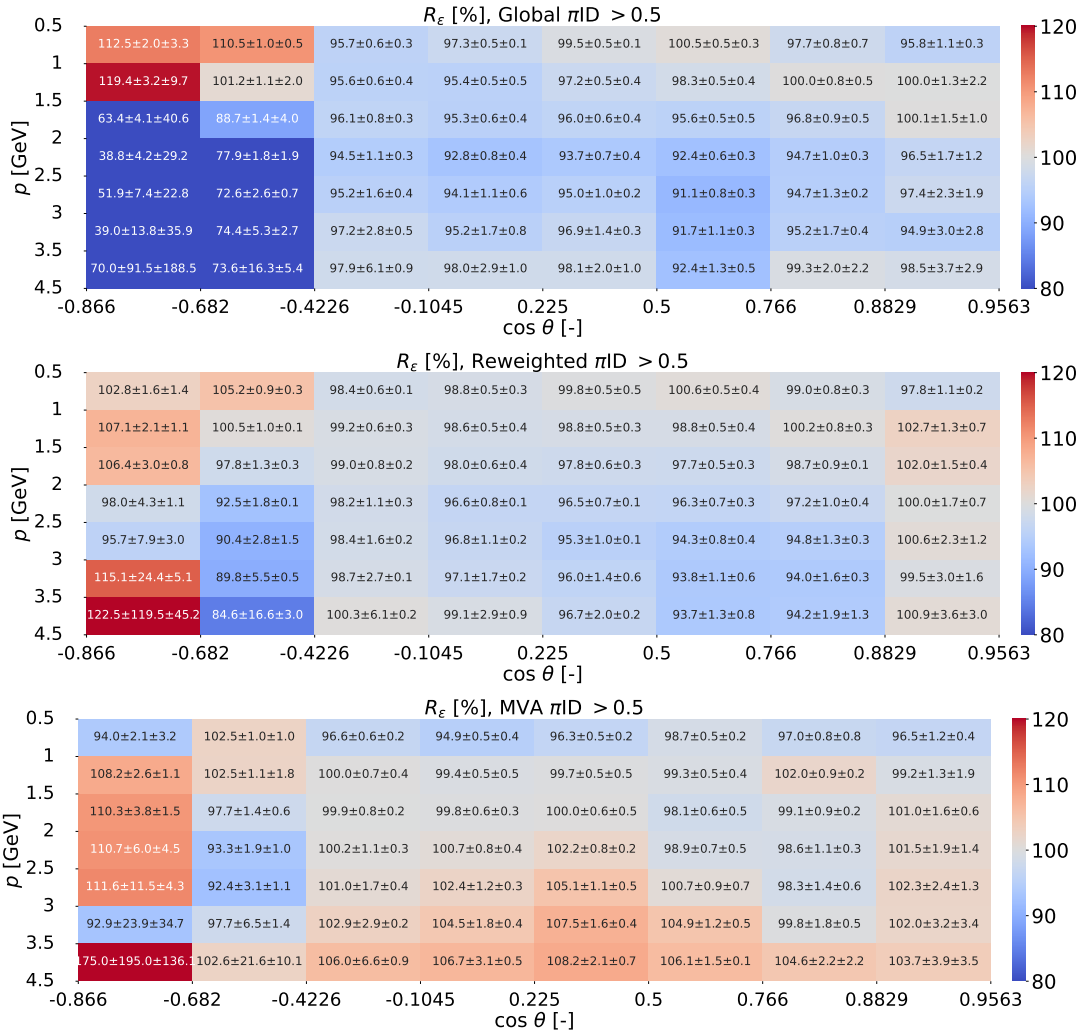


Figure 3.6: *Efficiency correction factors.* The PID variables used to get the correction factors binned in momentum and polar angle correspond to the Global (top), the Reweighted (middle) and the MVA (bottom) variables. The correction factors are multiplied by 100 in order to ease their readability.

The ratio alone in fact does not give a direct sense of measure of the agreement between the two samples. It is preferable to add pulls, which are computed as follows:

$$\text{pulls} = \frac{R_{\varepsilon_\tau} - R_{\varepsilon_{D^*}}}{\sigma_{\text{tot}}} \tag{3.13}$$

with

$$\sigma_{\text{tot}} = \sqrt{\sigma_\tau^2 + \sigma_{D^*}^2} \quad \text{and} \quad \sigma_X = \sqrt{\sigma_{X,\text{stat.}}^2 + \sigma_{X,\text{syst.}}^2}$$

The resulting comparison is shown in Figure 3.7. It demonstrate in particular that both samples are in good agreement.

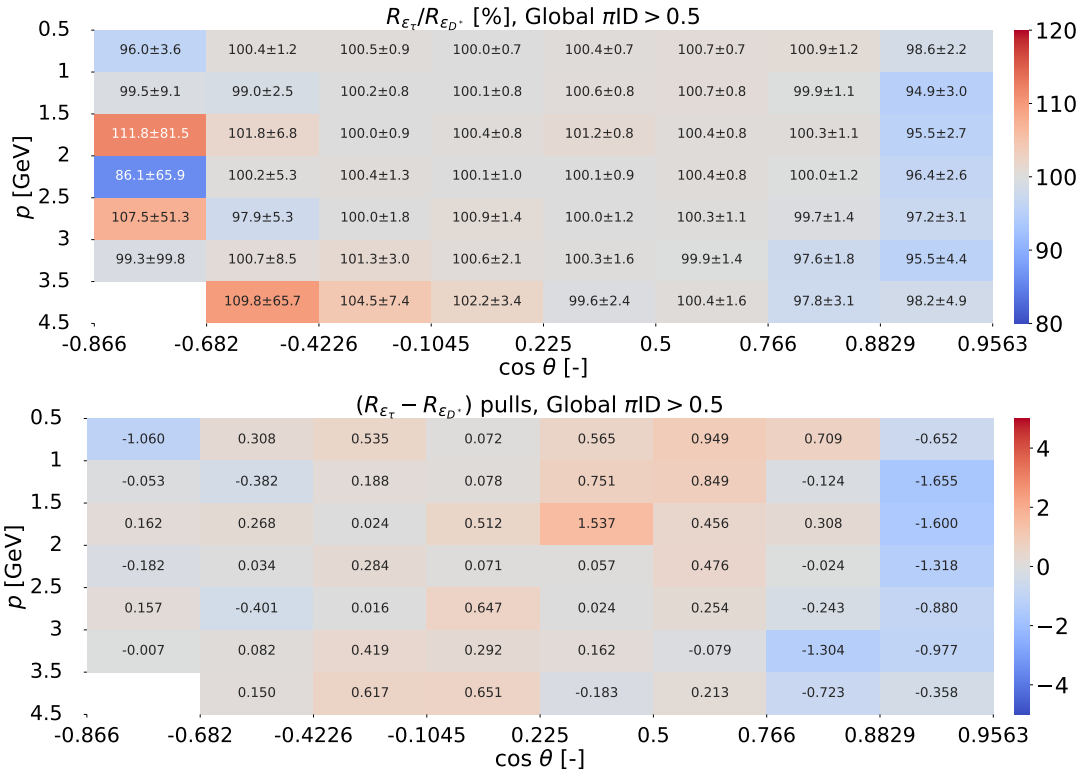


Figure 3.7: *Comparison with the D^* sample.* The ratio of correction factors (top), as well as the associated pulls (bottom) are displayed.

Another comparison of performance is done with results coming from the use of $K_s \to \pi^+\pi^-$ decays. As in the previous case, the ratio and pulls are computed, and are displayed in Figure 3.8. Here, a disagreement is observed between the two samples, and is located in low and medium momentum bins, and in medium cosine polar angle bins.

As to further understand the relation between the three samples, and in order to get a comprehensive result, a comparison between the D^* and K_s samples is provided in Figure 3.9. Here, discrepancies are also observed, although to a smaller extent.

3.6 Discussion

In this performance study, a 3×1 prong tau decay topology was used to compute correction factors in bins of p and $\cos\theta$ for three different PID variables.

The events were selected by applying a cut based selection. It reached a total purity of 98.749%.

The set of PID variables was composed of the Global hID, defined as the ratio of

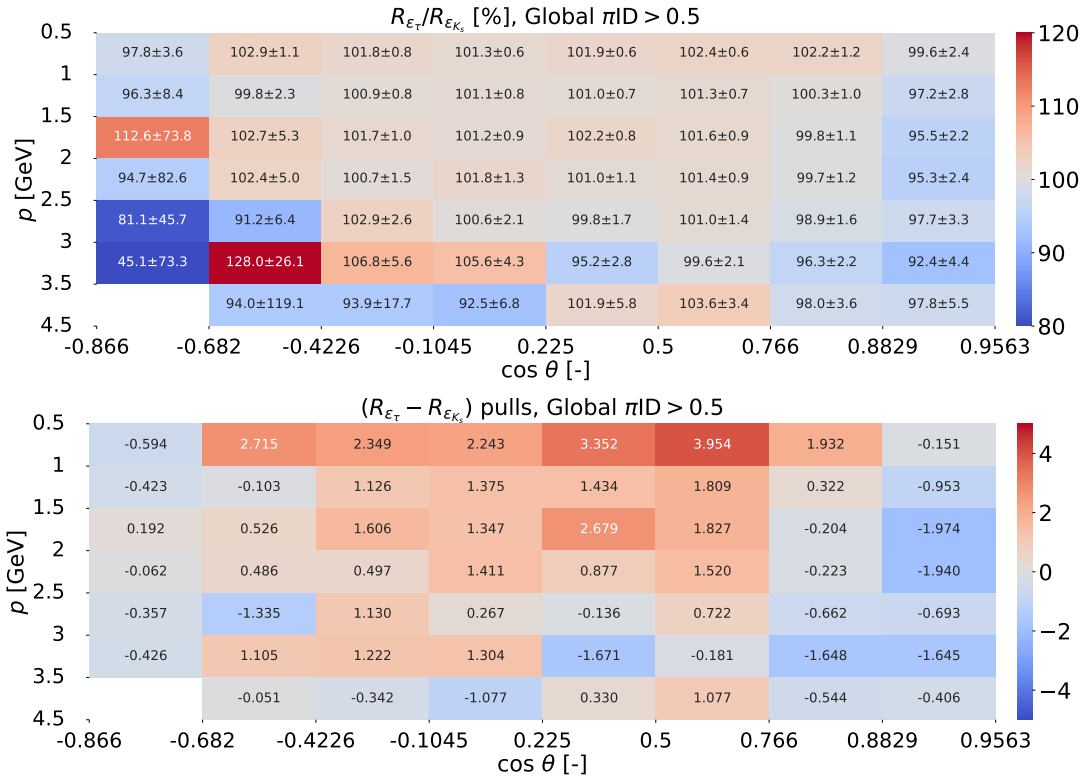


Figure 3.8: *Comparison with the K_s sample.* The ratio of correction factors (top), as well as the associated pulls (bottom) are displayed.

likelihoods of all detectors and particles, the Reweighted hID, being a weighted version of the Global hID, and the MVA hID, whose likelihoods $\mathscr{L}'$ are relying on an improved $e-\mu,\pi$ separation.

The performance of each hID variable was assessed by computing their corresponding πID efficiency and the $\pi \to K$ misID rate on integrated samples as well as in bins of p and $\cos\theta$. Several observations can be made on the results. First, the MVA hID performed the worst as its curved dropped significantly while the cut value increased. Second, the Reweighted hID outperforms the Global hID not only in terms of performance, but also in terms of discrepancy between data and Monte Carlo. Correction tables were produced for each of these hID variable, and for 9 values of hID cut in bins of p and $\cos\theta$. Their associated statistical uncertainties were assessed and are indicated on these tables as well as systematic uncertainties.

These are divided into four categories: systematic uncertainties related to the modeling of the $\tau \to \pi\pi\pi$ decay, to the fake rates of kaons, to the trigger selection and the $q\bar{q}$ background level. These assessments were done in a very conservative way.

A comparison with the D^* sample as been made with matching versions of samples. A first comparison in bins of momentum p or $\cos\theta$ has revealed a reasonable

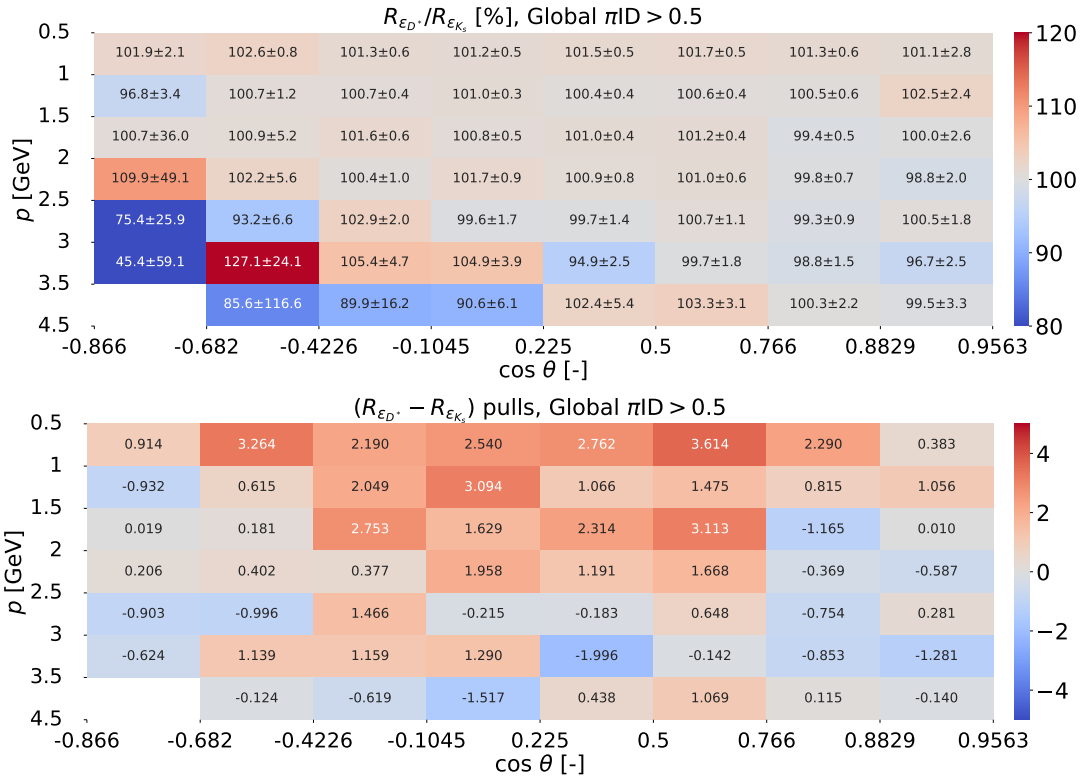


Figure 3.9: *Comparison between the D^* and K_s samples.* The ratio of correction factors (top), as well as the associated pulls (bottom) are displayed.

agreement between the two samples. On the contrary, the comparison with the K_s samples demonstrated a non negligible discrepancy between the two samples, which was especially visible for higher cut values, as also highlighted by the pulls. An additional comparison between the D^* and K_s sample provided a comprehensive understanding of the agreement between the different studies, by highlighting further discrepancies between these two channels.

Although the results of this performance study are not yet available through the *Systematic Corrections Framework*, it is nevertheless encouraging to see that they are in agreement with one of the two available corrections. This indicates that low multiplicity channels using tau lepton pairs can reliably use the available corrections.

There are several ways of improving this measure. Firstly, discrepancies were observed between data and Monte Carlo due to an issue related to the skimming process: several runs were duplicated in data but not in Monte Carlo samples. Additionally, the systematic framework tables were produced some time ago, and changes have occurred since then, leading to increased systematic uncertainties that could impact the computed pulls when comparing different channels.

In terms of assessing systematic uncertainties, the side-band used for $q\bar{q}$ factors may not exactly match the region selected for correction factors, potentially introducing uncertainty into the correction factor value. Furthermore, observed patterns in the computation of these factors during different data-taking periods are suspected to be related to changes in triggers.

The fake rates used as another source of systematic uncertainty are corrected for kaons but not for muons, which are a major component, and the none category still requires thorough understanding as it is also a significant contributor.

Regarding systematic modeling on the signal side, evaluation is limited to certain models due to inaccuracies in the CLEO form factors. However, using a conservative approach by leaving out certain models can be advantageous despite the limitations.

Additionally, the MVA PID variable was optimised for lepton identification rather than hadrons, resulting in less effective performance as anticipated.

4. Physics Analysis

The CKM matrix is subject to scrutiny aiming to detect deviations from unitarity that could signify the presence of New Physics (NP), while acknowledging the potential theoretical origins of the observed tension.

In this chapter, we discuss the measurement of the magnitude of the CKM matrix element V_{us}, which is the essence of this thesis. This measurement is performed using a ratio of branching fractions of hadronic tau decays, called $R_{K/\pi}$. Belle II lend itself well to the use of tau decays, as it both produces tau lepton pair events abundantly, and that systematic uncertainties are kept to a minimum through Belle II continuous efforts to reduce them. Charge conjugation is always implied, except where explicitly stated otherwise. Comparison with the BaBar collaboration [33] is provided when relevant, as a final measurement on data has not been obtained.

4.1 Methodology Overview

A tau decay topology similar to that presented in section Section 3.1 is used. This 3×1-prong topology, along with its corresponding Feynman diagram, is illustrated in Figure 4.1. Unlike in the previous chapter, the signal side is defined as the 1-prong side, and corresponds to the decays $\tau^- \to \pi^- \nu_\tau$ or $\tau^- \to K^- \nu_\tau$. The tag side, on the other hand, is the 3-prong side, and corresponds to $\tau^- \to X^- X^+ X^- \nu_\tau$, with X being any charged particle system. This topology allows for the selection of tau lepton pair events with reduced background level from other processes.

The selection of signal events involves three steps, strategically optimised for the given conditions. The initial step involves implementing cuts designed to suppress background events, but also regions where discrepancies exist between data and Monte Carlo simulations. These discrepancies arise due to incomplete simulation of all potential decay processes, rendering the simulations inadequate to accurately describe certain regions. Only 10% of data is used in this step to mitigate unblinding issues. The second step is directly related to the third and final step. In fact,

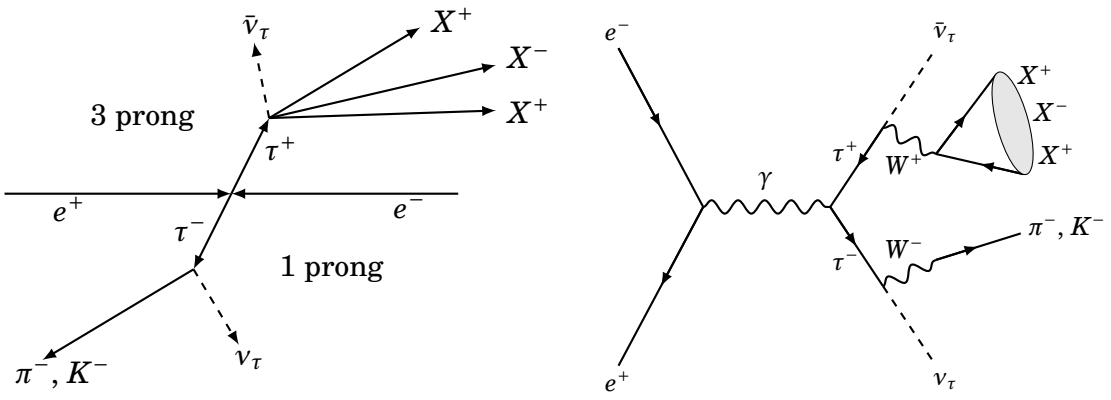


Figure 4.1: *Diagrams of targeted processes.* Illustration of the 3×1 prong topology and Feynman diagram for the $e^+e^- \to \tau^+\tau^-$ with on the signal side either $\tau^- \to \pi^-\nu_\tau$ or $\tau^- \to K^-\nu_\tau$, and on the tag side $\tau^+ \to X^+X^-X^+\bar{\nu}_\tau$, which is the process targeted by this analysis.

the last selection step uses multivariate methods such as Boosted Decision Trees (BDT). Since these methods depend on variable shapes, it is needed to ensure they accurately describe data. Hence, the second step focuses on identifying and computing correction factors related to the events or specific selection steps that are performed.

Differences might arise with a different selection that is tailored to each decay. Therefore, for the whole selection of the events, the samples are combined as to mitigate these differences. Several checks have been performed as to ensure that the performance of the selection is not impacted by this decision, especially in the case of the multivariate selection.

The measurement of $|V_{us}|$ is performed on these selected events. As demonstrated in Equation (1.19), this computation relies on the ratio $R_{K/\pi}$, as well as on other factors, such as the mass of the two hadrons (m_h) and the tau lepton (m_τ), hadronic form factors (f_h), radiative corrections ($\delta_{R_{\tau,K/\tau,\pi}}$), and another CKM matrix element magnitude ($|V_{ud}|$). These additional factors are independent from the measurement conducted with the selected events. Therefore, our primary emphasis lies in measuring $R_{K/\pi}$, and subsequently, we propagate uncertainties onto $|V_{us}|$ using the latest values available.

The ratio $R_{K/\pi}$ is measured with a template fit method. This allows for a comprehensive assessment of the ratio using maximum likelihoods estimations, since it simultaneously evaluates the associated uncertainties. We identify several sources of systematic uncertainties that are added as well-defined modifications of the template. Once all systematic uncertainties are implemented, the evaluation of the

total uncertainty on the $R_{K/\pi}$ ratio can be performed on Monte Carlo simulations. Regrettably, certain uncertainties could not be fully assessed within the given time constraints and availability of necessary corrections. The evaluation methods for these missing uncertainties are still discussed, as it is intended to address them in the near future.

4.2 Input Samples

For this measurement, the data collected by Belle II from 2019 to 2022 was used, and corresponds to a total integrated luminosity of 364.093 fb^{-1}. The details concerning this data set, called also "*LS1*", can be consulted in Table 3.1.

The simulated samples are taken from the official Belle II 15^{th} Monte Carlo campaign of run dependent samples. They correspond to the same generic and low multiplicity samples detailed in Table 3.2. They are reiterated here for the reader's convenience:

- ⋆ *Tau lepton pair samples*: from $e^+e^- \to \tau^+\tau^-$ process
- ⋆ $q\bar{q}$ *samples*: from $e^+e^- \to q\bar{q}$ processes, with $q = c, d, s$ and u
- ⋆ *B meson samples*: from $e^+e^- \to B^0\bar{B}^0$ and $e^+e^- \to B^+B^-$ processes
- ⋆ *Radiative dilepton samples*: from $e^+e^- \to \ell^+\ell^-(\gamma)$ processes, with $\ell = e, \mu$
- ⋆ *Two-photon samples*: from $e^+e^- \to e^+e^-\ell^+\ell^-$ processes, with $\ell = e, \mu, \tau$
- ⋆ *Low-multiplicity samples*: from $e^+e^- \to e^+e^-h^+h^-$ and $e^+e^- \to \mu^+\mu^-\ell^+\ell^-$ processes, with $h = K, \pi, p$ and $\ell = \mu, \tau$
- ⋆ *Initial state radiation samples*: from $e^+e^- \to h^+h^-(\gamma)$, $e^+e^- \to K^0\bar{K}^0(\gamma)$, and $e^+e^- \to \pi^+\pi^-\pi^0(\gamma)$ processes, with $h = K, \pi$

4.3 Event Selection

To accurately measure $R_{K/\pi}$ and therefore $|V_{us}|$, the signal events have to be selected. The signal in this case is defined by two decays: $\tau^- \to \pi^-\nu_\tau$ and $\tau^- \to K^-\nu_\tau$. These decays, as indicated on Figure 2.16, have a branching fraction of 10.816% and 0.696% respectively. The selection process is done on the combined signal samples.

The cuts employed in the first part of the selection are applied during the reconstruction of the decay, but also after this first selection to adjust more precisely the cuts. These two steps corresponds in what follows to the "*preselection*" and "*selection cuts*".

Before continuing and finalizing the selection with a multivariate selection, correction factors are computed in order to reshape the Monte Carlo simulation. These correction factors depend on the selection of the events. The multivariate selection is needed to enhance the purity of the sample. As this last sepp accomplishes a selection of events in a very effective way, we refrain from implementing any overly complex or finely tuned selections in the first step.

4.3.1 Selection Assessment

The concepts of purity, efficiency and figure of merit have already been introduced in Section 3.3.1 (in Equations (3.1), (3.2), and (3.3)). The only element that differs in this measurement from the performance study is that the branching fraction of the 3-prong and 1-prong sides are given by $\mathscr{B}_{3p}$ = 15.2%, $\mathscr{B}_K$ = 0.696% and $\mathscr{B}_\pi$ = 10.816%. Therefore, this section is intended to introduce multivariate methods, related concepts and ways of assessing their performance. They are introduced in a similar way to that used in [80].

Multivariate analysis techniques, utilizing machine learning algorithms, are employed to identify patterns and use this knowledge to predict outcomes for comparable, but unknown datasets. To accomplish this objective, MVA methods use three primary inputs: sets of events used to train and test the MVA method to recognise patterns, referred to as "*test*" and "*training*" sets, key variables defining these events, known as "*discriminating features*", and a set of computational configurations, known as "*hyper parameters*". Given the nature of MVA methods, which involve event classification, the output, representing a prediction, is commonly referred to as a "*classifier*". Various methods are available for supervised learning classification. This measurement employs in particular Boosted Decision Trees (BDT).

Decision trees are effective tools for classification tasks, operating on a recursive splitting approach based on input features. This method involves making decisions at each step (denoted as "*nodes*") and ultimately reaching endpoints (called "*leaves*") that correspond to class labels or probabilities. The criteria for splitting nodes aim to maximise information gain or minimise entropy [81]. The visualisation of this process in Figure 4.2 shows the tree-like structure characteristic of this method.

The ensemble of trees that leads to a classification is interpreted as boosting.

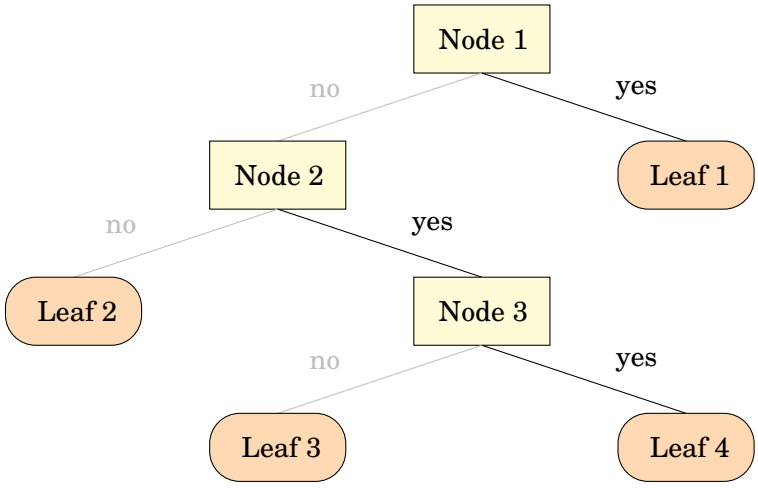


Figure 4.2: *Illustration of a decision tree.* Two branches arise from each decision, corresponding to whether or not the criterion is met. This process progressively draws branches constituting a so-called decision tree. This particular tree is growing leaf-wise, and not level-wise.

In supervised learning, the prediction $\hat{y}$ is computed using the input x. Generally, predictions are assumed to be proportional to inputs. The factors of proportionality θ are determined through model training, which aims to find the factors that best fit the inputs and predictions. The objective function $O(\theta)$ is defined to measure the accuracy with which the model fits the training data. This function is the sum of two terms, which are the training loss $L(\theta)$ and regularisation $\Omega(\theta)$, such that:

$$O(\theta) = L(\theta) + \Omega(\theta) \tag{4.1}$$

The regularisation term $\Omega(\theta)$ handles issues, related to the concept of "*overtraining*", that can arise when a model is too complex. It will be touched on later, in particular in Figure 4.4. The training loss term $L(\theta)$, on the other hand, measures the predictability of a model. The objective function is enhanced based on the choice of boosting technique selected for classification, which improves the process of creating decision trees [82, 83].

Ensuring the effectiveness of a BDT method involves a consideration that is sometimes overlooked: incorporating the necessary features while adhering to Occam's razor principle, which advocates for simplicity by including no more features than required. This implies an evaluation of each feature necessity, assessing both its separating power and possible redundancy, to ensure that only the most influential attributes are used as inputs.

These two conditions are evaluated respectively by performing a ranking of the features, and computing their correlations. The latter is assessed by a correlation coefficient $C \in [-1,1]$, which is attributed to two features at a time. It can be either positive or negative, depending on their relation of dependency. Mathematically, the Pearson's correlation between two features X and Y is computed using a data set reduced to these features, being A and B respectively. If two features have a weight w, their correlation is computed as follows:

$$C_{X,Y} = C_{Y,X} = \frac{\mathrm{cov}(A,B,w)}{\sqrt{\mathrm{cov}(A,A,w)\mathrm{cov}(B,B,w)}}, \tag{4.2}$$

where

$$\mathrm{cov}(A,B.w) = \frac{\sum_{i=1}^{N} w_i(a_i - \bar{a})(b_i - \bar{b})}{\sum_{i=1}^{N} w_i} \qquad \text{and} \qquad \bar{\alpha} = \frac{\sum_{i=1}^{N} w_i \alpha_i}{\sum_{i=1}^{N} w_i}, \tag{4.3}$$

where $\mathrm{cov}(A,B,w)$ is the weighted covariance, and a, b, the elements composing A, B respectively.

The performance of the BDT is evaluated through its ability to establish the different categories it is asked to distinguish, and its reproducibility on other samples. The receiver operating characteristic (ROC) curve visually represents the classification performance of a BDT. An example of ROC curve is represented in Figure 4.3.

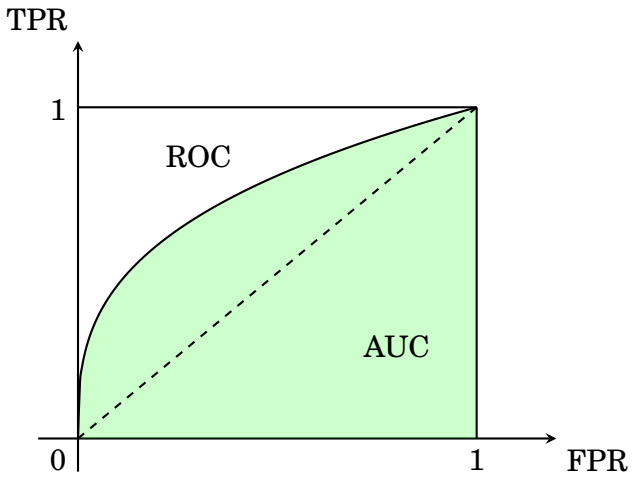


Figure 4.3: *Illustration of a ROC curve.* The area under the ROC curve (AUC) measures the BDT classification performance.

In the case of binary classification, the computation of the ROC curve is solely based on the classification interpretation [84]. Given binary class labels (positive, P; negative, N), four possible outcomes exist depending on whether the prediction matches

(true, T) or does not match (false, F) the actual value. These definitions yield the computation of a true positive rate (TPR) and a false positive rate (FPR) as:

$$\text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad \text{and} \quad \text{FPR} = \frac{\text{FP}}{\text{TN} + \text{FP}}. \tag{4.4}$$

The area under this curve (AUC), normalised to a value of 1, quantifies the effectiveness of the classification. An AUC close to 1 signifies excellent classification accuracy, while an AUC close to 0 indicates poor performance. An AUC value of 0.5 represents random classification.

The reproducibility of a BDT method is directly related to the concept of overtraining (also called "*overfitting*"). It occurs when a model's complexity is excessively high, leading it to capture every detail and nuance of the training dataset, resulting in a fit that may not generalise well to new or unseen data. An example of classification of events that could potentially lead to overtraining is provided in Figure 4.4.

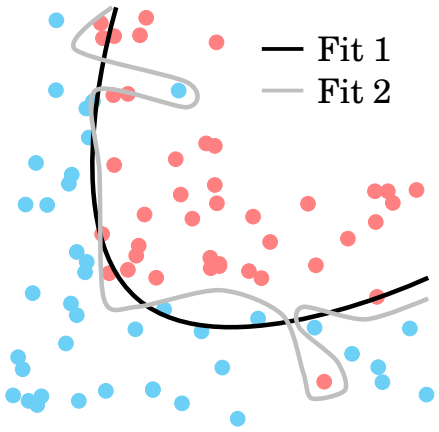


Figure 4.4: *Illustration of overtraining.* Fit 1 represents a regularized model, whereas Fit 2 represents an overtrained model.

While some kinds of BDTs are more prone to overtraining than others, there is a common belief that using simplified models can effectively prevent this issue during training. Applying the BDT on a test set enables verification of its reproducibility. If a significant difference is observed in the BDT's classification between the test and training sets, it suggests that overtraining is occurring. One widely used method to measure overtraining is the Kolmogorov-Smirnov test [85]. This test yields a uniformly distributed value between 0 and 1, assessing the similarity of outcomes between test and training sets. The BDT method is generally considered not to exhibit overtraining when the test result is equal to 0, but this value also depends on the implementation of this test.

4.3.2 Selection Variables

The variables selected for this measurement must efficiently separate the signal from background, while accurately representing real data.

The thrust, introduced in Section 2.6, Equation (2.15), is used to distinguish signal from other processes, especially $q\bar{q}$ and Bhabha events. Other variables defined using the thrust are the cosine of the polar angle component of the thrust axis ($\cos\theta_{\text{thrust}}$), and the momentum flow in cones defined around the thrust axis, known as "*CLEO cones*" [86] (CC_i). There is a total of nine cones separated each by a polar angle intervals of 10 degrees, covering the regions defined by angles from $i \times 10$ degrees to $-i \times 10$ degrees from the thrust axis. Forward and backward intervals are combined. Three of the cones are illustrated in Figure 4.5. The momentum flow in each cone is calculated as the scalar sum of all tracks and clusters in that cone.

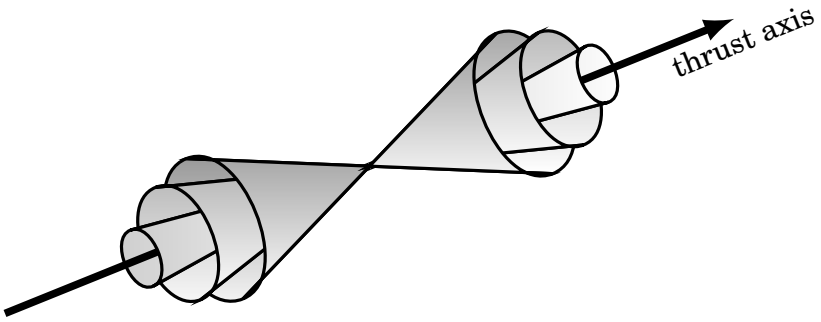


Figure 4.5: *CLEO cones.* Illustration of the first three of the nine cones.

Other variables related to event shape and kinematics used to remove two-photon and radiative dilepton events are the visible energy of the event in the CMS frame ($E_{\text{vis}}^{\text{CMS}}$), the missing mass squared of the event (m_{miss}^2), the missing momentum in the lab and CMS frames (p_{miss} and $p_{\text{miss}}^{\text{CMS}}$), the polar angle component of the missing momentum of the event in the CMS frame ($\theta_{p_{\text{miss}}^{\text{CMS}}}$) and the Fox-Wolfram normalised moments (R_ℓ).

Tag side variables consist of the mass ($m_{\tau,\text{tag}}$), the energy in the CMS frame ($E_{\tau,\text{tag}}^{\text{CMS}}$) and the sorted transverse momenta ($p_{T,\tau,\text{tag}}^{i\text{th}} \equiv p_{T,\text{tag}}^{i\text{th}}$). These momenta are defined as follows:

$$\begin{cases} p_{T,\text{tag}}^{1\text{st}} & \equiv & \max(p_{T,\text{track}\,1},\ p_{T,\text{track}\,2},\ p_{T,\text{track}\,3}) \\ p_{T,\text{tag}}^{2\text{nd}} & \equiv & \text{not}\big(\max \text{ or } \min\big) \\ p_{T,\text{tag}}^{3\text{rd}} & \equiv & \min(p_{T,\text{track}\,1},\ p_{T,\text{track}\,2},\ p_{T,\text{track}\,3}) \end{cases} \tag{4.5}$$

Signal side variables are needed to constrain the phase space region of the measurement. In particular, the momentum ($p_{\tau,\mathrm{sig}}$) and the cosine component of the polar angle ($\cos\theta_{\tau,\mathrm{sig}}$) are used.

Finally, the number of neutrals are valuable variables, which are notably needed to build the events during the reconstruction. More precisely, the number of neutral pions in the signal and tag sides ($n_{\pi^0,\mathrm{sig}}$ and $n_{\pi^0,\mathrm{tag}}$) and the number of photons in the signal and tag sides ($n_{\gamma,\mathrm{sig}}$ and $n_{\gamma,\mathrm{tag}}$) are used.

Regarding the particle identification variables, we use two types: the Global PID, corresponding to Equation (2.9), and the Reweighted PID, detailed in Equation (2.12). To facilitate readability, previously introduced equations are reiterated here for the reader's convenience. Assuming a set of sub-detectors $D \equiv$ { SVD, CDC, TOP, ARICH, ECL, KLM} and a set of long-lived particles $X \equiv \{e,\mu,\pi,K,p,d\}$, we obtain:

$$\begin{cases} \text{Glob.} \equiv P_i = \dfrac{\exp\left(\log \mathscr{L}_i\right)}{\sum\limits_{x\in X}\exp\left(\log \mathscr{L}_x\right)} = \dfrac{\exp(\sum_{d\in D}\log \mathscr{L}_i^d)}{\sum_{x\in X}\exp(\sum_{d\in D}\log \mathscr{L}_x^d)} \\ \text{Rew.} \equiv P_i' = \dfrac{\exp\left(\log \mathscr{L}_i'\right)}{\sum\limits_{x\in X}\exp\left(\log \mathscr{L}_x'\right)} = \dfrac{\exp(\sum_{d\in D} w_{i,d}\log \mathscr{L}_i^d)}{\sum_{x\in X}\exp(\sum_{d\in D} w_{x,d}\log \mathscr{L}_x^d)} \end{cases} \tag{4.6}$$

These two variables are not used simultaneously, but are explored separately, so that the selection that minimises errors can be chosen. The same kind of PID variable is used for both pions and kaons.

As for the trigger selection, CDC- and ECL-based trigger bits were selected. Using the definitions introduced in Section 2.4, they correspond to *ffy*, *fyo*, *fff*, *ffo*, *lml0*, *lml1*, *lml2*, *lml4*, *lml6*, *lml7*, *lml8*, *lml9*, *lml10*, *lml12* and *hie*, which are three charged tracks, low-multiplicity and cluster energy triggers. Their definition is given in Table 4.1.

4.3.3 Preselection

Events are reconstructed using charged tracks originating from the interaction region, which is delimited by restricting the radial and axial impact parameter as $\mathrm{d}r < 1.0$ cm and $|\mathrm{d}z| < 3.0$ cm. Events are required to contain exactly four charged tracks, and their sum is required to be zero.

Neutral particles are divided into neutral pion candidates and photons not coming from a neutral pion. Neutral particles are used to compute event shape and kinematics variables for the aforementioned charged tracks.

Name	Base	Definition
hie	ECL	total ECL energy > 1 GeV and no ECL Bhabha veto signal
lml0	ECL	$N_{\text{cluster}} \geq 3$, one with $E_{\text{CM}} > 300$ MeV, $1 < \theta_{\text{id}} < 17 \equiv 12.4^\circ < \theta_{\text{lab}} < 154.7^\circ$, not an ECL Bhabha
lml1	ECL	exactly 1 cluster with $E_{\text{CM}} > 2$ GeV and $4 < \theta_{\text{id}} < 14$ $\equiv 32.2^\circ < \theta_{\text{lab}} < 124.6^\circ$
lml2	ECL	$N_{\text{cluster}} \geq 1$ with $E_{\text{CM}} > 2$ GeV, $\theta_{\text{id}} = 2, 3, 15$ or 16 $\equiv 18.5^\circ < \theta_{\text{lab}} < 32.2^\circ$ or $124.6^\circ < \theta_{\text{lab}} < 139.3^\circ$ and not an ECL Bhabha
lml4	ECL	$N_{\text{cluster}} \geq 1$ with $E_{\text{CM}} > 2$ GeV, $\theta_{\text{id}} = 1$ or 17 $\equiv 12.4^\circ < \theta_{\text{lab}} < 154.7^\circ$ and not an ECL Bhabha
lml6	ECL	exactly one cluster with $E_{\text{CM}} \geq 1$ GeV, $4 \leq \theta_{\text{id}} \leq 15$ $\equiv 32.2^\circ < \theta_{\text{lab}} < 128.7^\circ$ and no other cluster with $E_{\text{lab}} > 300$ MeV anywhere
lml7	ECL	exactly 1 cluster with $E_{\text{CM}} > 1$ GeV, $\theta_{\text{id}} = 2, 3$ or 16 $\equiv 18.5^\circ < \theta_{\text{lab}} < 31.9^\circ$ or $128.7^\circ < \theta_{\text{lab}} < 139.3^\circ$ and no other cluster with $E_{\text{lab}} > 300$ MeV anywhere
lml8	ECL	two clusters with $170^\circ < \Delta\phi_{\text{CM}} < 190^\circ$ and $E_{\text{CM}} > 250$ MeV and no other cluster with $E_{\text{CM}} \geq 2$ GeV anywhere
lml9	ECL	two clusters with $170^\circ < \Delta\phi_{\text{CM}} < 190^\circ$ and one cluster with $E_{\text{CM}} > 250$ MeV, the other with $E_{\text{CM}} < 250$ MeV and no other cluster with $E_{\text{CM}} \geq 2$ GeV anywhere
lml10	ECL	two clusters with $160^\circ < \Delta\phi_{\text{CM}} < 200^\circ$ and $160^\circ < \sum\theta_{\text{lab}} < 200^\circ$ and no other cluster with $E_{\text{CM}} \geq 2$ GeV anywhere
lml12	ECL	$N_{\text{cluster}} \geq 3$, one with $E_{\text{CM}} \geq 500$ MeV, $2 \leq \theta_{\text{id}} \leq 16 \equiv 18.5^\circ < \theta_{\text{lab}} < 139.3^\circ$, not an ECL Bhabha
ffy	CDC	≥ 2 full tracks and one track reconstructed using a Neural Network
fyo	CDC	1 full track and 1 Neural Network reconstructed track, with $\Delta\phi > 90^\circ$
fff	CDC	≥ 3 full tracks
ffo	CDC	≥ 2 full tracks, with $\Delta\phi > 90^\circ$

Table 4.1: *ECL and CDC-based trigger bits.* A short definition is provided for each trigger bit.

Three different corrections are applied on the events. The correction applied on data involves a tracking momentum scale factor and a photon energy bias, and the correction applied on Monte Carlo is composed of a correction of the photon efficiency. At the time of writing this thesis, the photon efficiency correction was not available for this particular Monte Carlo campaign of run dependent samples; therefore, the version intended for run independent samples was used. Despite this difference, this version provides a highly accurate approximation.

A skim is applied on the event to reduce the data set size and accelerate the reconstruction of the events, as explained in Section 2.5.

4.3.4 Selection Cuts

Before applying a multivariate selection, as described in Section 4.3.6, data-driven and preselection cuts are applied. As highlighted in the distributions shown in Figure 4.6, the preselection is not sufficient to obtain a satisfying agreement between data and Monte Carlo. In these distributions, the data set of real data corresponds to 10% of the total data set, and the B meson, radiative dimuon, two-photon, low-multiplicity and initial state radiation samples are combined, as they represent a very small portion of the events.

This additional step ensures not only that the discrimination between signal and background events will be improved in the next step, but also that the variables (or features) entering as input of the BDT are correctly modeling data. The cuts that are applied are detailed in Table 4.2.

A cut is applied on the thrust to remove $q\bar{q}$ events on the lower side and Bhabha events that peak at around 1. The cut on $E_{\text{vis}}^{\text{CMS}}$ allows to cut some two-photon non simulated events in the lower region, and to remove Bhabha events in the higher region.

Two-dimensional cuts relying on $\cos\theta_{\text{thrust}}$, m_{miss}^2 and $\theta_{p_{\text{miss}}^{\text{CMS}}}$ are such that they remove two-photon events [87]. These cuts are defined by slopes, providing a systematic and mathematically derived approach to delimit regions of interest. To compute the slope between the point (x_1, y_1) and (x_2, y_2), the following formula is used:

$$y = x \cdot \frac{y_1 - y_2}{x_1 - x_2} - x_i \cdot \frac{y_1 - y_2}{x_1 - x_2} + y_i \tag{4.7}$$

This strategy enables a more nuanced selection process, leading to increased efficiency while removing problematic regions.

The motivation behind imposing a cut on $m_{\tau,\text{tag}}$ is rooted in the physical motivation

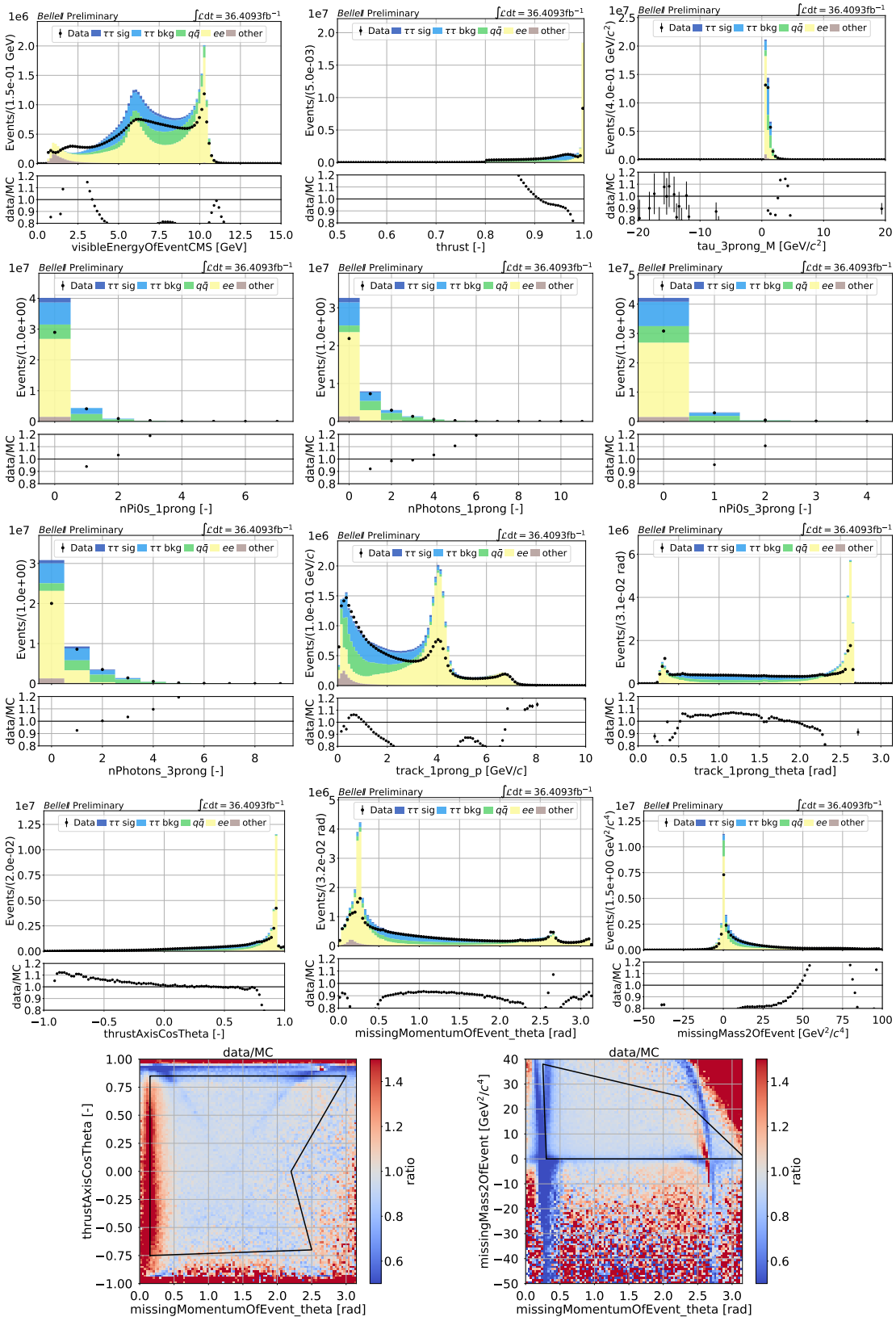


Figure 4.6: *Features distributions.* These distributions were obtained by applying the preselection on the combined pion and kaon samples. The data set representing real data corresponds to 10% of the total data set.

Feature(s)	Cut(s)	Unit(s)
thrust	$\in [0.9, 0.99]$	-
$E_{\text{vis}}^{\text{CMS}}$	$\in [4.5, 9]$	GeV
m_{miss}^2 and $\theta_{p_{\text{miss}}^{\text{CMS}}}$	$m_{\text{miss}}^2 > (-760.0 \times \theta_{p_{\text{miss}}^{\text{CMS}}}) + 228.0$ and $m_{\text{miss}}^2 < (-6.5 \times \theta_{p_{\text{miss}}^{\text{CMS}}}) + 39.6$ and $m_{\text{miss}}^2 < (-26.3 \times \theta_{p_{\text{miss}}^{\text{CMS}}}) + 84.2$ and $m_{\text{miss}}^2 > 0$ and $\theta_{p_{\text{miss}}^{\text{CMS}}} \in [0.15, 2.75]$	GeV$^2/c^4$ and rad
$\cos\theta_{\text{thrust}}$ and $\theta_{p_{\text{miss}}^{\text{CMS}}}$	$\cos\theta_{\text{thrust}} < 0.85$ and $(\cos\theta_{\text{thrust}} > (1.06 \times \theta_{p_{\text{miss}}^{\text{CMS}}}) - 2.34$ or $\cos\theta^{\text{thrust}} < (-2.33 \times \theta_{p_{\text{miss}}^{\text{CMS}}}) + 5.13)$ and $\cos\theta^{\text{thrust}} > (0.02 \times \theta_{p_{\text{miss}}^{\text{CMS}}}) - 0.75$	- and rad
$m_{\tau,\text{tag}}$	$\in [0, 2]$	GeV/c^2
$n_{\pi^0,\text{sig}}$	$= 0$	-
$n_{\gamma,\text{sig}}$	< 4	-
$n_{\pi^0,\text{tag}}$	< 2	-
$n_{\gamma,\text{tag}}$	< 4	-
$p_{\tau,\text{sig}}$	$\in [0.5, 4.5]$	GeV/c
$\cos\theta_{\tau,\text{sig}}$	$\in [-0.4226, 0.8829]$	-
hID	$> 0.5, 0.6, 0.7, 0.8, 0.9$	-
trigger	$= \bigwedge_{t \in \text{trigg}} t$, for trigg $= \{$ *ffy*, *fyo*, *fff*, *ffo*, *lml0*, *lml1*, *lml2*, *lml4*, *lml6*, *lml7*, *lml8*, *lml9*, *lml10*, *lml12*, *hie* $\}$	-
Experiment number	≤ 10	-

Table 4.2: *Selection cuts.* Features and ranges defining the cuts applied to select signal. The cut on the signal hadron ID variable is not defined as several possibilities are explored. The trigger selection requires that at least one of these CDC- and ECL-based trigger bits is activated.

to ensure the presence of a tau lepton on that particular side. Some discrepancies were also observed in higher region of $n_{\pi^0,\text{sig}}$, $n_{\gamma,\text{sig}}$, $n_{\pi^0,\text{tag}}$ and $n_{\gamma,\text{tag}}$. A limit on the momentum and polar angle in the signal side has been set such that it minimises the PID corrections, as they are the lowest in this phase space region.

The cuts were sequentially applied not only to assess the agreement between the

data and Monte Carlo samples, but also in evaluating the background level in the Monte Carlo sample, while accounting for feature correlations.

Several hID selections were tested. A high cut value has the advantage of reducing sources of background, but could have the disadvantage of having correction factors with higher average uncertainties. On the other hand, a lower cut value would reduce the errors associated with the correction factors, but could increase the quantity of background events. This same background could also be useful to retain in reasonable quantities for selection with a multivariate method, given that the latter also categorises the events using background events.

This last point raises another question, which is why the cut is not applied in this case after the multivariate selection, or even why the hID selection is not split into two parts, the first of which would be done before the multivariate selection with a very small value, and the second, after the selection, with a higher value. Checks have shown that it is preferable to keep the cut before selection, without splitting it into two parts, as it leads to better efficiencies and purities.

For simplicity's sake, the same type of variable and working point are used for both pions and kaons, such that:

$$h\text{ID} = K\text{ID} > x \text{ or } \pi\text{ID} > x, \tag{4.8}$$

where x is the working point. No PID requirement is applied on the tag side. The PID working point of 0.9 of the Reweighted hID for the signal side is employed subsequently in all cases where more points would create unnecessary repetitions, given its superior performance, particularly highlighted in Section 4.4.5.

Trigger selection uses several CDC- and ECL-based triggers. The combination of these two types of triggers, which are assumed to be independent, stems from the need to increase the trigger efficiency.

The Experiment 7 and 8 are removed, as they present serious discrepancies between data and Monte Carlo events. They represent 0.6% of the total integrated luminosity, and removing them yields $\mathscr{L}_{\text{int}} = 361.922 \pm 0.021$ (stat.) ± 2.331 (syst.) fb^{-1}.

By applying all selections mentioned, the purity amounts to 31.459%, and the efficiency to 18.663%. Distribution of discriminating features are displayed in Figure 4.7. The Monte Carlo simulations are in general quite in good agreement with the real data, which corresponds to 10% of the total data set. With the exception of the thrust that is showing a descending trend, and the tag side mass that is known to be mismodeled. Again, the B meson, radiative dimuon, two-photon, low-multiplicity and initial state radiation samples are combined.

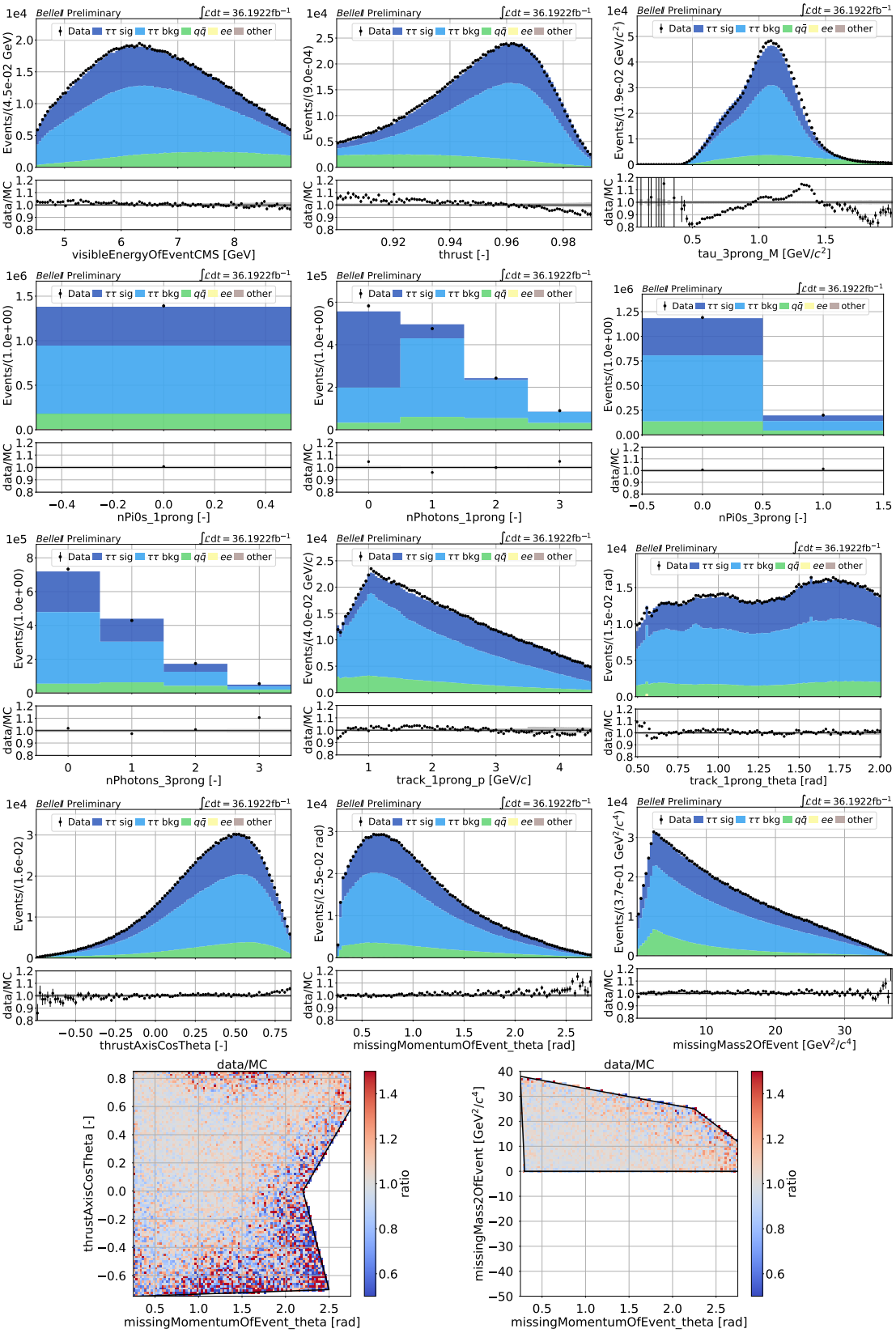


Figure 4.7: *Features distributions.* These distributions were obtained by applying all selections and corrections on the combined pion and kaon samples; in particular, the hID selection is using the Reweighted hID variable and a working point of 0.9.

4.3.5 Corrections

Four additional correction sources have been added to those already applied on Monte Carlo. The first one is related to the neutral selection that was put in place. Unlike in Chapter 3, the neutral pions were not corrected during the reconstruction of the events, because the associated corrections did not allow a straightforward implementation. The second source of correction is related to the background sources. The modeling of these processes sometimes require greater precision, which can be achieved using correction factors computed in side-bands. The two last ones are related to the hID and trigger selections that were used. These corrections are computed using the selected samples, and by using tables of corrections, such as the ones provided in Chapter 3 in the case of hID corrections.

4.3.5.1 Neutral Pion Efficiency

The neutral pion efficiency correction is not implemented in the reconstruction of the events, due to correction factors exceeding 100%, as demonstrated in Table 4.3. This specificity indeed would require that either additional Monte Carlo simulated events would have to be created, or that the data events would have to be removed to compensate the correction. Since neither of these two solutions seem to be conceivable, the corrections are applied after the reconstruction of the events.

	$\cos\theta_1$	$\cos\theta_2$	$\cos\theta_3$	$\cos\theta_4$	$\cos\theta_5$	$\cos\theta_6$
p_1	89.5 ± 3.1	100.5 ± 4.7	101.1 ± 2.7	100.6 ± 4.3	99.9 ± 3.6	88.5 ± 4.0
p_2	89.7 ± 2.1	103.6 ± 2.4	106.6 ± 2.1	103.2 ± 2.3	105.3 ± 2.3	99.3 ± 2.6
p_3	96.6 ± 1.7	103.4 ± 1.5	105.5 ± 2.3	105.7 ± 1.5	104.4 ± 1.3	99.3 ± 1.5
p_4	99.2 ± 2.6	101.1 ± 2.0	101.8 ± 1.5	105.0 ± 1.7	103.7 ± 1.4	99.0 ± 1.6
p_5	97.2 ± 3.1	100.8 ± 2.8	106.6 ± 2.9	101.7 ± 1.8	104.0 ± 1.5	101.1 ± 1.7

Table 4.3: *Correction factors attributed to neutral pions.* These factors were computed using tau decays, and for bins of momentum (p_i) and cosine of polar angle ($\cos\theta_j$). The edges of the bins correspond to 0.6, 0.8, 1.0, 1.5, 2.0, and 3.0 for p_i and -0.65, -0.3 0.0, 0.3, 0.5, 0.7, and 1.0 for $\cos\theta_j$. The corrections are given in percent. The uncertainty next to the central value is only statistical.

The disadvantage that comes with this decision is that the event kinematic and shape variables might not be reproduced properly. As these variables enter the selection of events, this potential issue is consequent. But thorough investigations

of the distribution of these features have revealed that the impact of this issue is negligible.

The corrections presented in Table 4.3 are applied as a weight that is computed in each momentum and cosine of the polar angle bin as:

$$\varepsilon_{\pi^0}^{\text{corr}} = f^m, \tag{4.9}$$

where f is the central value of the correction factor and m is the multiplicity of the neutral pion.

4.3.5.2 Background

Another important aspect involves addressing potential imperfections in the background modeling in order to get a representation of the observed data as accurate as possible. An easy and efficient way of assessing the correction of background modeling is to use side-bands.

This method is employed for non-tau lepton pair background such as $q\bar{q}$ and Bhabha events, which are the two dominant non-tau lepton pair backgrounds with 12.66% and 0.05% after the selection of events. Compared to the $q\bar{q}$ contamination, the Bhabha events represent a second order effect. But, because it is still one order higher than the rest of the background contributions, it was preferable to still assess a correction factor for this type of background process.

Side-bands are obtained by either taking samples orthogonal to the one used in the analysis, *i.e.*, with reversed cuts applied, or by using a very specific region where the events are concentrated. The Bhabha side-band is defined by applying the cuts listed in Table 4.4, whereas the $q\bar{q}$ side-band is defined with the cuts displayed in Table 4.5.

These cuts were determined by following as closely as possible the selection detailed in Section 4.3.4, while increasing as much as possible the number of events that we wanted to observe. The neutral pion efficiency correction computed in the previous section is applied.

In order to understand more the differences between data and Monte Carlo, the side-bands distributions are determined for Chunk1 to Chunk5, and Prompt data sets separately. It is worth noting that the Chunk1 sample corresponds to Experiment 10 only.

The distributions for the side-band that defines Bhabha events are shown in Figure 4.8, whereas the results for $q\bar{q}$ events are displayed in Figure 4.9.

Feature(s)	Cut(s)	Unit(s)
thrust	> 0.98	-
$E_{\text{vis}}^{\text{CMS}}$	> 9.5	GeV
$n_{\pi^0,\text{sig}}$	$= 0$	-
$n_{\gamma,\text{sig}}$	< 4	-
$n_{\pi^0,\text{tag}}$	< 2	-
$n_{\gamma,\text{tag}}$	< 4	-
$p_{\tau,\text{sig}}$	$\in [0.5, 4.5]$	GeV/c
$\cos\theta_{\tau,\text{sig}}$	$\in [-0.4226, 0.8829]$	-
$\lvert m_{\tau,\text{tag}}\rvert$	> 2	GeV/c^2
trigger	$= \bigwedge_{t\in\text{trigg}} t$, for trigg $=\{$ *ffy*, *fyo*, *fff*, *ffo*, *lml0*, *lml1*, *lml2*, *lml4*, *lml6*, *lml7*, *lml8*, *lml9*, *lml10*, *lml12*, *hie* $\}$	-
Experiment number	≤ 10	-

Table 4.4: *Bhabha side-band cuts.* The neutral and trigger selections were kept identical to the signal selection as to keep a comparable phase space.

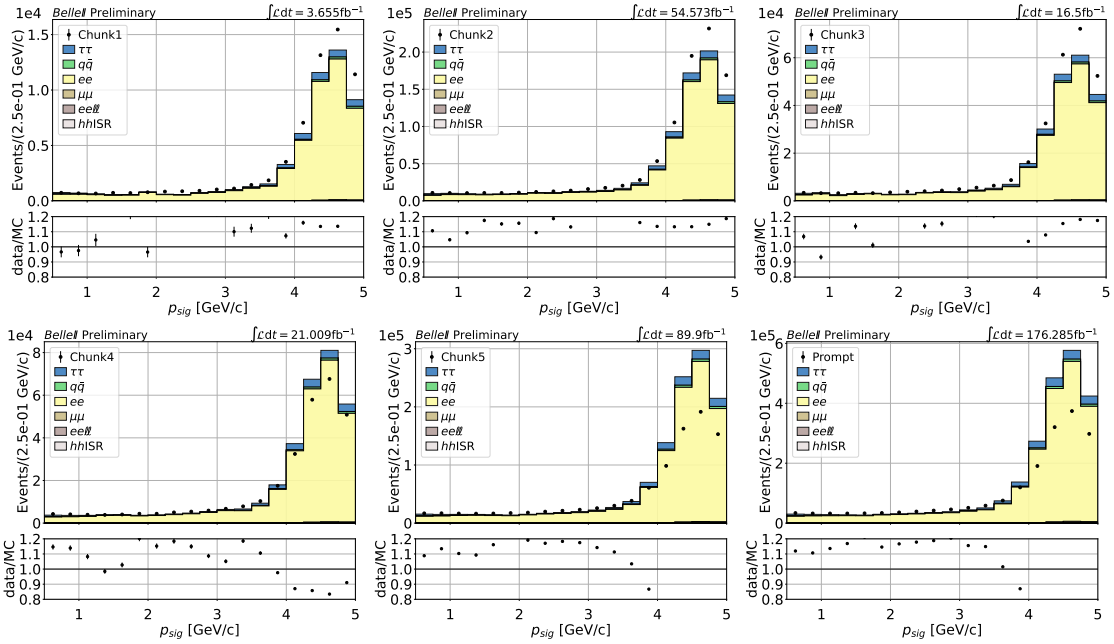


Figure 4.8: *Bhabha side-band.* The distributions are divided into the different main data samples. In this case, Chunk1 corresponds to Experiment 10 only.

Feature(s)	Cut(s)	Unit(s)
thrust	$\in [0.9, 0.99]$	-
$E_{\text{vis}}^{\text{CMS}}$	$\in [4.5, 9]$	GeV
m_{miss}^2 and $\theta_{p_{\text{miss}}^{\text{CMS}}}$	$m_{\text{miss}}^2 > (-760.0 \times \theta_{p_{\text{miss}}^{\text{CMS}}}) + 228.0$ and $m_{\text{miss}}^2 < (-6.5 \times \theta_{p_{\text{miss}}^{\text{CMS}}}) + 39.6$ and $m_{\text{miss}}^2 < (-26.3 \times \theta_{p_{\text{miss}}^{\text{CMS}}}) + 84.2$ and $m_{\text{miss}}^2 > 0$ and $\theta_{p_{\text{miss}}^{\text{CMS}}} \in [0.15, 2.75]$	GeV$^2/c^4$ and rad
$\cos\theta_{\text{thrust}}$ and $\theta_{p_{\text{miss}}^{\text{CMS}}}$	$\cos\theta_{\text{thrust}} < 0.85$ and $(\cos\theta_{\text{thrust}} > (1.06 \times \theta_{p_{\text{miss}}^{\text{CMS}}}) - 2.34$ or $\cos\theta^{\text{thrust}} < (-2.33 \times \theta_{p_{\text{miss}}^{\text{CMS}}}) + 5.13)$ and $\cos\theta^{\text{thrust}} > (0.02 \times \theta_{p_{\text{miss}}^{\text{CMS}}}) - 0.75$	- and rad
$n_{\pi^0,\text{sig}}$	$= 0$	-
$n_{\gamma,\text{sig}}$	< 4	-
$n_{\pi^0,\text{tag}}$	< 2	-
$n_{\gamma,\text{tag}}$	< 4	-
$p_{\tau,\text{sig}}$	$\in [0.5, 3]$	GeV/c
$\cos\theta_{\tau,\text{sig}}$	$\in [-0.4226, 0.8829]$	-
$\lvert m_{\tau,\text{tag}}\rvert$	> 2	GeV/c^2
trigger	$= \bigwedge_{t\in\text{trigg}} t$, for trigg $= \{$ *ffy*, *fyo*, *fff*, *ffo*, *lml0*, *lml1*, *lml2*, *lml4*, *lml6*, *lml7*, *lml8*, *lml9*, *lml10*, *lml12*, *hie* $\}$	-
Experiment number	≤ 10	-

Table 4.5: *$q\bar{q}$ side-band cuts.* The selection was kept very similar to the signal selection as to keep a comparable phase space.

From these distributions, it is possible to extract the correction factor for each sample by performing a ratio of the number of event in data, corrected by the other sources of Monte Carlo background, over the number of $q\bar{q}$ or Bhabha events, as previously done in Equation (3.12):

$$\varepsilon_{q\bar{q}}^{\text{corr}} = \frac{N_{data} - N_{MC}^{\text{other}}}{N_{MC}^{q\bar{q}}} \quad \text{and} \quad \varepsilon_{\text{Bhabha}}^{\text{corr}} = \frac{N_{data} - N_{MC}^{\text{other}}}{N_{MC}^{\text{Bhabha}}}. \tag{4.10}$$

The correction of the data event is recommended, as there is a non-negligible con-

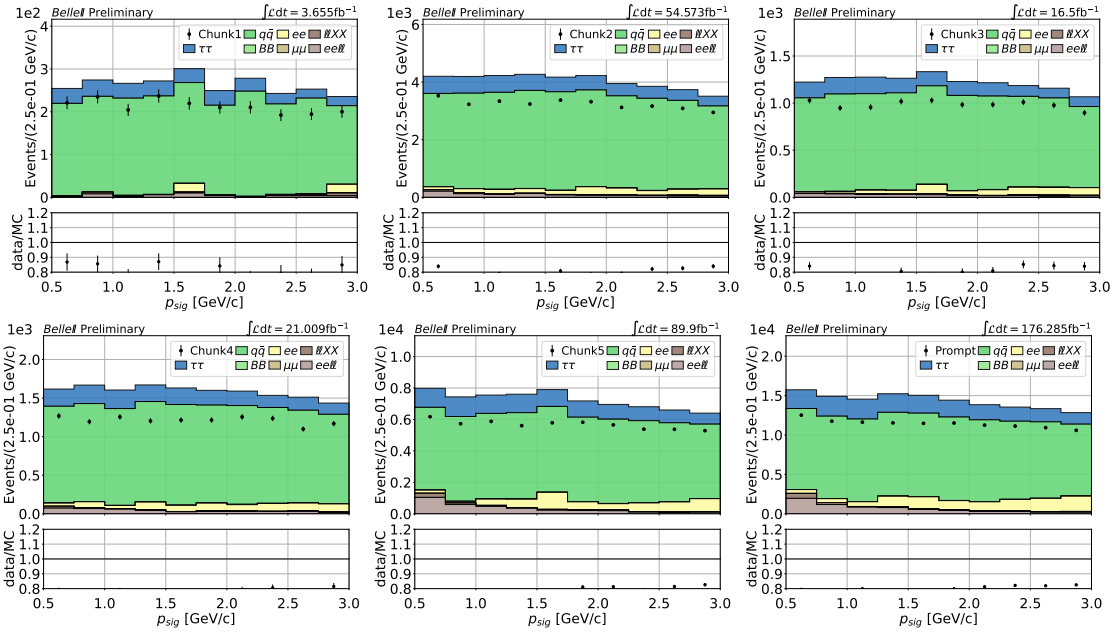
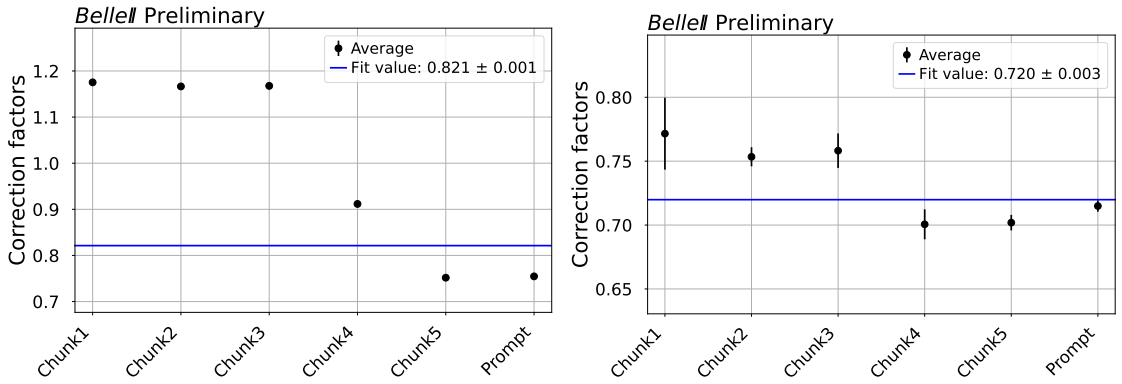


Figure 4.9: *$q\bar{q}$ side-bands.* The distributions are divided into the different main data samples. In this case, Chunk1 corresponds to Experiment 10 only.

tamination from tau lepton pair events. These factors are reported in Figure 4.10 for each sample.



(a) Bhabha correction factors

(b) $q\bar{q}$ correction factors

Figure 4.10: *Correction factors reported for each main data sample.* The global correction factor is determined by a fit, and is reported in the legend. On the left-hand side is shown the factors for Bhabha events, and the $q\bar{q}$ are on the right-hand side.

This factor was determined by fitting the averages with a polynomial of degree 0. A mean squared error (MSE) is subsequently estimated. The pattern observed in Bhabha events is unexpected, and can be due to several factors combined. One possible explanation could be that there might be a trigger bit that is prescaled

during the first data-taking periods, and which was not taken into account in the Monte Carlo simulation. But since the Bhabha correction represents a second order correction, this effect does not have a significant impact on our measurement.

These correction factors, used to correct the corresponding non-tau lepton pair background events are applied, and correspond to $\varepsilon_{q\bar{q}}^{\text{corr,fit}} = 0.720 \pm 0.003$ and $\varepsilon_{\text{Bhabha}}^{\text{corr,fit}} = 0.821 \pm 0.001$.

4.3.5.3 Trigger

The trigger selection is corrected with the ratio of efficiency in data and Monte Carlo simulation. Usually, the trigger efficiency is computed as

$$\varepsilon_t = \frac{N_{t_m \wedge t_r}}{N_{t_r}}, \tag{4.11}$$

where either CDC- or ECL-based trigger bits would be used as main triggers (t_m), and the other type would be used as a reference trigger (t_r), as done in Equation (3.11) for the performance study. In the case of this measurement, it is slightly different, as both types of triggers are combined, and instead requires the use of inefficiencies, defined as:

$$\varepsilon_{1-\text{inef}} = 1 - (1 - \varepsilon_{ECL})(1 - \varepsilon_{CDC}), \tag{4.12}$$

where ε_t, with $t = ECL, CDC$ is computed as in Equation (4.11). The ECL reference trigger bits correspond to the whole set of CDC main trigger bits, whereas the CDC reference triggers are only *hie* and *lml0*. The idea behind this computation is that this inefficiency should reproduce the absolute efficiency computed as:

$$\varepsilon^{abs} = \frac{N_{ECL \vee CDC}}{N_{tot}},$$

but unlike the latter, can be applied to the data sample. Indeed, the total number of events N_{tot} cannot be obtained in data reliably, as other selections are by default in place. To make sure that this inefficiency computation can be used, it is necessary to check that the inefficiency $\varepsilon_{1-\text{inef}}$ reproduces in Monte Carlo a similar behaviour as ε^{abs}.

Figure 4.11 shows the comparison between these two efficiencies. There are three different cases of samples, depending on the hID requirement: the samples employed were either combined, or separated into pion and kaon samples. In this

case, the hID requirement is defined by the Reweighted PID variable, with a working point of 0.9.

The bins are also different, as to analyse the comparison in momentum and polar angle bins separately. The maximum difference between $\varepsilon_{1-\text{inef}}$ and ε^{abs} is about $\mathcal{O}(0.2)\%$ for the momentum-binned results, and about the same for the polar angle-binned results, except at high cosine values, where the difference amounts to approximately 5%.

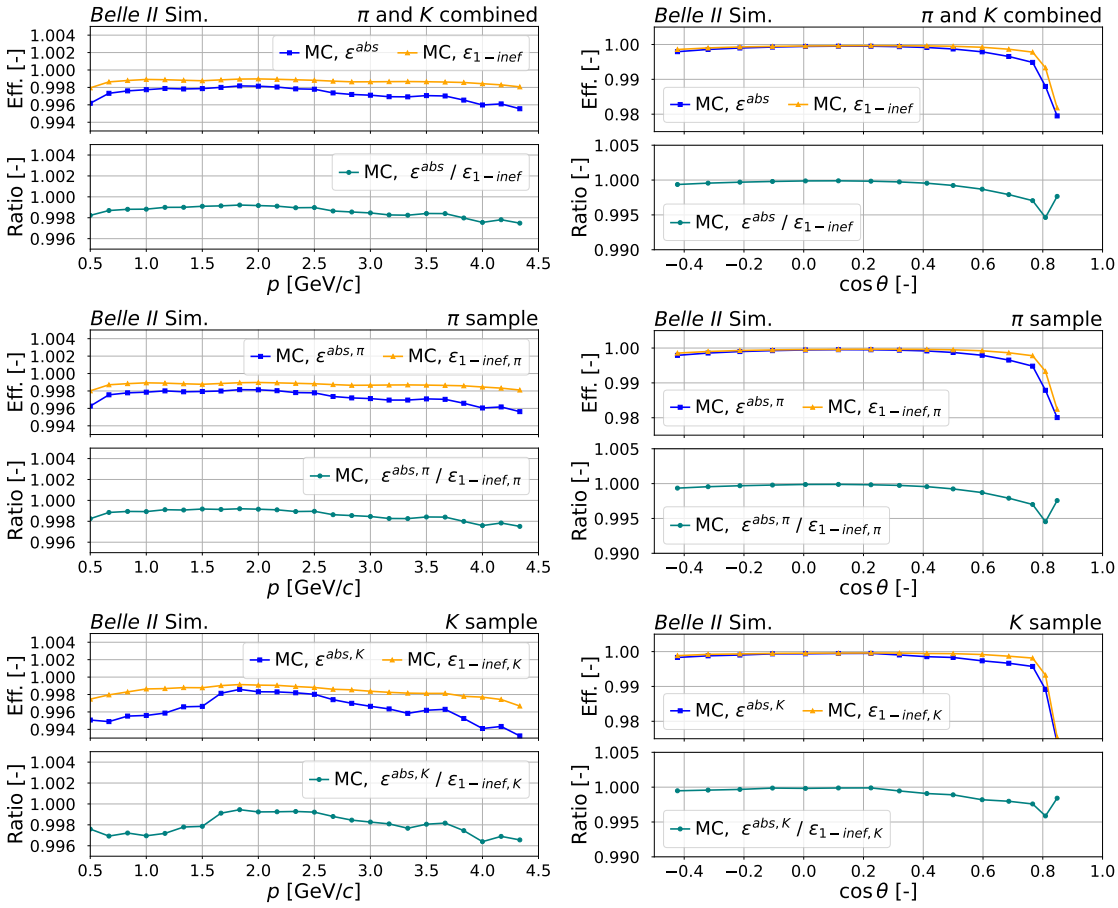


Figure 4.11: *Comparison in Monte Carlo sample between the inefficiency* $\varepsilon_{1-\text{inef}}$ *and absolute efficiency* ε^{abs}. A Reweighted hID cut of 0.9 is applied on the samples. The comparison is done between combined samples (top), pion samples (middle) and kaon samples (bottom), separated by the corresponding hID cut. Bins of momentum are used on the left-hand side, and bins of polar angle are used on the right-hand side.

This substitution technique is validated by these results, as they still translate a satisfying Monte Carlo modeling of the data. The resulting correction factors are deduced by computing the ratio of inefficiencies between data and Monte Carlo, as follows:

$$\varepsilon^{\text{corr}}_{1-\text{inef}} = \frac{\varepsilon^{\text{data}}_{1-\text{inef}}}{\varepsilon^{\text{MC}}_{1-\text{inef}}} \tag{4.13}$$

The resulting correction factors are shown in Figure 4.12. The corrections used for the trigger selection are only binned in momentum for simplicity. These correction factors are really close to unity ($\mathcal{O}(0.998)$), and will be translated into small systematic uncertainties, as discussed in Section 4.4.5.

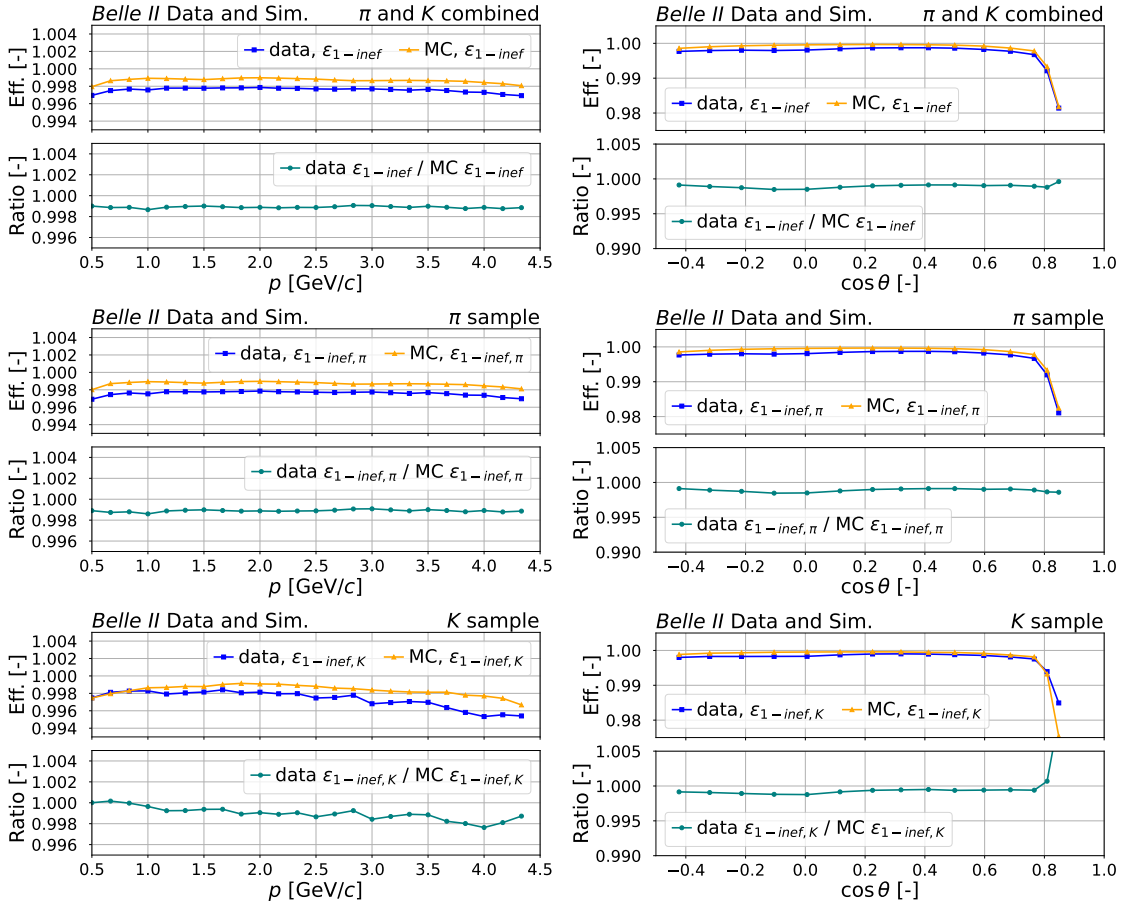


Figure 4.12: *Comparison of inefficiency* $\varepsilon_{1-\text{inef}}$ *between Monte Carlo and data samples.* A Reweighted hID cut of 0.9 is applied on the samples. The comparison is done between combined samples (top), pion samples (middle) and kaon samples (bottom), separated by the corresponding hID cut. Bins of momentum are used on the left-hand side, and bins of polar angle are used on the right-hand side.

4.3.5.4 Particle Identification (PID)

The particle identification selection is corrected using ratios of efficiency in data and Monte Carlo, similar to the ones provided in Chapter 3.

Several corrections have to be implemented, and comprises the pion and kaon efficiencies, as well as their corresponding misidentification rates. Using the notation introduced in Section 3.4.1, these rates are of hadronic nature, such as "$\pi \to K$" and "$K \to \pi$", or leptonic nature, as in the case of "$\pi \to e$", "$\pi \to \mu$", "$K \to e$" or "$K \to \mu$". For the sake of clarity, they are summarised in Table 4.6. All hID variables and working points (WPs) are accessible for the hID efficiencies, and hadronic misidentification rates. Whereas for the leptonic misidentification rates, only the Global hID variable and working point of 0.6 is available at the time of this thesis.

Correction type	Particles involved	Availability
particle ID efficiency	π, K	All hID variable, WPs
particle misID rate	$\pi \to K$, $K \to \pi$	All hID variable, WPs
	$e \to K$, $\mu \to K$, $e \to \pi$, $\mu \to \pi$	Global hID, WP 0.6

Table 4.6: *Summary of PID corrections.* The same convention to describe the fake rates is employed throughout this chapter. The availability of the corrections is also indicated.

The correction are binned in momentum and polar angle. The bins used for this measurement are the same as for the performance study. For the reader's convenience, they are repeated in Table 4.7. The sketch of the Belle II detector from the side with the regions corresponding to the polar angle bins is available in Figure 3.2, and will not be repeated here to avoid redundancy. If not otherwise specified in the following sections of this chapter, this binning is the one used by default for all aspects related to particle identification corrections.

4.3.6 Multivariate selection

For this measurement, a Boosted Decision Tree (BDT) from the *LightGBM* implementation [88] was used. One specific feature of this implementation is that the decision trees are not grown level-wise, but leaf-wise, such that not a maximum depth, but a maximum number of leaves is set as a hyper parameter. Another advantage is that it utilises two novel techniques called Gradient-Based One-Side Sampling (GOSS) and Exclusive Feature Bundling (EFB), which allow the algorithm to run faster while maintaining a high level accuracy.

The test and training sets are split as 30% and 70% of the total sample respectively. Other fractions are possible, but given the size of the total sample, this splitting should yield two data sets that can provide a reliable and reproducible method.

momentum	polar angle		
p [GeV/c]	θ [rad]	θ [deg]	$\cos\theta$ [-]
[0.5, 1]	[0.297, 0.489]	[17, 28]	[0.883, 0.956]
[1, 1.5]	[0.489, 0.698]	[28, 40]	[0.766, 0.883]
[1.5, 2]	[0.698, 1.047]	[40, 60]	[0.500, 0.766]
[2, 2.5]	[1.047, 1.344]	[60, 77]	[0.225, 0.500]
[2.5, 3]	[1.344, 1.676]	[77, 96]	[-0.104, 0.225]
[3, 3.5]	[1.676, 2.007]	[96, 115]	[-0.423, -0.104]
[3.5, 4.5]	[2.007, 2.321]	[115, 133]	[-0.682, -0.423]
	[2.321, 2.618]	[133, 150]	[-0.866, -0.682]

Table 4.7: *Momentum and polar angle binning.* This binning is used for both pion and kaon corrections. The polar angle bins are given in radian, degrees and as their cosine value.

The hyper parameters are first set to default as to mitigate the overtraining effects that can occur if the method is too specific. In a second step, the hyper parameters are optimised with the *Optuna* [89] library. As the results were showing a poorer performance compared to the use of default parameters, we decide to proceed with the default parameters. Their values are indicated in this documentation [90].

The multivariate selection is relying on distributions shapes defined by the features that are used as input. Weights are needed to allow these features to describe data as accurately as possible, to ensure that the selection trained on Monte Carlo is indeed reproducible on data. These weights translate the luminosity of the different samples that are used, but also all the corrections applied.

But another effect also has to be taken into account. As the pion and kaon samples are trained together to mitigate differences which would not cancel in the ratio $R_{K/\pi}$, a disparity arises. The kaon signal is by definition smaller than the pion signal, due to their different branching fractions. An approximate factor of 16 separates the expected number of pion and kaon events. The issue with this non negligible difference is that the multivariate method would most likely only catch features from the pion sample, and not the kaon. Therefore, an additional weight is attributed to the events, and Figure 4.13 illustrates this process.

First, the kaon signal events are weighted to match the number of pion signal events. Then, the total signal is weighted to match the total background number of events. These weights prevent effectively biases that can be introduced without,

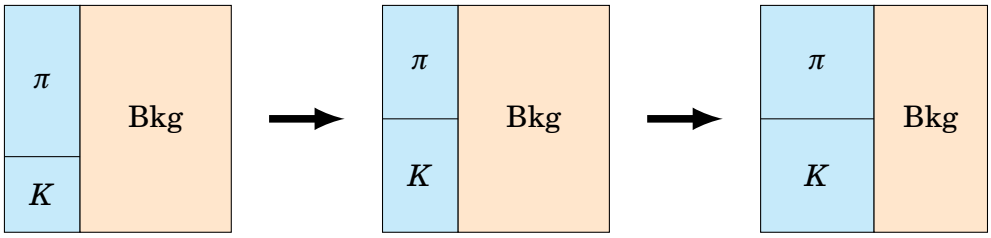


Figure 4.13: *Illustration of samples weighting.* The procedure is the following: the kaon signal events are weighted to match the number of pion signal events, and then the total signal is weighted to match the total background number of events.

and showed in particular a better performance than without this weight.

A total of 24 features are tested as input of the method. These features are the nine CLEO cones (CC_i), the four Fox-Wolfram normalised moments (R_i), the tag side sorted transverse momenta ($p^{i\text{th}}_{T,\text{tag}}$), the tag side energy in the CMS frame ($E^{\text{CMS}}_{\tau,\text{tag}}$), the missing mass squared of the event (m^2_{miss}), missing momentum in the lab and CMS frame (p_{miss} and $p^{\text{CMS}}_{\text{miss}}$), the polar angle component of the missing momentum of the event in the CMS frame ($\theta_{p^{\text{CMS}}_{\text{miss}}}$), the thrust, the cosine of the polar angle component of the thrust axis ($\cos\theta_{\text{thrust}}$) and the visible energy of the event in the CMS frame ($E^{\text{CMS}}_{\text{vis}}$). Figures 4.14 and 4.15 show the distribution of these features as they are used as input of the method. As in the previous distributions of features, the B meson, radiative dimuon, two-photon, low-multiplicity and initial state radiation samples are combined. Some CLEO cones distributions are shown with a logarithmic Y-axis scale to enhance their visibility.

The *scikit-learn* library [91] offers several methods to rank features. A 5-fold technique [92] with 3 repetitions is applied on the samples to split them into several test and training sets. The performance of the features is evaluated while reducing the dependency of the evaluation on the sample. The importance of each feature is extracted at each stage, and the average is computed. The result is visible in Figure 4.16.

It is worth mentioning that other hID selections, using both Global and Reweighted hID variables, were also employed to rank these features. The results are very similar to the one shown for the hID selection using the Reweighted hID variable and a working point of 0.9, and are available in Appendix B.1.

Correlation coefficients are computed using Equation (4.2). A matrix of correlation is obtained for both background and signal events. These matrices are shown in Figures 4.17 and 4.18.

The determination of the ranking of the features and their correlations raise impor-

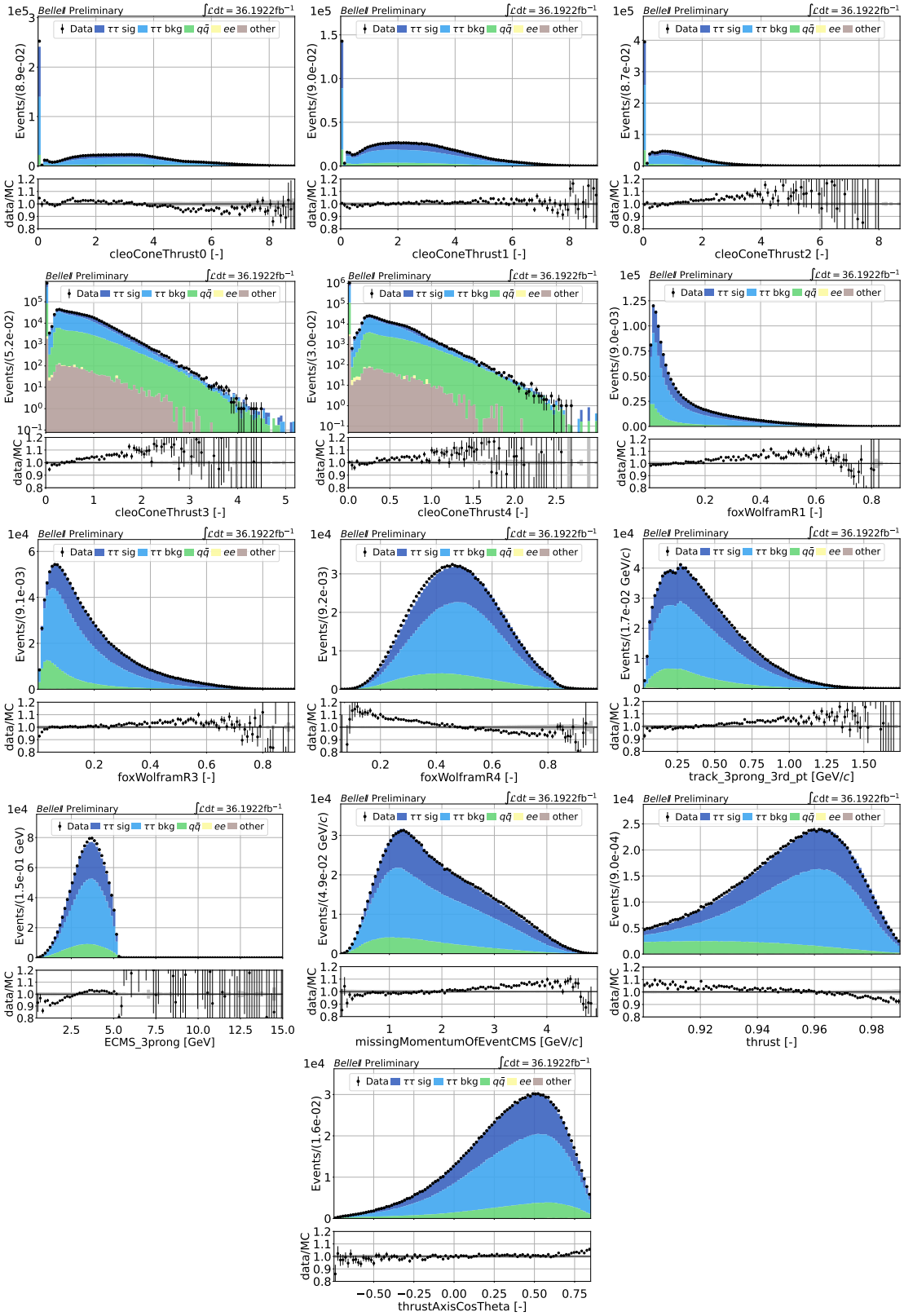


Figure 4.14: *Distribution of features used as BDT input (part 1).* The selection and corrections are applied; in particular, the hID selection is using the Reweighted hID variable and a working point of 0.9.

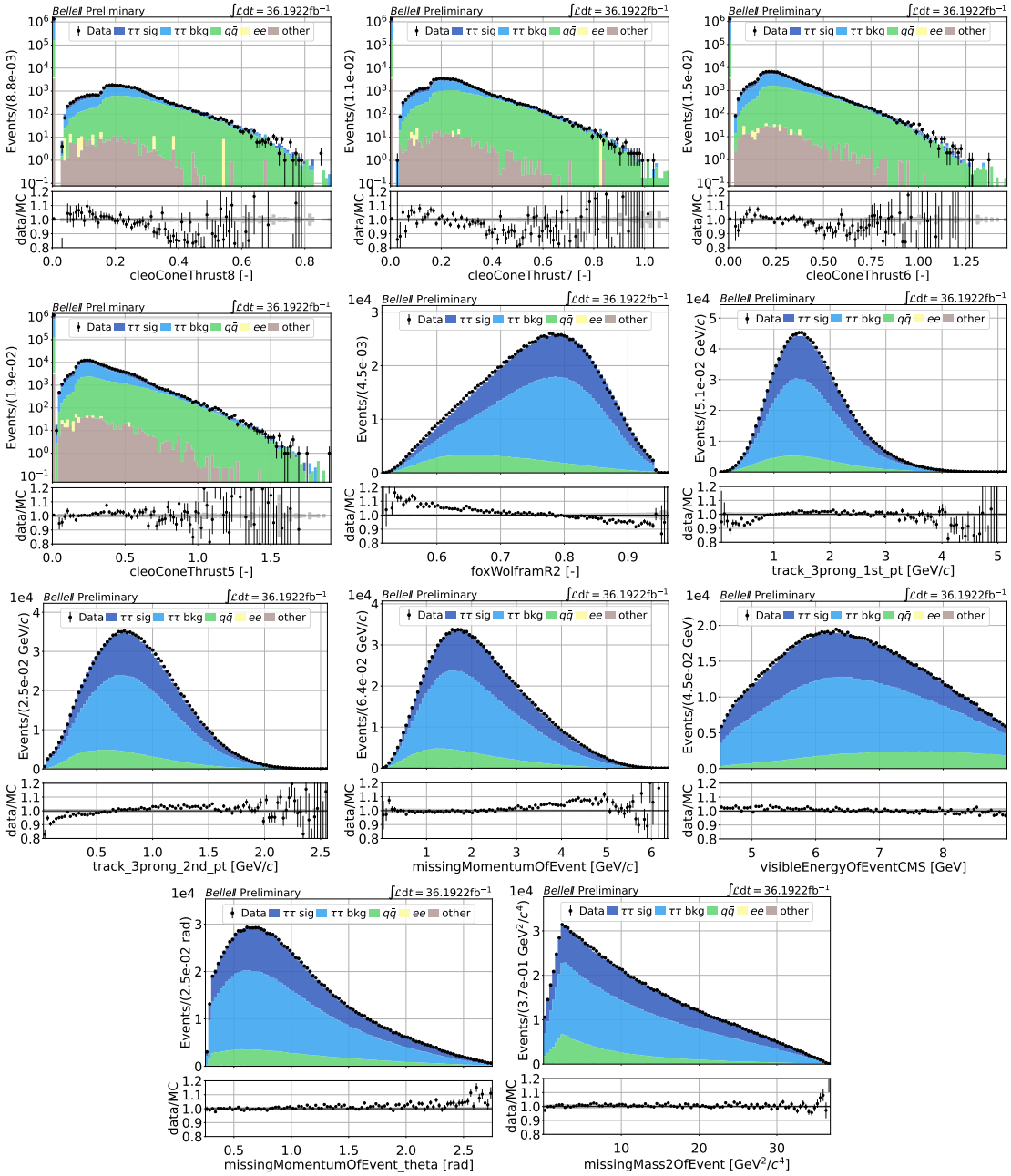


Figure 4.15: *Distribution of features used as BDT input (part 2).* The selection and corrections are applied; in particular, the hID selection is using the Reweighted hID variable and a working point of 0.9.

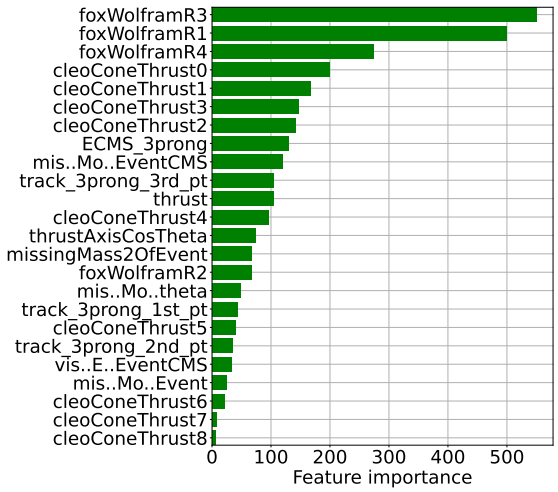


Figure 4.16: *Feature importance.* Computed using a k-fold technique and repetitions. The hID selection is using the Reweighted hID variable and a working point of 0.9.

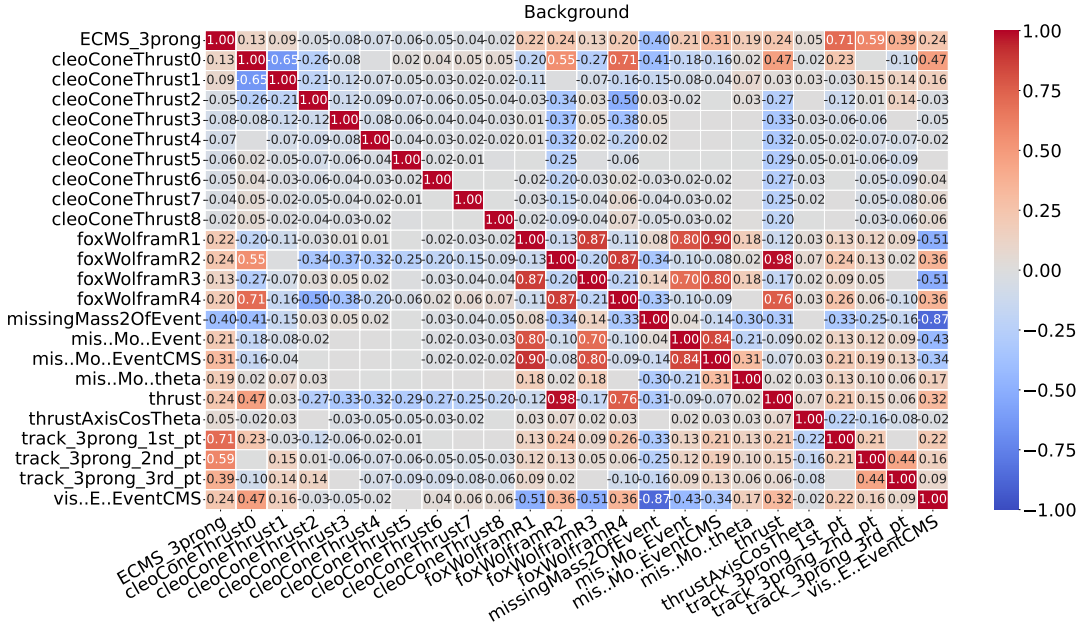


Figure 4.17: *Background correlations.* The correlations are computed using the weighted Pearson's correlations. The hID selection is using the Reweighted hID variable and a working point of 0.9.

tant points to consider when selecting the variables to be included in the method.

First, in Figures 4.17 and 4.18, the Fox-Wolfram normalised moment R_2 and the thrust are demonstrating a correlation value which is very close to one. Since the ranking of R_2 is lower than that of the thrust, as demonstrated in Figure 4.16, this feature is removed from the set of features entering the BDT selection.

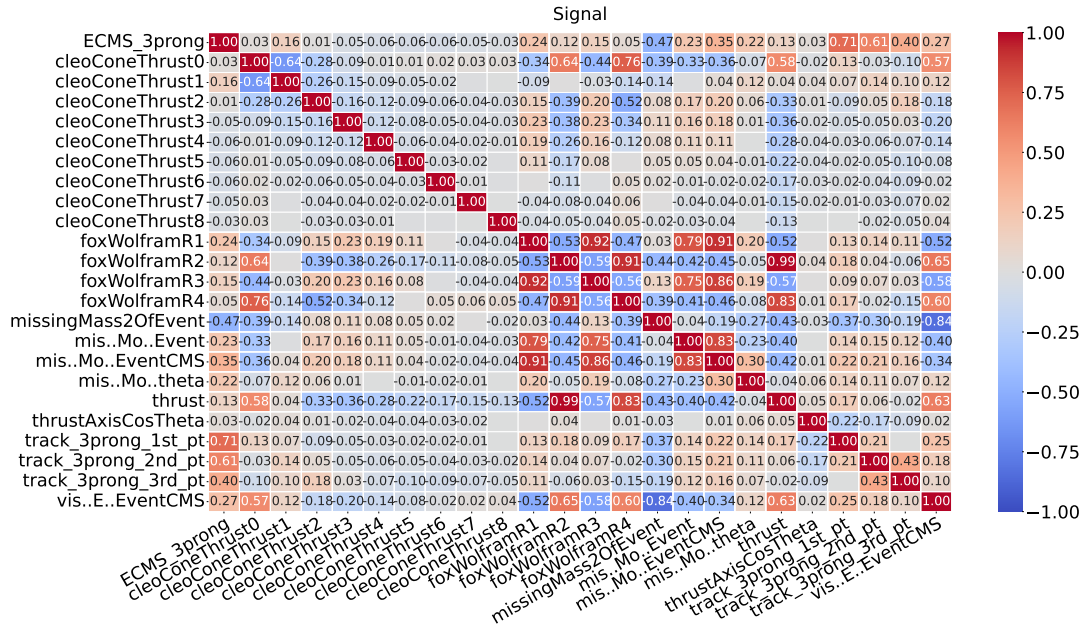
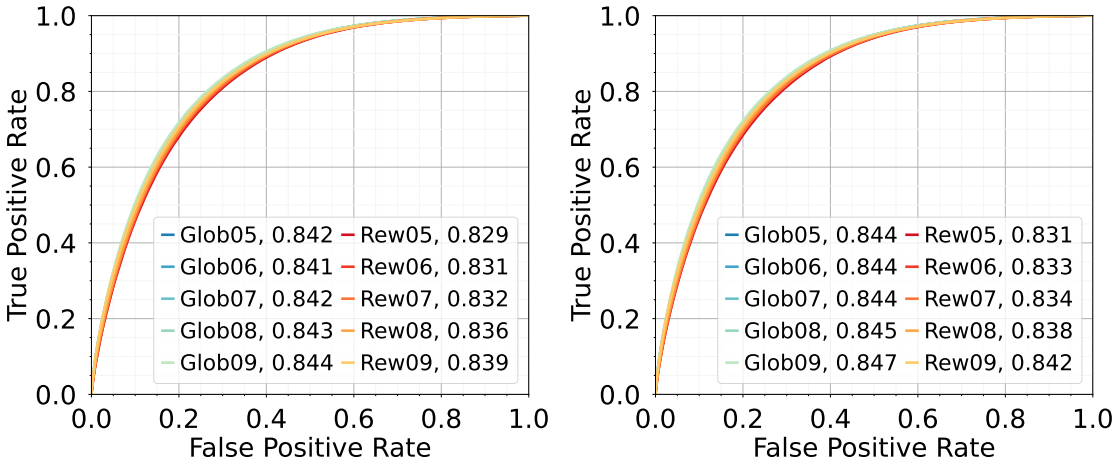


Figure 4.18: *Signal correlations.* The correlations are computed using the weighted Pearson's correlations. The hID selection is using the Reweighted hID variable and a working point of 0.9.

A closer look to the ranking makes us realise that there is a significant difference of importance observed between the initial and final features. This observed difference prompts the removal of features, starting from the least important ones. The question arises, however, as to how many features can be removed, given that a clear separation is not observed. Several attempts lead to the conclusion that removing 10 features does not significantly impact the performance of the BDT. This is particularly highlighted in Figures 4.19 and 4.21, where both configurations using either 23 or 13 features are shown. And this applies to all hID selections tested. The features kept for the multivariate selection, and therefore used as input of the BDT, are shown in Figure 4.14. The 11 features removed, on the other hand, are shown in Figure 4.15. The process of removing features is not only aimed at simplifying the selection criteria but also at avoiding the use of features that could potentially impact the accurate assessment of systematic uncertainty, a topic that will be further discussed in Section 4.4.5, and is more particularly relevant for Section 4.4.5.5.

The performance of the multivariate method is evaluated through the AUC computation and Kolmogorov-Smirnov test. The method demonstrated AUC values ranging from 0.829 to 0.847, as represented in Figure 4.19. These values indicate that the method struggles to classify the events. This difficulty may be coming from different reasons. One possibility is that the features used as input of the method

are not powerful enough to discern all patterns in signal and background. Unfortunately, features that are more discriminating are also not modeled well enough, and therefore cannot be added to the set of features entering the BDT.



(a) BDTs with 13 features

(b) BDTs with 23 features

Figure 4.19: *ROC curves.* The area under the curve (AUC) corresponding to the hID selection using the Reweighted hID variable and a working point of 0.9 is equal to 0.839 in the case of the BDT trained with 13 features (left). Also shown are the performances obtained with 23 features (right).

The Kolmogorov-Smirnov test is applied on test and training sets of signal and background samples. Figure 4.20 shows the distribution of these samples. The test yields a very small statistic value, and a p-value that would indicate that the hypothesis of overtraining remains improbable. For a more complete picture, we have evaluated the statistic and a p-values for all hID selections. The result is shown in Figure 4.21, and demonstrate the same conclusion formulated before. This result is expected as no hyper parameter tuning was introduced in our models, which helps prevent excessive complexity and in turn mitigates overfitting effects.

The output of the Boosted Decision Tree (BDT) is a value that falls within the range of 0 to 1. But, in Figure 4.22, the BDT output distribution does not extend until 1. There are several reasons that can explain a reduced power of the BDT in identifying with certainty the signal events. A first hypothesis was that the use of combined $\tau^- \to K^- \nu_\tau$ and $\tau^- \to \pi^- \nu_\tau$ signals would have introduced some supplementary difficulties, as these decays have different kinematics. But training a method on the separated samples did not result in an extended output distribution. Another convincing explanation that was already discussed for the AUC value, is that the variables that are used as input of the method are not powerful enough to discriminate the signal more accurately. In any case, this particularity does not in-

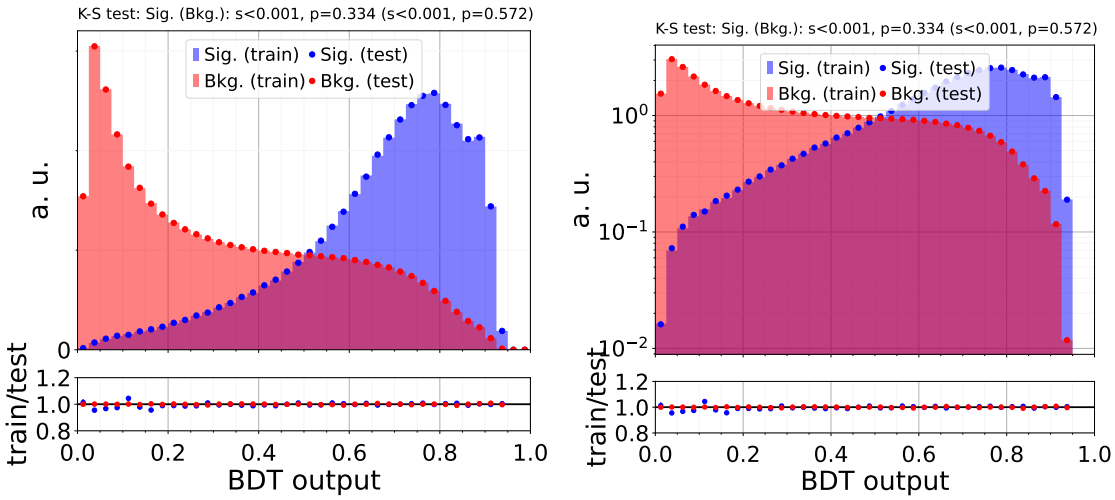
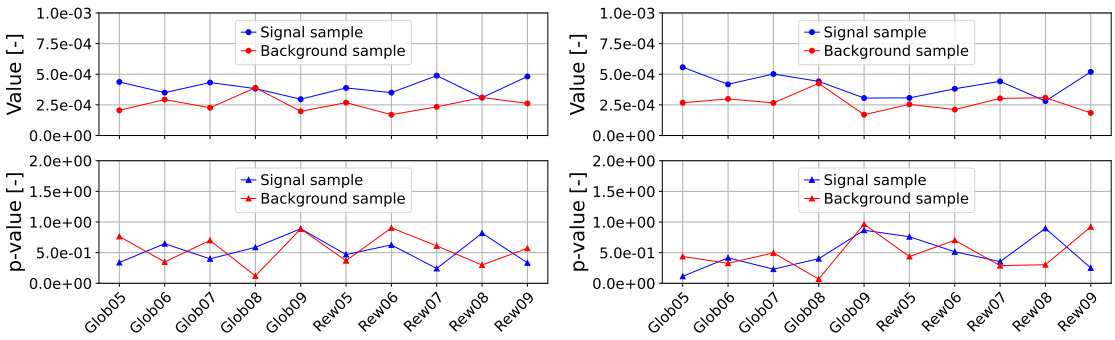


Figure 4.20: *Distributions of signal and background data sets divided into test and training sets.* The method is trained on the training set, and is applied on the test set to see if the method is reproducible. The distributions are displayed in linear and logarithmic Y-axis scale. A Kolmogorov–Smirnov test value is indicated for both data sets. The *h*ID selection is using the Reweighted *h*ID variable and a working point of 0.9.



(a) BDTs with 13 features

(b) BDTs with 23 features

Figure 4.21: *Kolmogorov–Smirnov test values.* The (statistical) value as well as the p-values are indicated for all *h*ID selections in the case of the BDT trained with 13 features (left). Also shown are the performances obtained with 23 features (right).

terfere with the rest of the analysis, as only a cut on the BDT output is needed, and which could be placed at different values. To make the distribution of the different samples clearer, they are separated and shown individually as well. As previouly done, the B meson, radiative dimuon, two-photon, low-multiplicity and initial state radiation samples are combined into one sample.

Since the tau lepton pair background is the dominant source of background, the different decay modes associated to the signal side are studied in Figure 4.23 for

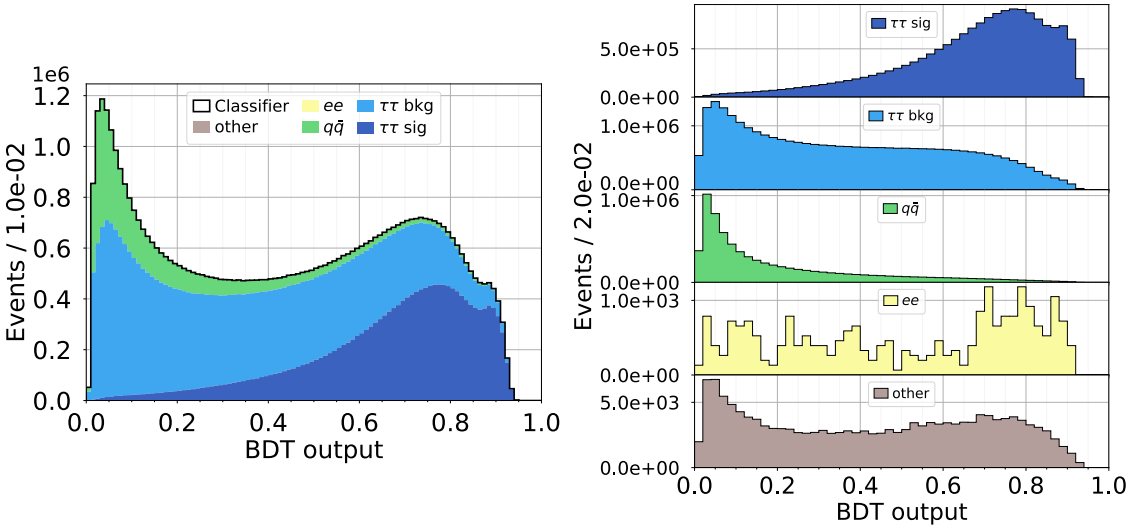


Figure 4.22: *BDT output distributions.* The distributions are displayed either in stacked or unstacked format. The hID selection is using the Reweighted hID variable and a working point of 0.9.

the combined sample, as well as the kaon and pion samples separately. In the case of the combined sample, it appears that the mode associated to the decay $\tau^- \to \pi^-\pi^0\nu_\tau$ is the most dominant one, and has a slight peak around 0. The four consecutive modes $\tau^- \to \pi^-\pi^0\pi^0\nu_\tau$, $\tau^- \to \mu^-\nu_\tau\bar{\nu}_\mu$, $\tau^- \to \pi^- K^0_L\nu_\tau$ and $\tau^- \to K^-\pi^0\nu_\tau$ are also present, but in lower quantities, and are more constant across the whole BDT output range. Whereas the rest of the events, characterised by "*other*" in the legend, are more concentrated towards lower values.

Given that the pion sample constitutes the majority of the combined sample, it is expected to observe that the dominant modes of the combined samples are also present in the pion sample. As for the kaon sample, it exhibits a broader range of modes compared to the pion sample, reflecting its inherently suppressed nature.

From the BDT output distribution, it is possible to evaluate the purity, efficiency and figure of merit, all described in Equations (3.1), (3.2) and (3.3), in order to assess further the performance of the BDT method. Figure 4.24 shows the efficiency and the purity performances, obtained for a serie of cuts applied on the BDT output of several working points of Reweighted and Global hID variables. The result is separated into the two different channels, with the value of the BaBar collaboration [33] indicated as a reference. In particular, their efficiency is recalculated to align with Equation (3.2) to ensure a more equitable comparison.

It is demonstrated later, in Section 4.4.5, that the hID selection relying on the Reweighted hID variable with a working point of 0.9 used in the results shown pre-

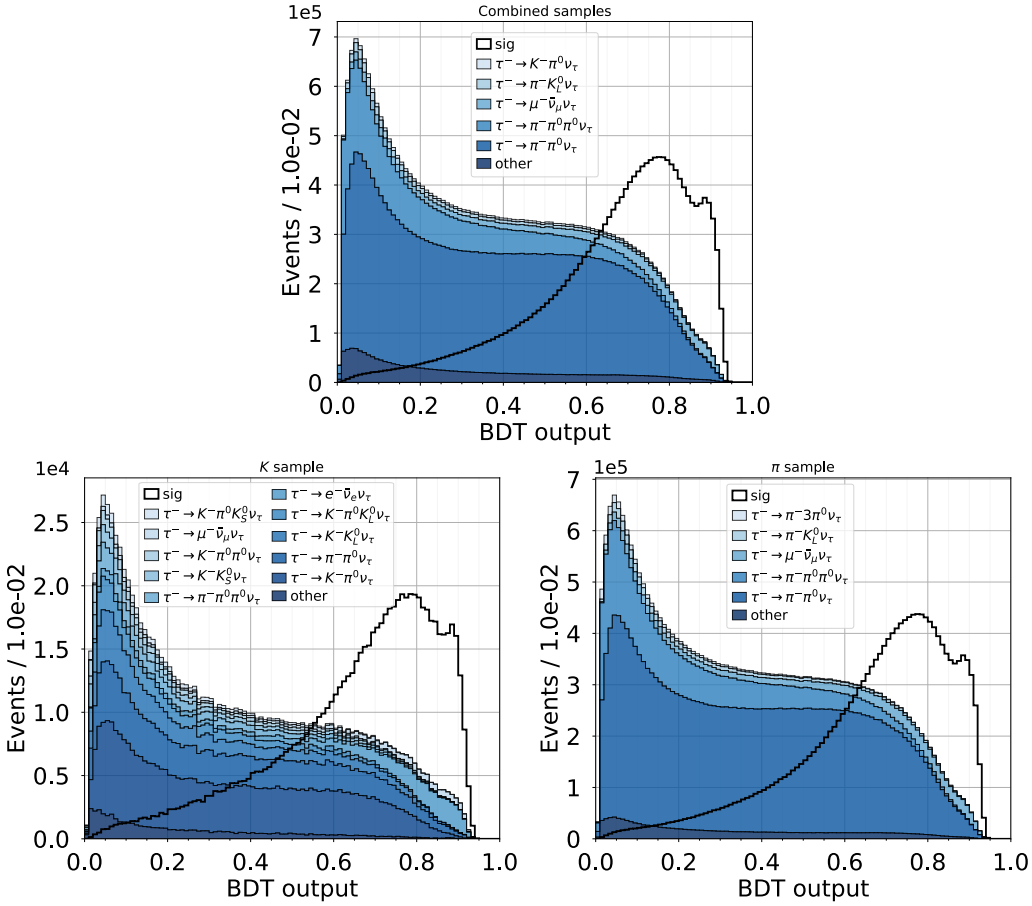
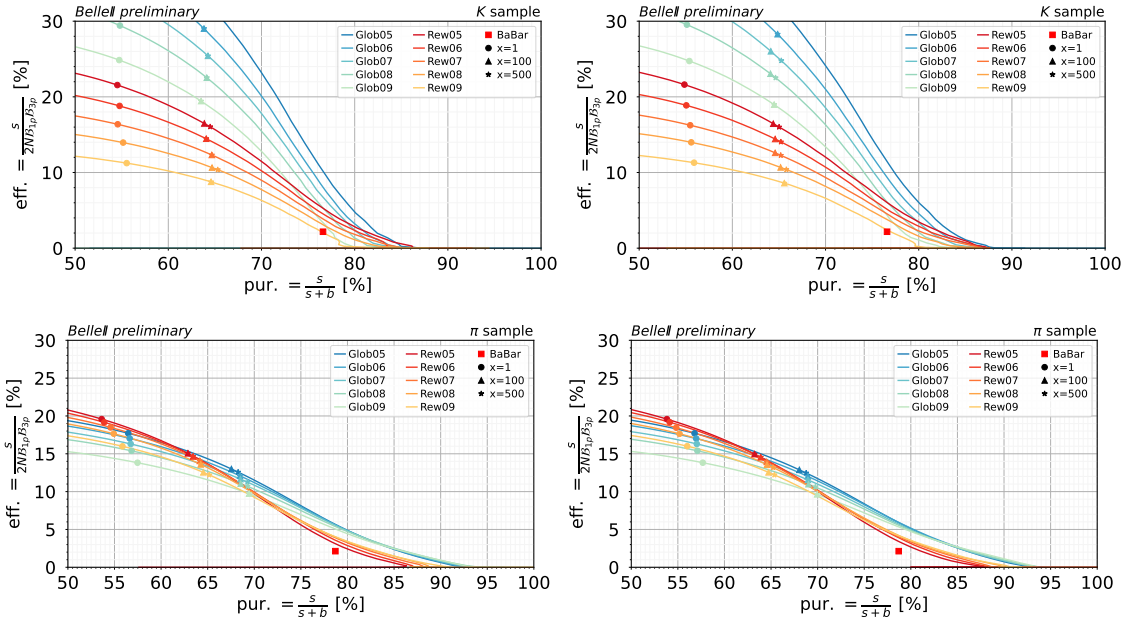


Figure 4.23: *Tau lepton pair BDT output distributions.* The stacked distributions are displayed for the combined, K and π samples, with the main decays shown. They are sorted in the legend according to their size, apart from "*other*", which represents all other modes. The hID selection is using the Reweighted hID variable and a working point of 0.9.

viously indeed minimises the systematic uncertainty. Even though in Figure 4.24, the very same selection seems to be leading to one of the worst efficiency and purity performance.

The selection of an optimal cut on the BDT output is essential. One approach involves evaluating a figure of merit (FOM) with varying penalty factors on the background, but this method exclusively considers the statistical aspect of the sample, and does not account for systematic uncertainties. To address this limitation, multiple cut values are explored through subsequent analysis as a first approach. The method is relatively time-consuming, but allows a very comprehensive evaluation of the best BDT output cut.

In order to have an idea of the content of the sample we use for the rest of the



(a) BDTs with 13 features

(b) BDTs with 23 features

Figure 4.24: *BDT output cuts translated into efficiency and purity.* The performance obtained by all hID selections in the case of the BDT trained with 13 features (left) are shown, as well as the performances obtained with 23 features (right). Three markers indicate the optimisation done through the FOM using different penalty factors, as introduced in Equation (3.3).

analysis, we provide in Table 4.8 the decays observed on the signal side that are observed once a BDT cut of 0.5 is applied. This BDT working point is the same used in Section 4.4.5, as it yields the best measurement.

4.4 Measurement of $|V_{us}|$

The measurement of $|V_{us}|$ is performed through a ratio of branching fraction of hadronic tau decays, denoted as $R_{K/\pi}$. Using Equation (1.19), $|V_{us}|$ can be expressed as:

$$|V_{us}| = R_{K/\pi}^{1/2}\,|V_{ud}|\,\frac{f_\pi}{f_K}\,\frac{1-m_\pi^2/m_\tau^2}{1-m_K^2/m_\tau^2}\left(\frac{1}{1+\delta_{R_{\tau,K/\tau,\pi}}}\right)^{1/2} \propto R_{K/\pi}^{1/2}, \tag{4.14}$$

where

$$R_{K/\pi} = \frac{\mathcal{B}(\tau^- \to K^-\nu_\tau)}{\mathcal{B}(\tau^- \to \pi^-\nu_\tau)}. \tag{4.15}$$

Process	Events	Fraction [%]
$\tau^+\tau^-$	24381938.71	96.51
- $\pi^-\nu_\tau$	13794700.72	54.60
- $\pi^-\pi^0\nu_\tau$	7367497.90	29.16
- $\mu^-\bar{\nu}_\mu\nu_\tau$	936905.93	3.71
- $\pi^-\pi^0\pi^0\nu_\tau$	857776.81	3.40
- $K^-\nu_\tau$	595841.89	2.36
- $e^-\bar{\nu}_e\nu_\tau$	212194.94	0.84
- $\pi^-K^0_L\nu_\tau$	192669.97	0.76
- $K^-\pi^0\nu_\tau$	113253.24	0.45
- $\pi^-\pi^0\gamma\nu_\tau$	72570.02	0.29
- $\pi^-K^0_S\nu_\tau$	59984.40	0.24
- $\pi^-3\pi^0\nu_\tau$	39497.27	0.16
- $K^-K^0_L\nu_\tau$	27412.84	0.11
- $\pi^-\pi^0K^0_L\nu_\tau$	27971.77	0.11
- $K^-\pi^0K^0_L\nu_\tau$	11291.43	0.04
- $2\pi^-\pi^+\nu_\tau$	10837.97	0.04
- $\tau^- \to$ other	61531.61	0.24
$q\bar{q}$	801761.67	3.17
$e^+e^-\ell^+\ell^-$	63580.56	0.25
$\ell^+\ell^-(\gamma)$	12575.40	0.05
BB	3493.49	0.01
h^+h^-ISR	207.23	<0.01
total	25263557.07	100.00

Table 4.8: *Signal for the combined channels, after all selections.* The hID selection is using the Reweighted variable with a working point of 0.9. The BDT output working point is set to 0.5. The "*other*" category represents all other decays that are not explicitly indicated.

As highlighted in Equation (4.14), this measurement is related to the events that were selected only through the ratio $R_{K/\pi}$. The other independent factors will be used only to assess the resulting $|V_{us}|$ measurement.

The ratio $R_{K/\pi}$ is measured with a template fit method, detailed in Section 4.4.1.

The ratio is evaluated using a maximum likelihoods fit, and simultaneously evaluates the associated uncertainties.

Several sources of systematic uncertainties have been identified, and are detailed in Section 4.4.5. Due to time constraints, a choice of uncertainties had to be made. As a result, not all uncertainties could be fully implemented, but priority was given to the uncertainties expected to be major in this analysis. However, if a specific result is not provided, the approach is elaborated to explain how to obtain it and what outcome to anticipate.

4.4.1 Template Fit

Measurements typically rely on assessing the consistency of observed events with theoretical predictions. In our approach, we employ the *pyhf* [93, 94] library to perform a template fit that incorporates systematic uncertainties as nuisance parameters. It is an implementation of *HistFactory* [95], and can be used to build models with binned likelihood methods.

Likelihoods capture the probability of observing the given data under the assumed model. They are constructed to incorporate templates, which serve as theoretical predictions or reference distributions for each bin in the histogram. Binning is an important aspect of this analysis, and its choice can impact the sensitivity and precision of the measurement. The templates are defined based on the expected signal and background contributions to the data, and the fitting process adjusts the model parameters to find the best agreement between the templates and the observed data.

The general form of the fit function $f(x|\phi)$ is a product of two terms: one dedicated to the simultaneous measurement of multiple channels, and one that summarises the constraint terms for auxiliary measurements. It is expressed as:

$$f(\boldsymbol{n},\boldsymbol{a}\,|\,\boldsymbol{\eta},\boldsymbol{\chi}) = \prod_{\substack{c\in\text{channels},\\ b\in\text{bins}(c)}} \text{Pois}\left(n_b^c \,|\, \nu_b^c(\boldsymbol{\eta},\boldsymbol{\chi})\right) \quad \prod_{\chi\in\boldsymbol{\chi}} c_\chi\left(a_\chi \,|\, \chi\right), \tag{4.16}$$

where the calculated event rates of the channel c and the bin b associated to the channel c are computed as:

$$\nu_b^c(\boldsymbol{\eta},\boldsymbol{\chi}) = \sum_{s\in\text{samples}} \nu_b^{c,s}(\boldsymbol{\phi}) = \sum_{s\in\text{samples}} \left(\prod_{\kappa\in\boldsymbol{\kappa}} \kappa_b^{c,s}(\boldsymbol{\phi})\right)\left(\nu_b^{c,s,0}(\boldsymbol{\phi}) + \sum_{\Delta\in\boldsymbol{\Delta}} \Delta_b^{c,s}(\boldsymbol{\phi})\right). \tag{4.17}$$

where $f(\boldsymbol{x}\,|\,\boldsymbol{\phi})$ is the model, $\boldsymbol{n}$ is the channel data (or event counts), $\boldsymbol{a}$ is the auxiliary

data, $\boldsymbol{\phi} = \{\boldsymbol{\eta}, \boldsymbol{\chi}\}$ represents all parameters, with $\boldsymbol{\eta}$ and $\boldsymbol{\chi}$ free parameters and constrained parameters respectively, $c_\chi\left(a_\chi \mid \chi\right)$ is the constraint term for constrained parameter, ν^0 is the nominal event rates, $\nu(\boldsymbol{\phi})$ is the calculated event rates, $\boldsymbol{\kappa}(\boldsymbol{\phi})$ is a multiplicative rate modifier and $\boldsymbol{\Delta}(\boldsymbol{\phi})$ is an additive rate modifier.

An overview of the available modifiers in *pyhf* and the corresponding constraint terms is shown in Tables 4.9 and 4.10. The choice of modifier used for each systematic uncertainty hinges on the correlation between the template bins, distinguishing between uncorrelated and correlated shapes. In the context of *pyhf*, it's possible to employ the same nuisance parameter to influence multiple templates, which is critical for correlating a systematic variation across various samples or channels. Given that nuisance parameters are inherently independent by design, modifiers are utilised specifically to model effects that are either fully correlated or completely independent. However, there is a way to model arbitrary correlation across template bins using multiple correlated shape modifiers, as described in Section 4.4.3.2.

	Description	Modifier	Name
1	Uncorrelated Shape	*shapesys*	$\kappa_{scb}\left(\gamma_b\right)$
2	Correlated Shape	*histosys*	$\Delta_{scb}(\alpha)$
3	Normalisation Uncertainty	*normsys*	$\kappa_{scb}(\alpha)$
4	MC Statistical Uncertainty	*staterror*	$\kappa_{scb}\left(\gamma_b\right)$
5	Luminosity	*lumi*	$\kappa_{scb}(\lambda)$
6	Normalisation	*normfactor*	$\kappa_{scb}\left(\mu_b\right)$

Table 4.9: *Relevant modifiers and constraints (part 1).* The indices correspond to sample (s), channel (c), and related bin (b).

4.4.2 Measurement of $R_{K/\pi}$

Employing a template fit enables us to obtain both the maximum likelihood estimate and the associated uncertainty for the parameter $R_{K/\pi}$. Notably, estimating systematic uncertainties becomes significantly more straightforward compared to methods based on event counting; uncertainties can be incorporated directly as modifications to the templates and the results are then obtained through the fitting process. To gain insights into the impact of individual systematic sources considered in this study, a toy approach is employed. This method allows for the computation of the relative expected uncertainty on $R_{K/\pi}$ attributable to each systematic

	Modification	Constraint Term c_χ	Input
1	$\kappa_{scb}\left(\gamma_b\right) = \gamma_b$	$\prod_b \text{Pois}\left(r_b = \sigma_b^{-2} \mid \rho_b = \sigma_b^{-2}\gamma_b\right)$	σ_b
2	$\Delta_{scb}(\alpha) = f_p(\alpha \mid \Delta_{\alpha=-1}, \Delta_{\alpha=1})$	$\text{Gaus}(a = 0 \mid \alpha, \sigma = 1)$	$\Delta_{scb,\alpha=\pm1}$
3	$\kappa_{scb}(\alpha) = g_p(\alpha \mid \kappa_{\alpha=-1}, \kappa_{\alpha=1})$	$\text{Gaus}\ (a = 0 \mid \alpha, \sigma = 1)$	$\kappa_{scb,\alpha=\pm1}$
4	$\kappa_{scb}\left(\gamma_b\right) = \gamma_b$	$\prod_b \text{Gaus}\left(a_{\gamma_b} = 1 \mid \gamma_b, \delta_b\right)$	$\delta_b^2 = \sum_s \delta_{sb}^2$
5	$\kappa_{scb}(\lambda) = \lambda$	$\text{Gaus}\ (l = \lambda_0 \mid \lambda, \sigma_\lambda)$	$\lambda_0, \sigma_\lambda$
6	$\kappa_{scb}\left(\mu_b\right) = \mu_b$	-	-

Table 4.10: *Relevant modifiers and constraints (part 2).* The indices correspond to sample (s), channel (c), and related bin (b). Here, $\Delta_{\alpha=\pm1} \equiv \Delta_{scb,\alpha=\pm1}$ and $\kappa_{\alpha=\pm1} \equiv \kappa_{scb,\alpha=\pm1}$

source, providing a comprehensive understanding of their respective contributions. Another advantage is that the evaluation of the systematic uncertainties can be done by steps, and therefore allows a more flexible approach, as it does not require the implementation of all systematic sources.

4.4.2.1 Likelihoods

To perform a fit of $R_{K/\pi}$, we require two distinct channels: one for pions and one for kaons. These channels are defined in both data and Monte Carlo based on a cut applied to the corresponding hadron identification variable. In addition, as $R_{K/\pi}$ is expressed as:

$$R_{K/\pi} = \frac{\mathcal{B}(\tau^- \to K^- \nu_\tau)}{\mathcal{B}(\tau^- \to \pi^- \nu_\tau)} \equiv \frac{\mathcal{B}_K}{\mathcal{B}_\pi}, \tag{4.18}$$

we need to introduce

$$\kappa_{(\mathcal{B}_\pi)} = \frac{\mathcal{B}_\pi}{\mathcal{B}_\pi^{gen}} \qquad \text{and} \qquad \kappa_{(R_{K/\pi})} = \frac{R_{K/\pi}}{R_{K/\pi}^{gen}} \tag{4.19}$$

as the multiplicative rate modifiers that are needed in Equation (4.17). $\kappa_{(\mathcal{B}_\pi)}$ is multiplied to all event rates of the signal template in the pion channel and $\kappa_{(R_{K/\pi})} \cdot \kappa_{(\mathcal{B}_\pi)}$ is multiplied to all event rates of the signal template in the kaon channel, as shown in Equation (4.22). If we now normalise the calculated signal event rates ν^{sc} in the kaon and pion channel by dividing them with the corresponding branching fraction of the signal tau decay used in the generator, $\kappa_{(R_{K/\pi})}$ will be the estimate of $R_{K/\pi}$ we are interested in. Additionally, the branching fractions used in the tau

lepton pair generator are given by:

$$\mathscr{B}_K^{gen} \equiv \mathscr{B}^{gen}(\tau^- \to K^- \nu_\tau) = 0.108164 \quad \text{and} \quad \mathscr{B}_\pi^{gen} \equiv \mathscr{B}^{gen}(\tau^- \to \pi^- \nu_\tau) = 0.006964. \tag{4.20}$$

Therefore, in the case of $R_{K/\pi}$, Equation (4.16) becomes:

$$f(\boldsymbol{n}, \boldsymbol{a} \mid R_{K/\pi}, \boldsymbol{\chi}) = \prod_{\substack{c \in K,\pi \\ b \in \text{bins}(c)}} \text{Pois}\left(n_b^c \mid \nu_b^c(R_{K/\pi}, \boldsymbol{\chi})\right) \quad \prod_{\chi \in \boldsymbol{\chi}} c_\chi\left(a_\chi \mid \chi\right), \tag{4.21}$$

with the calculated event rates of the pion and kaon channel and the bin b computed as:

$$\begin{aligned} \nu_b^K(R_{K/\pi}, \boldsymbol{\chi}) = \nu_b^K(\boldsymbol{\phi}) &= \frac{R_{K/\pi}}{R_{K/\pi}^{gen}} \cdot \frac{\mathscr{B}_\pi}{\mathscr{B}_\pi^{gen}} \cdot \left(\nu_b^{K,\text{sig}}(\boldsymbol{\phi}) + \nu_b^{K,\text{bkg}}(\boldsymbol{\phi}) + \nu_b^{K,\text{fake}}(\boldsymbol{\phi})\right) \text{ and} \\ \nu_b^\pi(R_{K/\pi}, \boldsymbol{\chi}) = \nu_b^\pi(\boldsymbol{\phi}) &= \frac{\mathscr{B}_\pi}{\mathscr{B}_\pi^{gen}} \cdot \left(\nu_b^{\pi,\text{sig}}(\boldsymbol{\phi}) + \nu_b^{\pi,\text{bkg}}(\boldsymbol{\phi}) + \nu_b^{\pi,\text{fake}}(\boldsymbol{\phi})\right). \end{aligned} \tag{4.22}$$

Three templates are anticipated in the computation of the event rates of both channels, and correspond to signal (sig), background with hadron correctly (bkg) or wrongly (fake) identified. They are discussed in Section 4.4.2.3.

4.4.2.2 Binning

We define our templates based on the signal hadron momentum in the laboratory frame. This approach is motivated by the fact that hadron identification corrections and associated uncertainties are binned according to momentum, as discussed in Section 4.3.5.4. By focusing on laboratory frame momentum, we aim to minimise the influence of bins where hadron identification is poorly understood and uncertainties are significant. The bins used in the fit are detailed in Table 4.11.

To leverage the shape of the templates effectively in the fit, we subdivided each bin into two equal-sized bins, thereby incorporating more detailed information. Increasing the number of bins generally captures more information and enhances sensitivity. However, this could also introduce more nuisance parameters and could extend the optimisation duration. In channels with low statistical significance, like the kaon channel, the likelihood of encountering bins with zero events increases. Moreover, excessive binning can render the fit more susceptible to template mismodeling. Therefore, splitting into two bins was deemed a balanced compromise between advantages and disadvantages. Using a range from 0.5 GeV/c to 4.5 GeV/c, each template consists of 14 bins.

momentum		momentum	
number	p [GeV/c]	number	p [GeV/c]
1	[0.5, 1]	1	[0.5, 0.75]
		2	[0.75, 1]
2	[1, 1.5]	3	[1, 1.25]
		4	[1.25, 1.5]
3	[1.5, 2]	5	[1.5, 1.75]
		6	[1.75, 2]
4	[2, 2.5]	7	[2, 2.25]
		8	[2.25, 2.5]
5	[2.5, 3]	9	[2.5, 2.75]
		10	[2.75, 3]
6	[3, 3.5]	11	[3, 3.25]
		12	[3.25, 3.5]
7	[3.5, 4.5]	13	[3.5, 4]
		14	[4, 4.5]

Table 4.11: *Momentum binnings.* These binnings are used for both pion and kaon channels. The nominal bins (left) are divided into two equal-sized bins (right).

4.4.2.3 Templates

To measure $R_{K/\pi}$, we require one template to represent the signal component, along with one or more templates to describe the background, for each channel. Given that the majority of the background originates from tau pair events, subdividing the background templates by types of process is not conceivable. However, a meaningful distinction can be made based on the particle type of the signal hadron.

Some of the background events will contain accurately identified hadrons, specifically pions or kaons, while others will involve misidentified particles, known as fakes. For backgrounds with correctly identified hadrons, the uncertainty in hadron identification due to efficiency corrections will be fully correlated with the signal template. Conversely, for templates involving fakes, we will apply separate corrections for misidentification and uncertainties based on samples containing the respective particle type. As indicated in Table 4.8, a majority of the fakes consist of muons or even electrons misidentified as hadrons. Therefore, three templates are defined for each channel, as illustrated in Figure 4.25.

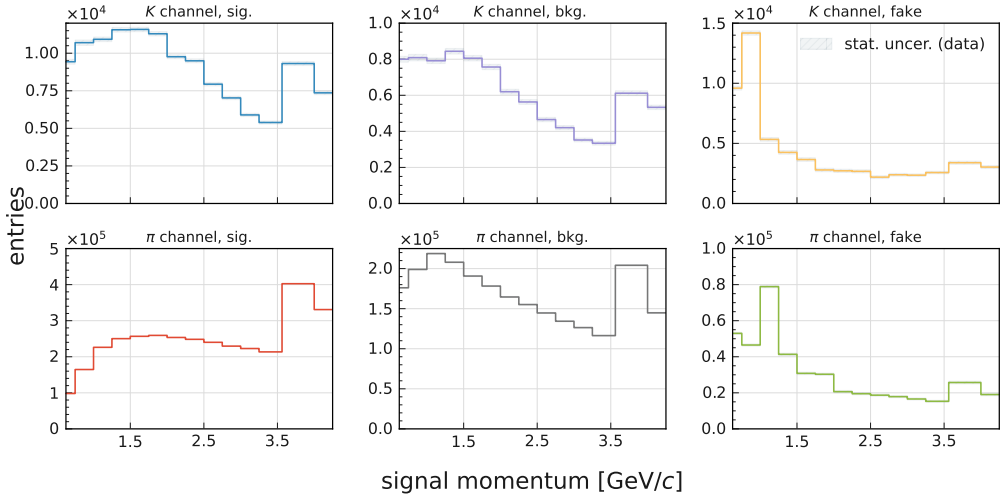


Figure 4.25: *Templates used for the fit.* The templates for the kaon channel are presented in the top row, while those for the pion channel are in the bottom row. Each row includes three templates: one representing the signal components (left), another depicting background with correctly identified hadrons (middle), and a third showing background with other particles misidentified as hadrons (right). These three templates per channel are combined and fitted simultaneously to the data. The hID selection is using the Reweighted variable with a working point of 0.9. The BDT output working point is set to 0.5.

4.4.3 Treatment of template variations

Template fit methodology encompasses various considerations to handle systematic uncertainties effectively. Modifier vetos may be implemented to control the impact of certain variations, ensuring a more robust analysis. Additionally, addressing correlated uncertainties is carefully taken into account to avoid overlooking potential interactions. Common treatments of template variations include symmetrisation or smoothening techniques, which are applied to ensure stability and mitigate fluctuations in the templates. It even extends to the nuanced handling of uncertainties through techniques such as a split of variations.

4.4.3.1 Split of variations

The templates variations can be split into into normalisation (*normsys*) and shape (*histosys*) components, both of which are fully correlated with each other. This split is mathematically equivalent to using a single *histosys* component, but it offers the advantage of potentially improving the behavior of the fit. The normalisation component $\kappa_{\text{norm}}^{\text{up/dn}}$ and the shape component $n_{b,\text{shape}}^{\text{up/dn}}$ are computed as:

$$\kappa_{\text{norm}}^{\text{up/dn}} = \sum_{b \in \text{bins}} \frac{n_b^{\text{up/dn}}}{n_b^{\text{nom}}} \quad \text{and} \quad n_{b,\text{shape}}^{\text{up/dn}} = \frac{n_b^{\text{up/dn}}}{\kappa_{\text{norm}}}. \tag{4.23}$$

where $n_b^{\text{up/dn}}$ are the up or down variations and n_b^{nom} are the nominal values of the templates in bin b.

In our approach, whenever we use a correlated shape modifier (*histosys*), we consistently split it into two components. In *pyhf*, if the nuisance parameter for these components shares the same name (denoted as α in Equation (4.21)), it is shared across the modifiers. The allocation of nuisance parameters by both *normsys* and *shapesys* components counts as one combined parameter.

4.4.3.2 Correlated uncertainties

Effectively representing the correlation is not a trivial task. This requires the combination of two mathematical expressions. First, we consider the relation between the covariance matrix C_{ik}^{sys} and a correlated shape variation Γ_{ij}, such as:

$$C_{ik}^{\text{sys}} = \sum_{j=1}^{N_s} \Gamma_{ij}\Gamma_{kj}, \tag{4.24}$$

where Γ_{ij} is a variation of bin i associated to some systematic source j and the sum is over N_s independent systematic sources. Secondly, we use the fact that any arbitrary covariance matrix C^{sys} can decomposed into $N_s = \dim(C^{\text{sys}})$ shape variations via eigendecomposition, as follows:

$$C^{\text{sys}} = VUV^T = (V\sqrt{U})(V\sqrt{U})^T, \tag{4.25}$$

where V is a matrix with columns corresponding to eigenvectors, and U is a diagonal matrix with the corresponding eigenvalues. Combining Equations (4.24) and (4.25) yields the following expression:

$$\Gamma = V\sqrt{U}. \tag{4.26}$$

Each eigenvector, multiplied by the square root of the corresponding eigenvalue, gives a correlated shape variation that we can implement with *histosys* modifiers. Collectively, these distinct shape variations encapsulate the overall correlation. In certain instances, eigenvalues may approach zero or be exactly zero. This situation can occur with highly correlated uncertainties, where only one predominant

shape variation is evident. To avoid incorporating unnecessary parameters into the fitting process, one can establish a threshold for the eigenvalues below which the corresponding variation is disregarded. This technical detail is discussed in the next section.

4.4.3.3 Modifier veto

To avoid potential fit instabilities, we we want to identify parameters that are creating minor template modifications, or having a negligible impact on the likelihood, in order to remove them. To handle these situations, we define certain conditions under which systematic uncertainties are excluded from the model. If a modifier affects multiple samples, it must be removed from all templates to be excluded.

For correlated shape variations (*histosys*), both of the following conditions must be satisfied:

- ⋆ emphEither the p-value of a χ^2 test comparing the shape component to the nominal histogram exceeds 0.99, *or* the largest deviation of the shape component from the nominal histogram is less than 10% of the smallest expected statistical uncertainty in the data, defined as:

$$\sigma^{\text{stat}}_{\text{exp},cb} = \sqrt{\sum_{s \in \text{samples}} \nu^0_{scb}}.$$

  where ν^0_{scb} represents the nominal event rates in Monte Carlo. The smallest expected statistical uncertainty in the data is always determined from all bins within the corresponding channel.

- ⋆ The deviation of the normalisation component from unity is less than 10% of the smallest relative expected statistical uncertainty in the data.

For uncorrelated shape variations (*shapesys*), the veto is applied to individual bins if the uncertainty in a bin is less than 10% of the expected statistical uncertainty in the data for that specific bin.

In the figures shown in Section 4.4.5, the relative expected statistical uncertainty in the data will be represented by hatched areas corresponding to template variations. The legend will display the smallest relative expected statistical uncertainty for each template.

For correlated shape variations, the figures will also include separate displays of p-values from χ^2 tests and normalisation components (minus one) for both up and

down variations. Additionally, each plot will indicate whether individual variations pass (green) or fail (red) veto conditions. If all samples are vetoed, the template modification will not be included in the model, as indicated by a checkbox on the plots.

For uncorrelated shape variations, the legend will reflect the largest relative variation observed for each template. Since the veto is applied bin-wise in this case, the color of individual variations (red or green) will indicate whether the modification is included.

4.4.3.4 Symmetrisation

All template variations with separate up and down components are made symmetrical before integration into the model. Symmetrisation is achieved by maintaining a constant difference between the up and down variations, adjusting only to ensure that their deviation from the nominal template is consistent. This process typically has minimal effect on most scenarios covered. In cases where only the up variation is provided, the down variation is always included as the symmetric counterpart of the up variation.

4.4.3.5 Smoothening

A smoother fits a curve through data, separating it into a smooth signal and rough residuals, where the rough component includes additive noise and outliers; methods like moving averages or running medians address this task but differ in their susceptibility to outliers and preservation of level shifts. In all generality, data from an equidistant set $\{x_1, x_2, ...x_n\}$ are smoothed by replacing each value x_i with a calculated smoothed value y_i derived a function f taking a window around x_i of width defined by the minimum and maximum elements $x_{i\pm k}$, such that:

$$y_i = f(x_{i-k}, x_{i-k+1}, ..., x_i, ..., x_{i+k-1}, x_{i+k})$$

The methods available are typically identified by a sequence of alphanumeric characters, which correspond to operations detailed in [96]. The meanings of these operations are provided in Table 4.12.

In particular, the *353QH* algorithm [97] is applied specifically to the relative variations involving the nominal histogram to address fluctuations within the variation itself. It is applied twice on histogram variations before incorporating them into the model. This smoothening process is used to mitigate statistical fluctuations, espe-

Operation	Description
2, 3, 4, 5, ...	Window size
R	Median smooth: Follow a median smooth, if no changes occur
H	Hanning: Convolution with a symmetrical kernel such that $y_i = \frac{1}{4}x_{i-1} + \frac{1}{2}x_i + \frac{1}{4}x_{i+1}$, with end values unchanged
G	Conditional hanning: Only non-monotone sequences of length 3
S	Split: Dissect the sequence into shorter subsequences at all places where two successive calues are identical, apply a $3R$ smooth to each sequence, reassemble and apply a 3 smooth
Q	Quadratic interpolation: Follow a quadratic interpolation of flat 3-areas

Table 4.12: *Main smoothening operations.* Possible operations extracted from algorithm names. The operations are detailed in [96].

cially when the up and down variations of the histogram result from independent reprocessing of Monte Carlo or data samples with parameter adjustments.

4.4.4 Statistical Uncertainties

Unlike systematic uncertainties that must be explicitly incorporated into the model, statistical uncertainty is already included in template fitting using *pyhf*. Consequently, any measurement inherently encompasses this uncertainty. When conducting the fit on Asimov data, the statistical uncertainty is calculated to align with the expected data luminosity.

4.4.5 Systematic Uncertainties

Systematic uncertainties are included as modifications of the templates through nuisance parameters. A toy approach is used to compute the relative expected uncertainty on $R_{K/\pi}$ that can be attributed to each of the sources that are described in the next sections. In particular, Table 4.13 shows the treatment applied on each systematic uncertainty source. If not explicitly stated otherwise, all uncertainties put in bold are implemented. They correspond mostly to uncertainties expected to be dominant, and were complemented with uncertainties that could be added rela-

tively quickly. Asimov data has be used to assess the uncertainties later propagated on $R_{K/\pi}$. Real data has only been used to assess up and down variations, that are then propagated on Monte Carlo samples.

Uncertainty source	Mod.	Split	Corr.	Veto	Smooth.	Symm.
Luminosity	*lumi*	✗	✗	✗	✗	✗
Charged PID	*hsys*	✓	✓	✓	✗	✗
Trigger	*hsys*	✓	✗	✓	✗	✗
Misalignment	*ssys*	✗	✗	✓	✗	✗
Tag side MC modeling	*ssys*	✗	✗	✓	✓	✗
BDT efficiency	*hsys*	✓	✗	✓	✓	✗
Bkg normalisation	*nsys*	✗	✗	✗	✗	✗
MC stat	*stat*	✗	✗	✗	✗	✗
Signal side p modeling	*hsys*	✓	✗	✓	✓	✗
Tracking p scale factor	*hsys*	✓	✗	✓	✓	✓
Tracking efficiency	*nsys*	✗	✗	✗	✗	✗
π^0 efficiency	*hsys*	✓	✗	✓	✓	✓
γ efficiency corr.	*hsys*	✓	✗	✓	✓	✓
γ energy bias	*hsys*	✓	✗	✓	✓	✓

Table 4.13: *Treatments applied to the different systematic sources.* The corresponding modifiers are also indicated, and are written in a succinct form, such that *hsys* corresponds to *histosys*, *ssys* to *shapesys*, *nsys* to *normsys*, and *stat* to *staterror*. The X mark is deliberately smaller to make the table easier to read. If not explicitly stated otherwise, all uncertainties put in bold are implemented.

4.4.5.1 Luminosity

The luminosity value corresponding to the experiment number going from 10 to 26 is given by:

$$\mathscr{L}_{\text{int}} = 361.922 \pm 0.021\ (\text{stat.}) \pm 2.331\ (\text{syst.})\ \text{fb}^{-1}.$$

This yields a total relative uncertainty of 0.6%. This uncertainty is implemented in the model using the *lumi* modifier, affecting all templates uniformly. Despite this modifier acting as a global scaling factor, one might anticipate minimal impact on $R_{K/\pi}$. However, unlike the expected full cancellation seen in simple ratio calculations with background-subtracted yields, a small effect on $R_{K/\pi}$ can still occur

during fitting. This is because a global scaling of the data can influence the determination of correction factors for other sources of mismodeling, such as $q\bar{q}$ normalisation or π^0 efficiency, which may vary differently across templates and thereby affect $R_{K/\pi}$.

4.4.5.2 Charged PID

Systematic uncertainties resulting from the use of hID corrections have to be assessed on Monte Carlo templates. We adopt the method outlined in Section 4.4.3.2 to integrate correlated uncertainties into our model. However, for effective implementation, understanding the correlation of the hID uncertainties is essential. The statistical component of the hID corrections is inherently independent across all correction bins. Currently, the correlation of the systematic component remains unknown, necessitating conservative assumptions to ensure a model capable of accurately describing the data. The corrections are binned across three dimensions, which correspond to:

- ⋆ Charge: Combining positive and negative charged tracks in our measurement while assuming fully correlated uncertainties will result in the largest overall uncertainties.
- ⋆ Polar angle: Similarly, when integrating over polar angle bins, assuming fully correlated uncertainties for this dimension will result in the largest uncertainties.
- ⋆ Momentum: Performing the fit with templates binned in momentum makes it highly sensitive to the correlation of hadron identification uncertainties. A priori, it is unclear which correlation assumptions are the most conservative. However, after conducting toy studies to explore different correlation assumptions, it seems that assuming uncertainties as fully correlated is the optimal choice given our current model.

We obtain the covariance matrices that we need as input to our model with a toy approach. We create 2000 sets of correction factors using k-dimensional multivariate Gaussian distributions, ensuring that the correlations between different dimensions are represented in the generated variations. We describe the statistical component of the corrections with a random vector $\mathbf{X}$, such as:

$$\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}^{\text{stat.}}), \tag{4.27}$$

where $\mathbf{\Sigma}^{\text{stat}}$ is a diagonal matrix containing the corresponding variances and $\boldsymbol{\mu}$ is the central value of the individual weights. For the systematic component, we use a random vector $\mathbf{Y}$, such that:

$$\mathbf{Y} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}^{\text{syst.}}), \tag{4.28}$$

where $\mathbf{\Sigma}^{\text{syst.}}$ contains all the information about the assumed correlations. In the most general case, it can be expressed as:

$$\mathbf{\Sigma}^{\text{syst.}} = \begin{bmatrix} \sigma_{x_1} & & 0 \\ & \sigma_{x_2} & \\ & & \ddots \\ 0 & & \sigma_{x_k} \end{bmatrix} \begin{bmatrix} 1 & \rho_{x_1,x_2} & \cdots & \rho_{x_1,x_k} \\ \rho_{x_2,x_1} & 1 & \cdots & \rho_{x_2,x_k} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{x_k,x_1} & \rho_{x_k,x_2} & \cdots & 1 \end{bmatrix} \begin{bmatrix} \sigma_{x_1} & & 0 \\ & \sigma_{x_2} & \\ & & \ddots \\ 0 & & \sigma_{x_k} \end{bmatrix}, \tag{4.29}$$

where σ_{x_i} are the systematic uncertainties corresponding to each bin. If the uncertainty is asymmetric, we consider the larger magnitude between the up and down components as σ_{x_i}. The correlations between uncertainties are denoted as ρ_{x_i,x_j}. Under the mentioned assumptions, this leads to $\rho_{x_i,x_j} = 1$. The covariance matrices encompass all momentum, theta and charge bins associated with a specific type of hadron ID correction. The correction types used in this analysis are categorised as outlined in Table 4.6, and they are treated as completely independent from one another. We obtain a k-dimensional random weight vector $\mathbf{w}$ at the end, such that:

$$\mathbf{w} = \mathbf{X} + \mathbf{Y}, \tag{4.30}$$

and from which we generate 2000 toy weights. We perform this process separately for each type of correction, applying corresponding toy weights to our Monte Carlo sample. This allows us to generate 2000 sets of templates binned according to signal momentum, following the binning scheme outlined in Section 4.4.2.2. Using the set of toy templates, we can compute the sample covariance between bin j and bin k as follows:

$$C_{jk} = \frac{1}{N-1} \sum_{i=1}^{N} \left(x_{ij} - \bar{x}_j\right)\left(x_{ik} - \bar{x}_k\right), \tag{4.31}$$

where x_{ij} denotes the i^{th} toy value of the j^{th} bin, $\bar{x}_j$ is the sample mean of the corresponding bin, and the sum is over all N toys. Therefore, from the six templates,

we derive four covariance matrices that characterise the efficiencies and misidentification rates listed in Table 4.14, which follows the same distinctive parameters mentioned already in Table 4.6.

Correction(s)	Templates	Channel	Description
πID efficiency	2	pion	sig + bkg with true π
KID efficiency	2	kaon	sig + bkg with true K
$e \to \pi$ misID $\mu \to \pi$ misID $K \to \pi$ misID	1	pion	bkg with fake π
$e \to K$ misID $\mu \to K$ misID $\pi \to K$ misID	1	kaon	bkg with fake K

Table 4.14: *Corrections used for each template.* A description is provided with the different categories, as well as the channel.

To illustrate the toy study, the total number of events in the signal template for the pion and kaon channels are shown in Figure 4.26.

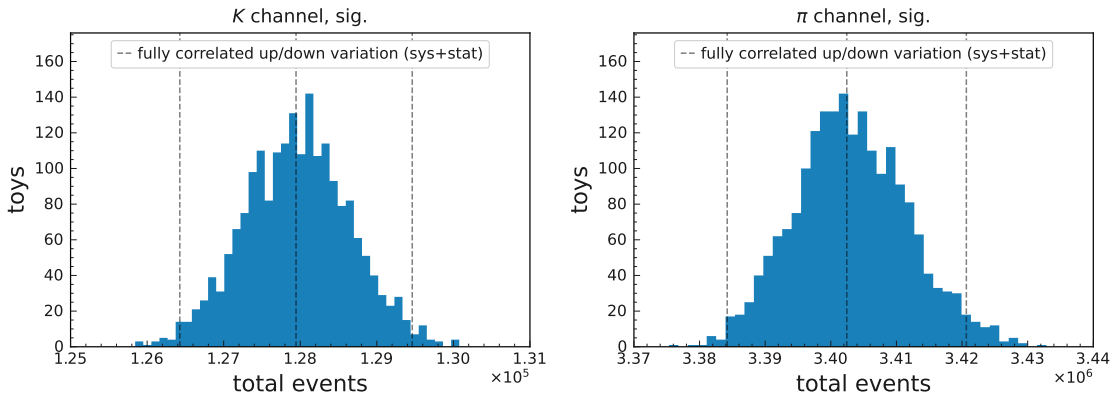


Figure 4.26: *Total number of events in the signal templates obtained with different generated toy weights.* Dashed lines show the values obtained when varying the weights up/down by their total uncertainty. The hID selection is using the Reweighted variable with a working point of 0.9. The BDT output working point is set to 0.5.

The vertical dashed lines represent values obtained by adjusting weights up and down based on their uncertainties, combining statistical and systematic components in quadrature. This approach corresponds to a fully correlated variation,

including for the statistical component. The uncertainties in both efficiency corrections are primarily driven by the systematic component. Therefore, in Figure 4.26, the fully correlated variations fall within the distribution observed from toy simulations.

The corresponding correlation matrices are shown in Figures 4.27 and 4.28 for the K and π efficiencies, and K and π fakes respectively.

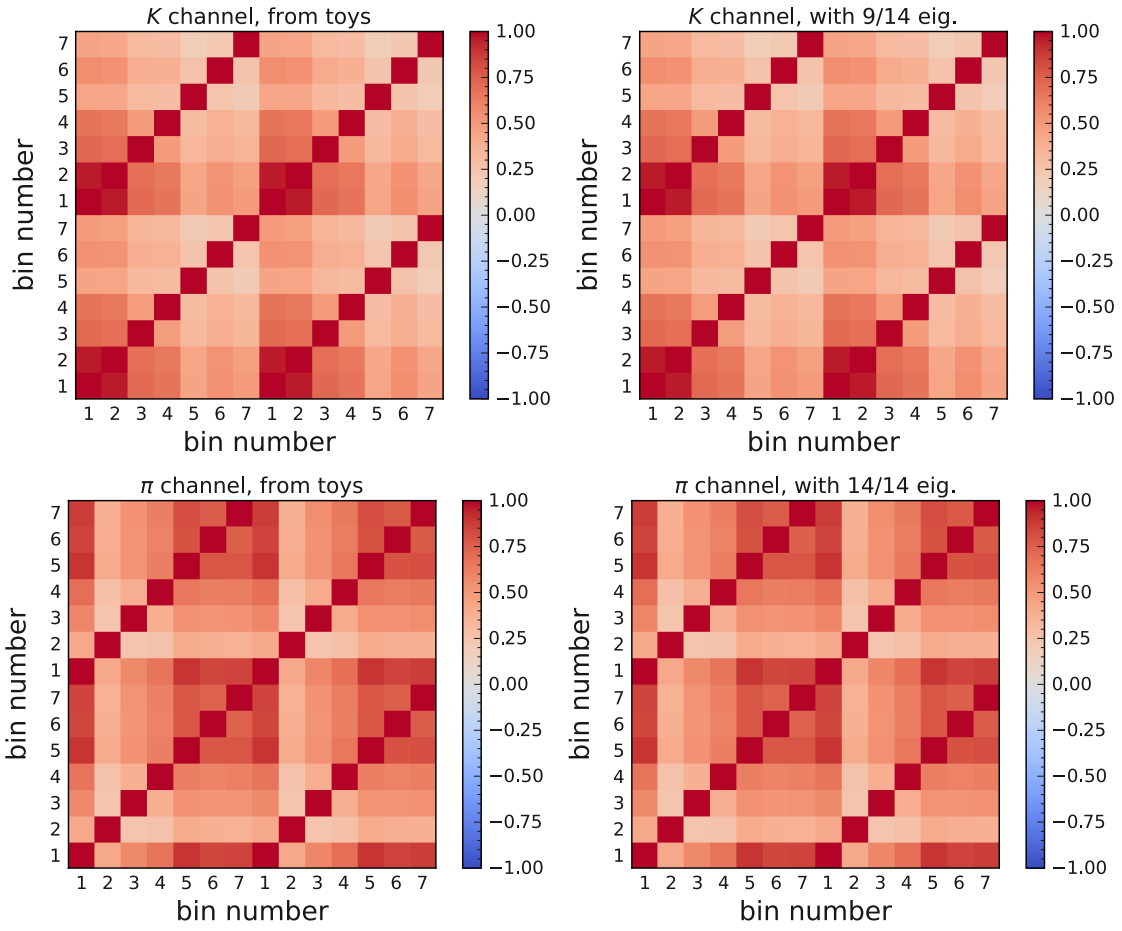


Figure 4.27: *Correlation matrices for K and π efficiencies, obtained from toys.* On the right-hand side, the correlation is shown when using eigenvalues. The nominal momentum bins are used. The hID selection is using the Reweighted variable with a working point of 0.9. The BDT output working point is set to 0.5.

To reduce the number of required parameters, we can disregard shape variations corresponding to small eigenvalues, thereby decreasing the parameter count. As a verification step, we reconstruct the covariance matrix using the remaining shape variations according to Equation (4.24) and compare it to the original matrix. The rebuilt correlation is displayed alongside the eigendecomposed matrices, revealing no significant differences compared to the original.

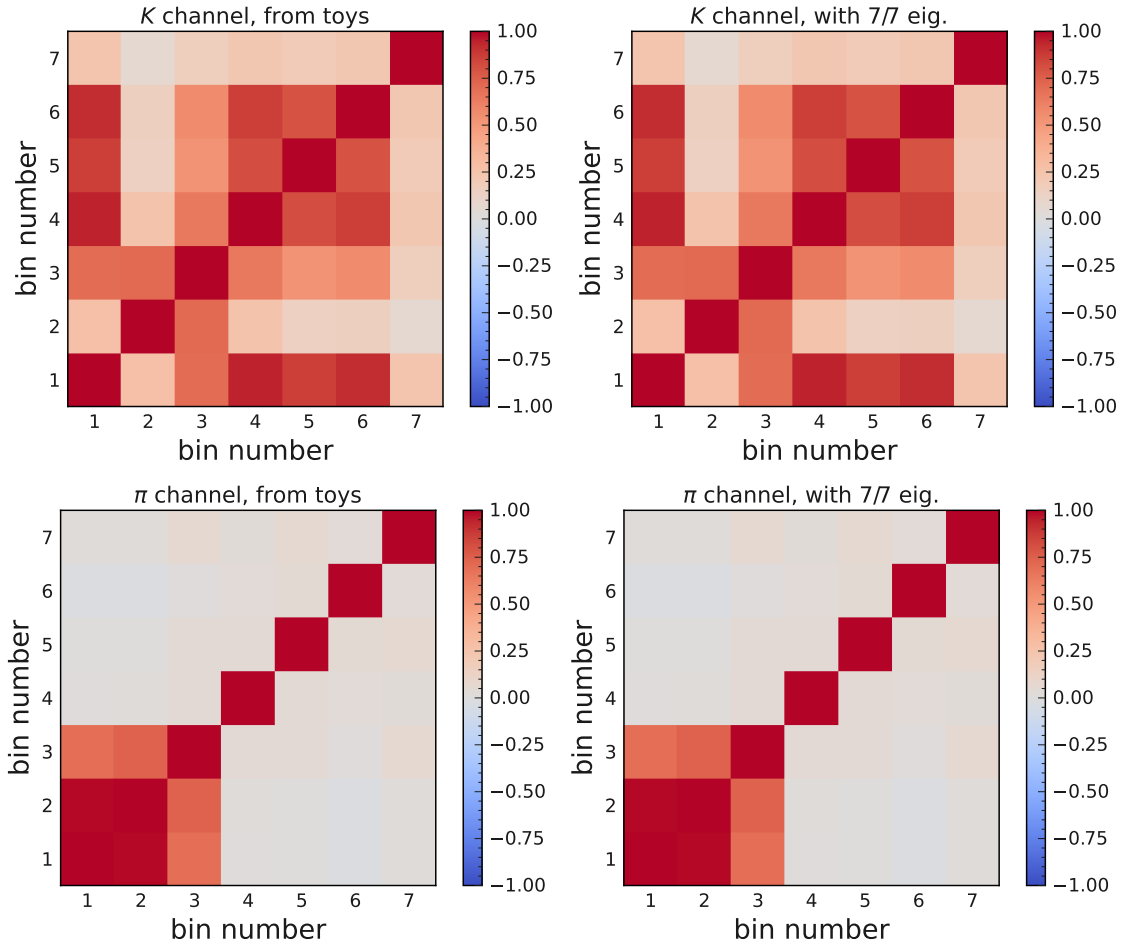


Figure 4.28: *Correlation matrix for K and π fakes, obtained from toys.* On the right-hand side, the correlation is shown when using eigenvalues. The hID selection is using the Reweighted variable with a working point of 0.9. The BDT output working point is set to 0.5.

The resulting relative variations of individual templates are illustrated in Figure 4.29. Each color represents an independent *histosys* variation, collectively modeling the desired correlation. Notably, the individual shape variations for signal and background templates of the same hadron type are fully correlated within each channel, as these shape variations are identical for both templates.

One key benefit of template fits is their ability to generate results even when not all uncertainties are accounted for. Exploiting this advantage, we conduct a comparison of various hID selections, as described in Section 4.3.6. To do so, the statistical uncertainty, as well as the luminosity uncertainties are also implemented.

The propagation of the resulting uncertainty on the ratio $R_{K/\pi}$ is shown in Figure 4.30. The hID selection criteria are provided at the end of each section of the

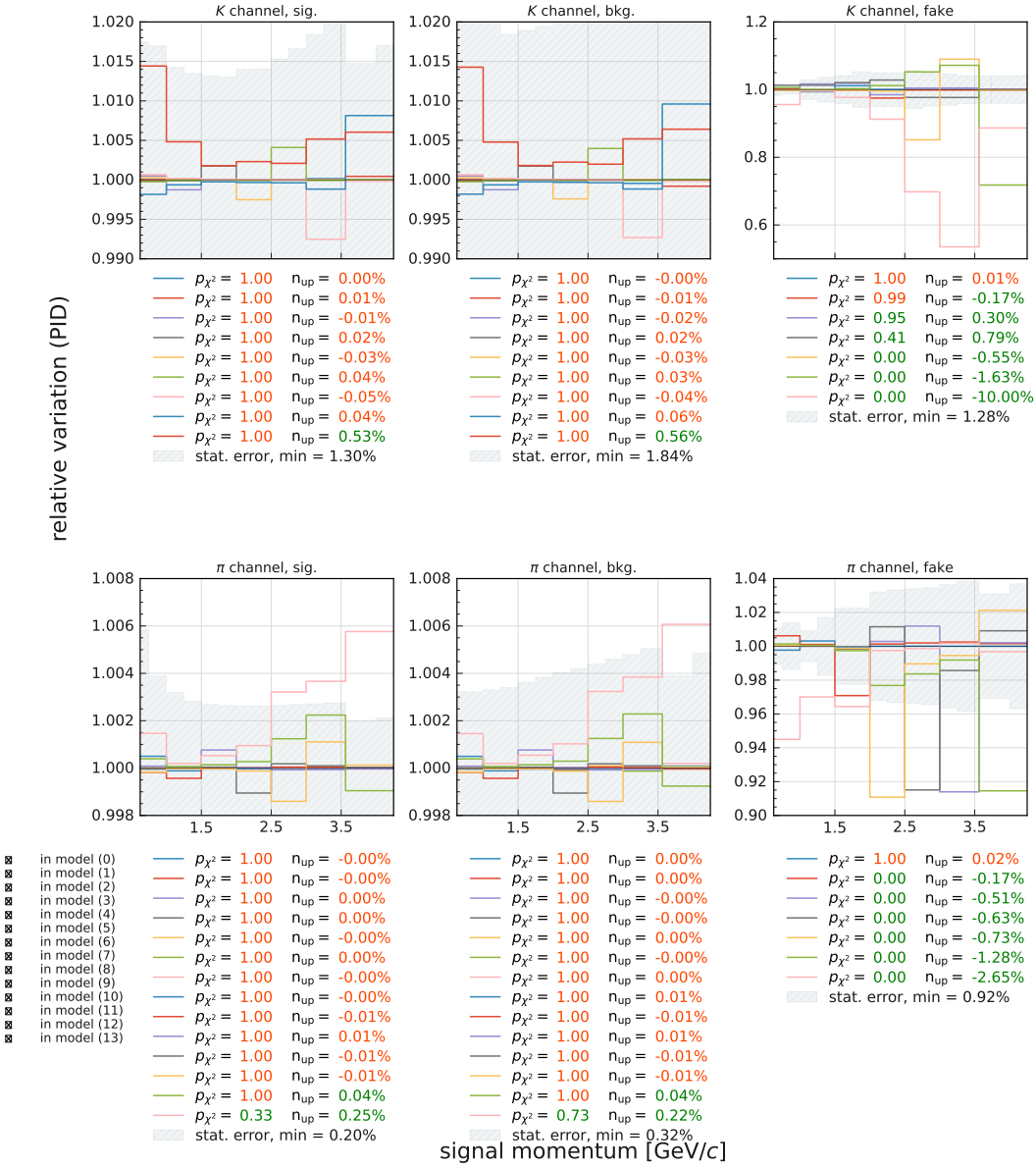


Figure 4.29: *Relative shape variations of the templates due to hID uncertainties.* Each template is affected by independent *histosys* modifiers, shown in different colours. These shape variations model together the desired correlation between the template bins. The *h*ID selection is using the Reweighted variable with a working point of 0.9. The BDT output working point is set to 0.5.

plot, with the variations obtained within this selection corresponding to different BDT output working points. A pattern emerges indicating that the propagated uncertainty decreases as the *h*ID selection becomes tighter, for both types of variables. The Reweighted *h*ID was selected as it appears to yield the minimum value, indicated in red. Other related results are available in Appendix B.2, where the different uncertainties are detailed.

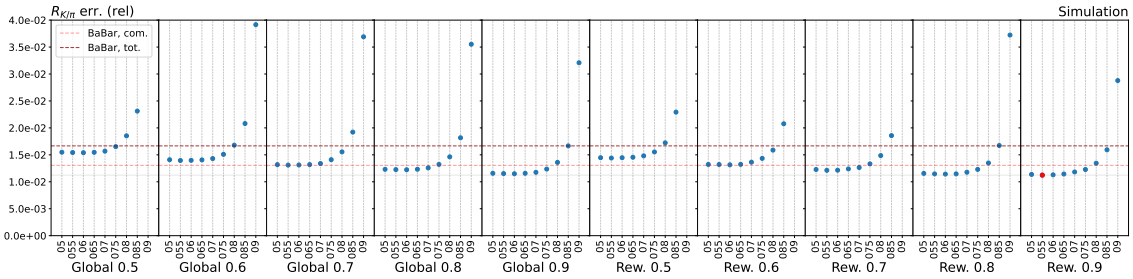


Figure 4.30: *Propagation of uncertainties on $R_{K/\pi}$ using different hID selections.* Each hID selection is decomposed into several BDT output working points. Two limits provided by the BaBar collaboration [33] are shown: a comparable limit that takes into account the hID and statistical uncertainties, and their total uncertainties. The red dot corresponds to the minimum value achieved across all selections. A grey line indicates the position of the minimum value.

4.4.5.3 Trigger

To address discrepancies between trigger simulation and data, we apply a correction to each bin individually of the Monte Carlo templates. This correction includes an associated uncertainty, which serves as a systematic uncertainty for our measurement. Details about the triggers bits are provided in Section 4.3.4.

Figure 4.31 displays the trigger efficiency comparison between data and Monte Carlo for both kaon and pion channels. In data, we only have access to the trigger inefficiency coming from Equation (4.12). However, in Monte Carlo simulations, we can also compute the absolute efficiency, as in Equation (4.3.5.3). The difference in efficiency observed in Monte Carlo between the two methods represents a systematic uncertainty.

The ratio of trigger inefficiency between data and Monte Carlo is used as a correction factor for adjusting the Monte Carlo templates. This correction factor is illustrated in the subplot of Figure 4.31, with error bars indicating the statistical uncertainty derived solely from the sample sizes of the data and Monte Carlo simulations.

A total of 42 *histosys* modifiers are needed to describe these variations. They are implemented such that they are independent across momentum bins and channels, but fully correlated for the individual momentum bins across the three samples in each channel. The relative uncertainty assigned to each bin is shown in Figure 4.32.

The reason for using the inefficiency computation in the first place becomes understandable when looking at the resulting small systematic uncertainty. Tests us-

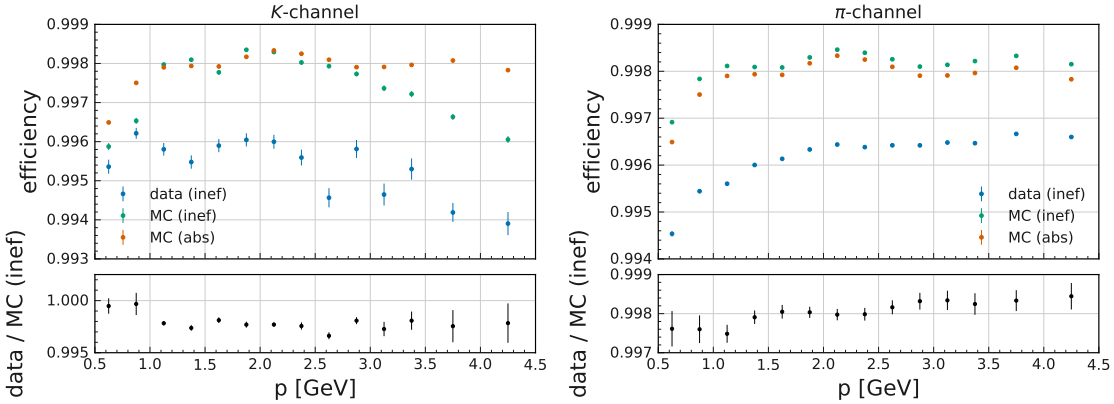


Figure 4.31: *The trigger (in)efficiency in data and Monte Carlo for both channels.* For Monte Carlo, both inefficiency and absolute efficiency are shown. The ratio of data over Monte Carlo of the trigger inefficiency, shown in the bottom part of the plot, is used to correct the templates. The error bars on the ratio include the statistical uncertainties. The hID selection is using the Reweighted variable with a working point of 0.9. The BDT output working point is set to 0.5.

ing the typical trigger efficiency computation, such as outlined in Equation (4.11), would increase this uncertainty significantly.

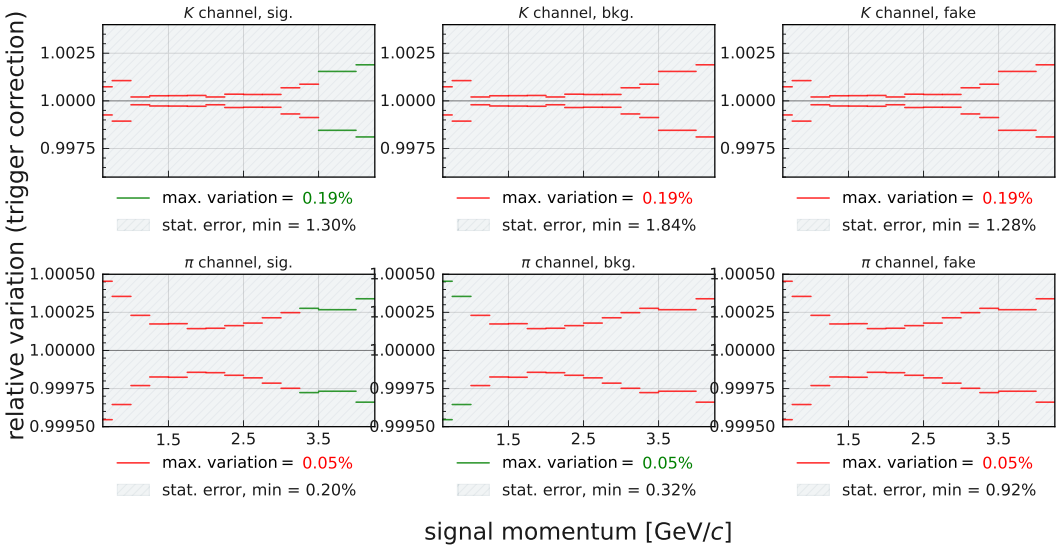


Figure 4.32: *Relative uncertainty associated with the trigger correction.* The uncertainty is assigned independently to each template bin. The hID selection is using the Reweighted variable with a working point of 0.9. The BDT output working point is set to 0.5.

4.4.5.4 Detector Misalignment

The alignment of the Belle II tracking system involves optimizing 60'000 parameters, ranging from local sensor and wire alignments to the relative global alignment of sub-detectors. These alignment parameters are determined during the calibration process using cosmic ray tracks and various data sample topologies. More information on the alignment procedures can be found in [98, 99].

The systematic distortions of the detector, known as weak modes, can adversely affect physical observables. To assess the impact of residual misalignment on the determination of $R_{K/\pi}$, we generated Monte Carlo samples of tau lepton pair decays under conditions matching the official Belle II 15^{th} Monte Carlo campaign of run independent samples, using both nominal and varied alignment configurations. These configurations can be described as the following:

- ⋆ *Configuration 1*: Residual misalignment in the iteration of the full global alignment (from zeros), including CDC wires.
- ⋆ *Configuration 2*: Difference between the prompt alignment and reprocessing alignment. This is expected to be a realistic misalignment estimation for the prompt calibration and a safe upper limit for prompt reprocessing.
- ⋆ *Configuration 3*: Supplementary alignment with all parameters for VXD sensors scaled by factor 0.3.
- ⋆ *Configuration 4*: Supplementary alignment with all parameters for VXD sensors scaled by factor 0.5.
- ⋆ *Configuration 5*: Difference from data in the day-to-day alignment of VXD and CDC detectors.
- ⋆ *Configuration 6*: Residual misalignment after VXD and CDC detectors alignment.

In each alignment configuration, the recalibration of the beam spot was incorporated. Using the same set of events from generation and simulation across all samples, we applied specific configurations to emulate misalignment during the reconstruction step.

Figure 4.33 shows the distribution of signal momentum in bins of 1 GeV/c width with the different alignment configurations. The ratio plot shows the ratio of the special alignment configurations to the nominal alignment configuration.

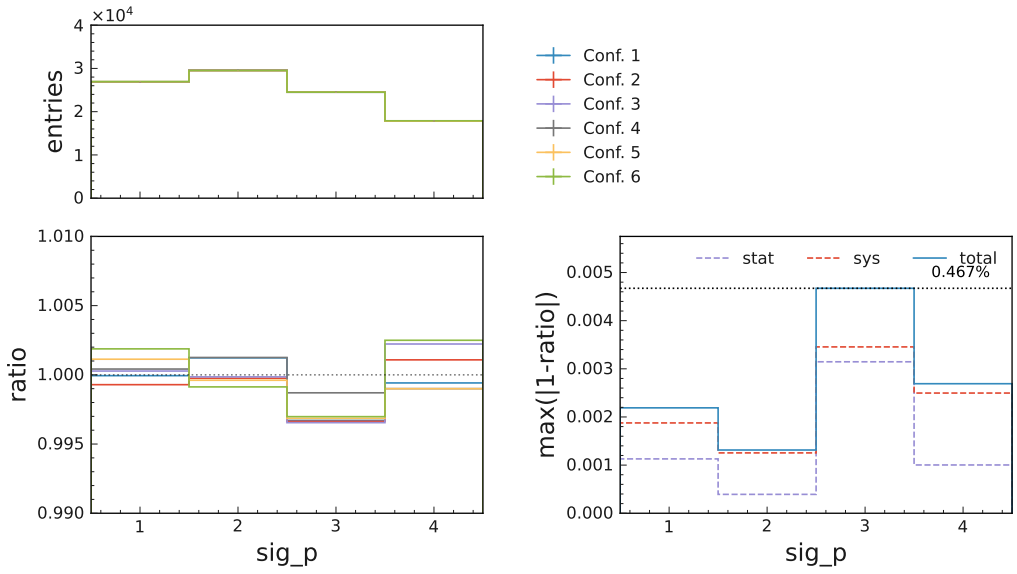


Figure 4.33: *Estimation of the misalignment systematic with different alignment configurations.* The distribution of signal momentum with the different alignment configurations is shown in the upper plot. The ratio of the special alignment configurations to the nominal alignment configuration is shown in the lower left shows. The resulting statistical and systematic uncertainties are shown in the lower right plot. They are computed based on the deviation of the ratio from 1. The hID selection is using the Reweighted variable with a working point of 0.9. The BDT output working point is set to 0.5.

Following recommendations regarding alignment uncertainties, we use the variation from *Configuration 5* to estimate the statistical component of the uncertainty, and the largest bin-wise variation of *Configuration 1*, *2*, *3*, *4* and *6* to estimate the systematic component. We then sum both components in quadrature, and obtain the result shown on the right side of Figure 4.33. We take the largest variation of the four bins and apply it as an uncorrelated shape variation to each bin of the templates. The uncertainty resulting from this conservative assumption amounts to 0.467%, and is implemented using a *shapesys* modifier.

4.4.5.5 Modeling of Monte Carlo on the tag side

To assess the influence of Monte Carlo modeling on the tag side, we use a signal embedding method. This involves taking the combined sample of K and π events and embedding truth-matched signal events (either $\tau^- \to K^- \nu_\tau$ or $\tau^- \to \pi^- \nu_\tau$) into both data and Monte Carlo, as illustrated in Figure 4.34.

The embedding process entails randomly selecting a true signal event and replacing

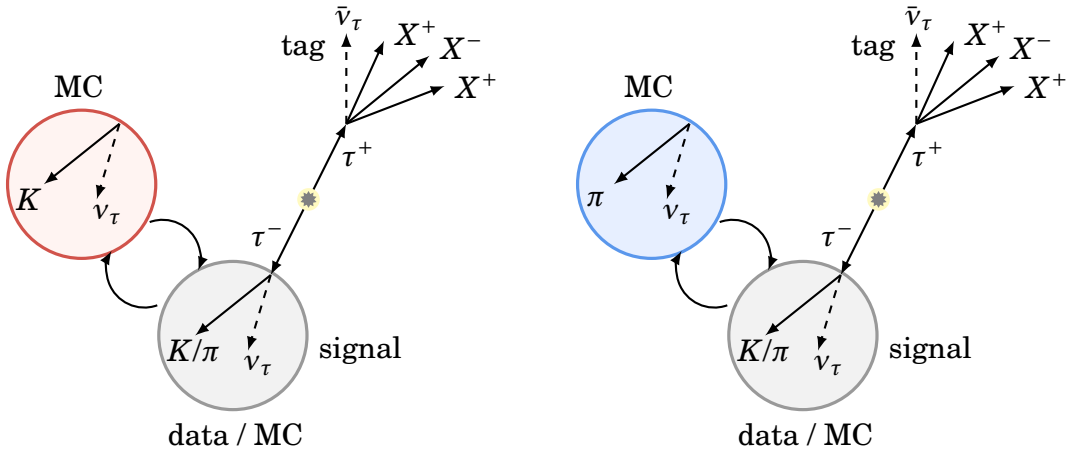


Figure 4.34: *Illustration of the embedding procedure.* As two signals are embedded in both Monte Carlo and data, four different samples result from this operation.

the corresponding signal track in the target sample that is being embedded. This technique allows for a direct comparison of modeled signals within the Monte Carlo simulation against real data.

We perform the replacement of signal tracks within quintiles of the signal momentum in the CMS frame, ensuring that the momentum magnitude of the new track is reasonably similar to the original. The replaced signal track is then rotated to align its momentum direction with that of the original track in the CMS frame. The sample on which the exchange of track is operated has a minimal selection applied, and in particular, has a loose momentum, polar angle and BDT selection. The loose momentum and polar angle selections correspond to $p_{\tau,\text{sig}} \in [0.1, 5.6]$ and $\cos\theta_{\tau,\text{sig}} \in [-0.6, 0.95]$.

Subsequently, variables that depend on signal momentum, especially event kinematics, are updated to reflect the characteristics of the replaced signal. As this procedure is done once all properties are reconstructed, it is not straightforward to update all event kinematics. In particular, the thrust, CLEO cones CC_i and Fox-Wolfram normalised moments R_i are not currently updated.

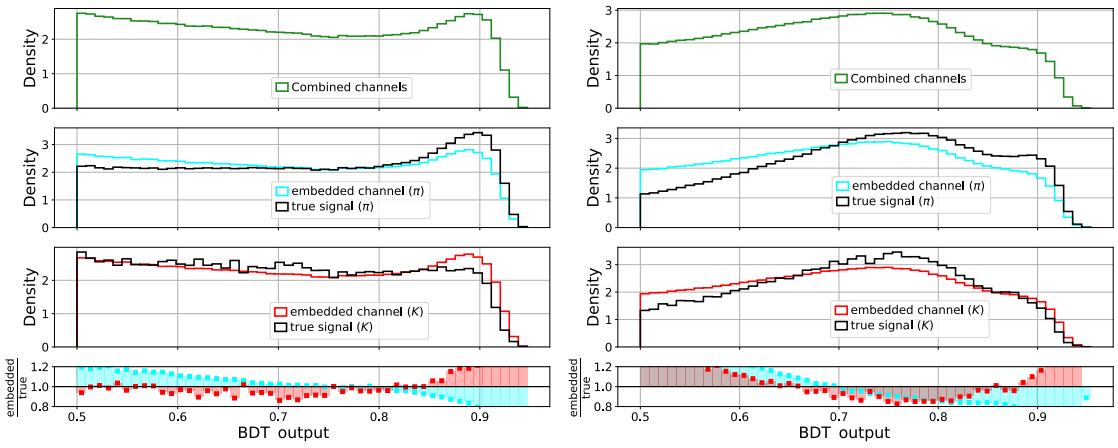
Following this, we apply again all analysis selections computed on the updated variables, and thus obtain four embedded samples.

Before comparing data and Monte Carlo embedded samples, we check how well the embedding works by comparing the distributions of the embedded Monte Carlo samples to the distributions of the true signal events used for the embedding. To gain meaningful insights without excessive computational resources, only a reduced sample of the tau lepton pair sample is used currently. This reduced sample corresponds to approximately to 10% of the whole sample, and events are selected

randomly. Several attempts were made to obtain a satisfactory agreement between the two samples. In these attempts, the following parameters were tuned:

- ⋆ Quintiles: the number of quintiles was set nominally to 6, but 2 and 10 were also tested.
- ⋆ BDT selection: as the BDT selection takes place before and after the embedding procedure with nominal and updated features respectively, it has a significant influence on the result. The number of features was set nominally to 23, but a BDT selection using 13 features appeared to be preferable. The BDT cuts were nominally set at 0.4 and 0.5. However, alternative settings of 0.75 and 0.8 were also evaluated during the testing process.

Given the multitude of features used in the selection, we decide to focus on the BDT output first, as it provides a direct indicator of any mismodeled features entering it as input. The nominal setting is displayed in Figure 4.35 along the configuration where only the number of features entering the BDT selection is changed to 13. The configuration using more BDT features is showing a greater discrepancy, as it is using more event kinematics features that are not necessarily updated in the embedding process.



(a) BDTs with 13 features

(b) BDTs with 23 features

Figure 4.35: *Distributions of combined, embedded and true signal samples of Monte Carlo.* The same configuration in both cases is employed, except the BDT selection, which uses either 13 or 23 features. The hID selection is using the Reweighted variable with a working point of 0.9.

As these numerous checks provided a result similar to that displayed in Figure 4.35, this systematic uncertainty could not be implemented in the model.

As next steps, we would compare the data and Monte Carlo agreement within the pion and kaon sample using the embedded samples. This can be done safely as the same signal sample is used for embedding data and Monte Carlo, any discrepancies we might observe come from the tag side and will cancel when we measure $R_{K/\pi}$. To quantify the difference between the discrepancies in the two samples, we would use the signal momentum distribution with the same binning as for the template fit, with the whole selection applied. This uncertainty would be implemented with uncorrelated shape (*shapesys*) modifiers.

4.4.5.6 BDT efficiency

The embedded samples explained in Section 4.4.5.5 would also be employed to derive a variation that is fully correlated across all templates, reflecting discrepancies in the ratio of data over Monte Carlo. While many of these discrepancies are likely accounted for by other systematic uncertainties or are simply statistical fluctuations, we avoid treating the ratio as a shape variation. Instead, we aim to mitigate most of these effects by calculating the BDT efficiency, expressed as:

$$\varepsilon_{\text{BDT}} = \frac{N_{\text{pass}}}{N_{\text{pass}} + N_{\text{fail}}}, \tag{4.32}$$

where N_{pass} is the number of events passing BDT selection and N_{fail} is the number of events removed by the cut on the BDT output. This efficiency would be calculated separately for each momentum bin within both the π and K samples. The ratio of the BDT efficiency in data to the BDT efficiency in Monte Carlo would be used to quantify the change in the ratio of data over Monte Carlo before and after applying the BDT cut.

4.4.5.7 Background normalisation

We intend to address systematic effects affecting the overall normalisation in all bins by introducing a normalisation uncertainty (*normsys*) for each physics process. To estimate a normalisation uncertainty for the tau lepton pair process, we would use a control sample. To account for potential normalisation differences arising from non-tau lepton pair processes, we would conduct a simple two-component fit using templates corresponding to tau lepton pair and non-tau lepton pair events. We would assign a normalisation uncertainty (*normsys*) to the non-tau lepton pair template and an uncorrelated shape uncertainty (*shapesys*) to the tau lepton pair template. Additionally, we would assign an overall normalisation uncertainty to

the non-tau background. To achieve this, we would group different backgrounds together, so that we obtain the following samples:

- $q\bar{q}$ samples: $u\bar{u}$, $d\bar{d}$, $c\bar{c}$ and $s\bar{s}$
- $e^+e^-\ell^+\ell^-$ samples: $e^+e^-e^+e^-$ and $e^+e^-\mu^+\mu^-$
- $\ell^+\ell^-(\gamma)$ samples: $e^+e^-(\gamma)$ and $\mu^+\mu^-(\gamma)$
- BB samples: $B^0\bar{B}^0$ and B^+B^-
- h^+h^-ISR samples: $h^+h^-(\gamma)$, $K^0\bar{K}^0(\gamma)$ and $\pi^+\pi^-\pi^0(\gamma)$

For each of these groups, we will independently assign a normalisation uncertainty based on the ratio of data over Monte Carlo discrepancies observed, some of which were identified during the computation of correction factors in Section 4.3.5.2. We plan to implement these normalisation uncertainties with correlated shape modifiers (*histosys*).

4.4.5.8 Monte Carlo statistics

The Monte Carlo sample on which the model is determined has a finite size. To account for fluctuations arising from this, *pyhf* provides a dedicated modifier, called *staterror*. This modifier is used to implement this uncertainty in the model.

4.4.5.9 Modeling of p on the signal side

Systematic effects that equally impact both the pion and kaon channels will cancel out in the ratio. To address any effects that could alter the shape of the distribution in a manner not covered by the included shape variations, we introduce a global shape systematic uncertainty. This involves adding uncertainties for each of the 14 bins in our template, which are fully correlated across all six templates but affect only one momentum bin at a time. This is implemented by incorporating correlated shape modifiers (*histosys*) shared across all templates, where the shape deviates from the nominal template in a single bin. This introduces an additional nuisance parameter per template bin. The magnitude of this uncertainty is estimated in a data-driven manner based on the ratio of data over Monte Carlo spread of the combined π and K samples, using the same binning scheme intended for the fit. To avoid bias from the shape or normalisation of the ratio of data over Monte Carlo in the combined sample, we compute the standard deviation without directly inspecting the ratio itself.

4.4.5.10 Tracking p scale factor

As discussed in Section 4.3.3, a tracking momentum correction is applied using a scale factor. To assess the associated systematic uncertainty, we reprocess the data using both the up and down variations of this scale factor, derived from the addition or subtraction of the combined statistical and systematic uncertainties. We then compute the relative variation of the templates corresponding to the varied scale factor compared to the data obtained with the nominal scale factor. This relative variation is then used as a correlated shape variation (*histosys*) applied to the Monte Carlo templates. These shape variations are shown in Figure 4.36.

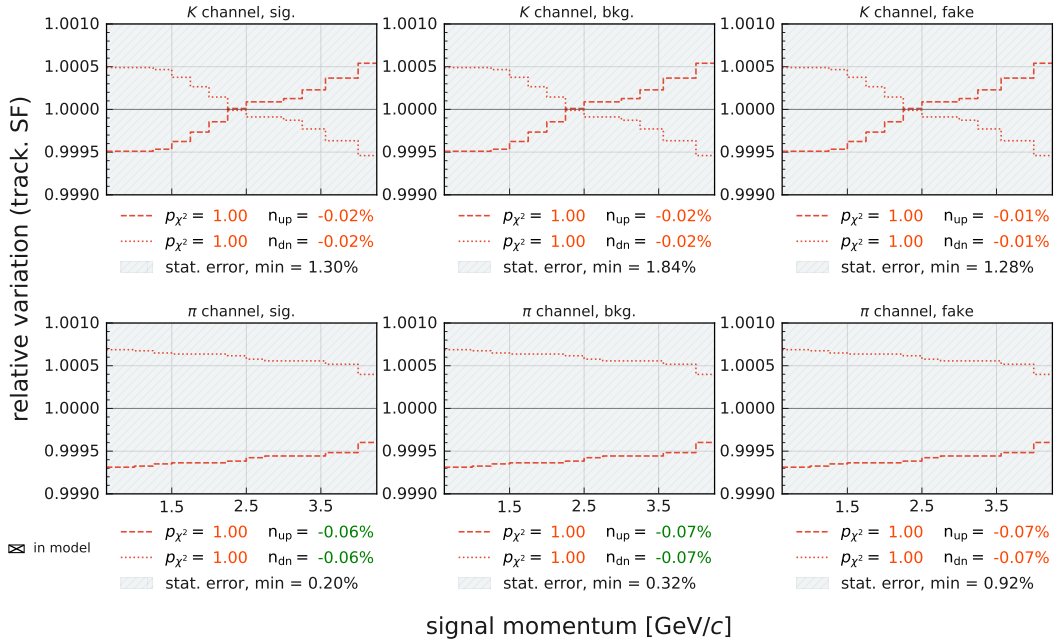


Figure 4.36: *Shape variations originating from the varied tracking momentum scale factor in data.* The smoothed variations are shown. The hID selection is using the Reweighted variable with a working point of 0.9. The BDT output working point is set to 0.5.

4.4.5.11 Tracking efficiency

Following recommendations regarding the tracking efficiency uncertainty, we assign a per-track systematic uncertainty of 0.27 % to Monte Carlo. This uncertainty is fully correlated across all samples and is included in the model through a *normsys* modifier.

4.4.5.12 π^0 efficiency

To correct for π^0 efficiency in Monte Carlo, we use a technique similar to that explained in Section 4.4.5.10. The correction associated to π^0 efficiency along their statistical uncertainties are provided in Table 4.3. As the systematic uncertainty has not been evaluated yet, only the statistical uncertainty is taken for the up and down variation operated on Monte Carlo. These variations are produced using additional Monte Carlo samples with varied π^0 efficiency corrections. The resulting relative variation is displayed in Figure 4.37. This relative variation is then used as a correlated shape variation (*histosys*) applied to the Monte Carlo templates.

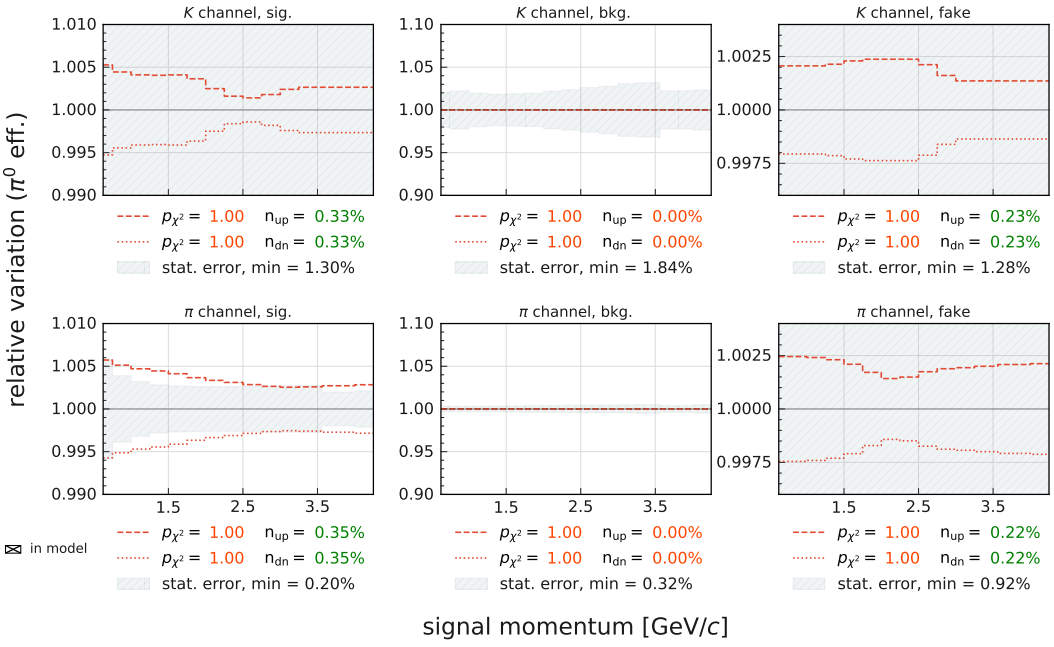


Figure 4.37: *Relative variation of the templates due to the π^0 efficiency correction uncertainty.* The smoothed variations are shown. The *h*ID selection is using the Reweighted variable with a working point of 0.9. The BDT output working point is set to 0.5.

4.4.5.13 Photon efficiency correction

The corrections related to photon efficiency pertain to an earlier version of Monte Carlo samples, as described in Section 4.3.3. Since these corrections are anticipated to be comparable to those for the Monte Carlo version used in this chapter, we would estimate systematic uncertainties by analyzing the corresponding up and down variations. These variations would be produced using additional Monte Carlo samples with varied photon efficiency corrections considering the statistical and

systematic uncertainties combined. The resulting relative shape variation would be implemented in the model as a correlated shape variation (*histosys*).

4.4.5.14 Photon energy bias

A photon energy bias correction, as discussed in Section 4.3.3, is applied on data. To assess the associated systematic uncertainty, we would reprocess the data using both the up and down variations of this correction. We would then compute the relative variation of the templates corresponding to the varied scale factor compared to the data obtained with the nominal correction. This relative variation would then be used as a correlated shape variation (*histosys*) applied to the Monte Carlo templates.

4.5 Results

Several intermediate result have been discussed and shown in the previous sections. Here, we will elaborate on the outcomes following the template fit. Among the thirteen identified sources of systematic uncertainties, eight contributed to these results. The remaining systematic uncertainties are planned for implementation in the near future, pending any potential obstacles.

4.5.1 Template Fit Results

Once the statistical model is created using the available systematic uncertainties, we fit it on Asimov data (Monte Carlo sample). The resulting pre-fit and post-fit distributions are shown in Figures 4.38 and 4.39. The distributions are separated into kaon and pion channels. Since the model is applied on Monte Carlo, it is expected to obtain such a good agreement between the two distributions, as shown in the bottom part of the plots. This step will be more relevant when real data will be used.

Nevertheless, this fit offers an assessment of all implemented systematic uncertainties propagated on $R_{K/\pi}$, which are detailed in Table 4.15. In particular, uncertainties highlighted in bold are integrated into the model. These uncertainties are presented as both relative and absolute values, along with the associated modifier and the number of parameters used.

The particle identification uncertainty appears to be the leading source of uncertainty. It is further categorised into its different contributions, which reveals that

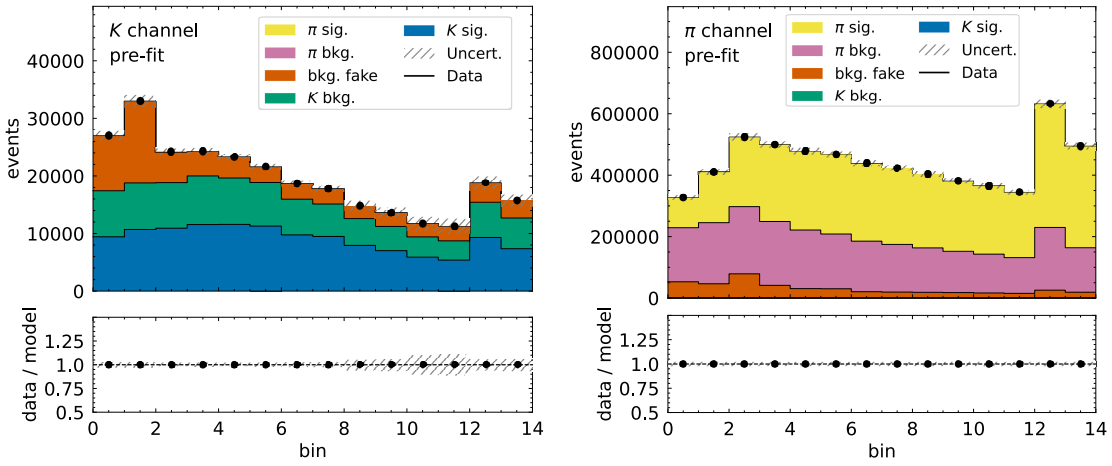


Figure 4.38: *Pre-fit distributions of templates with Asimov data.* The different template contributions are separated. The fit distributions are split into the kaon channel on the left-hand side, and the pion channel on the right-hand side. The hID selection is using the Reweighted variable with a working point of 0.9. The BDT output working point is set to 0.5.

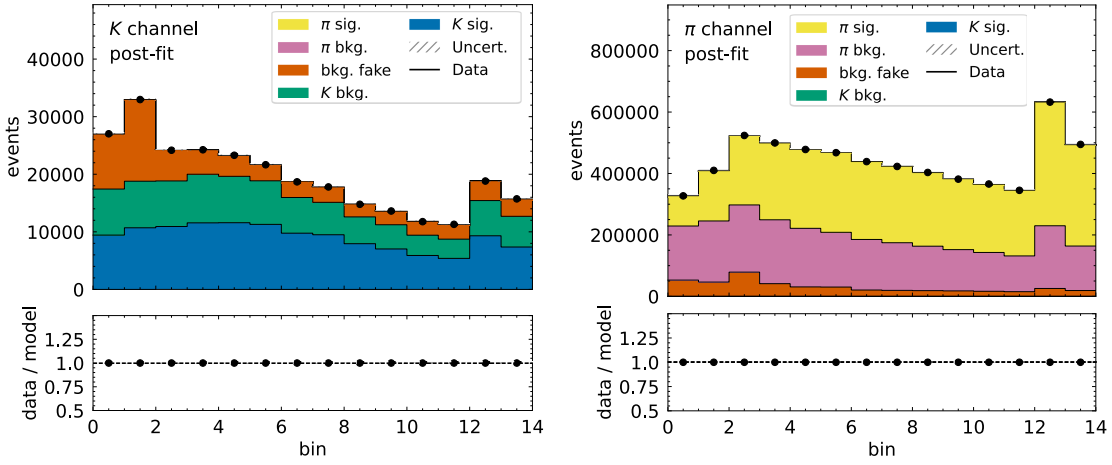


Figure 4.39: *Post-fit distributions of templates with Asimov data.* The different template contributions are separated. The fit distributions are split into the kaon channel on the left-hand side, and the pion channel on the right-hand side. The hID selection is using the Reweighted variable with a working point of 0.9. The BDT output working point is set to 0.5.

the $X \to K$ misID rate is the leading component. The second most significant source of uncertainty is attributed to misalignment, while the remaining sources are one magnitude lower in magnitude.

Uncertainty source	Modifier	Parameters	Uncertainty	
			Rel. [%]	Abs.
Luminosity	*lumi*	1	0.057	0.00004
Charged PID	*hsys*	37	0.896	0.00058
- π**ID**	*hsys*	14	0.112	0.00007
- K**ID**	*hsys*	9	0.301	0.00019
- $X \to \pi$ **misID rate**	*hsys*	7	0.319	0.00021
- $X \to K$ **misID rate**	*hsys*	7	0.692	0.00045
Trigger	*hsys*	42	0.010	0.00001
Misalignment	*ssys*	42	0.388	0.00025
Tag side MC modeling	*ssys*	(42)	-	-
BDT efficiency	*hsys*	(1)	-	-
Bkg normalisation	*nsys*	(6)	-	-
MC stat	*stat*	1	0.330	0.00021
Signal side p modeling	*hsys*	21	0.104	0.00007
Tracking p scale factor	*hsys*	1	0.062	0.00004
Tracking efficiency	*nsys*	1	0.050	0.00003
π^0 **efficiency**	*hsys*	1	0.020	0.00001
γ efficiency corr.	*hsys*	(1)	-	-
γ energy bias	*hsys*	(1)	-	-
Total syst. uncertainty			1.022	0.00066
Total stat. uncertainty			0.706	0.00045
Total uncertainty			1.243	0.00080

Table 4.15: *SUmmary of uncertainties*. The modifiers as well as the number of parameters used to model the corresponding systematic uncertainty are detailed. If the uncertainty is not indicated, the systematic uncertainty has not been yet fully implemented. The hID selection is using the Reweighted variable with a working point of 0.9. The BDT output working point is set to 0.5.

4.5.2 Estimation of $|V_{us}|$

The fitting procedure yields:

$$R_{K/\pi} = 0.06438 \pm 0.00045\ (\text{stat.}) \pm 0.00066\ (\text{syst.}). \tag{4.33}$$

This value is used to derive a value of $|V_{us}|$, following Equation (4.14). This es-

timation involves several additional factors, and their central values along with uncertainties are shown in Table 4.16. These values are the most updated ones as of the time of this thesis. The mass values are from [19], while the other factors values are from taken [21].

Factor	Central value	Uncertainty	Unit
$\lvert V_{ud}\rvert$	0.97373	0.00031	[-]
f_K/f_π	1.1932	0.0021	[-]
$\delta_{R_{\tau,K/\tau,\pi}}$	0.001	0.008	[-]
m_π	139.57039	0.00018	[MeV/c^2]
m_K	493.677	0.016	[MeV/c^2]
m_τ	1776.86	0.12	[MeV/c^2]

Table 4.16: *Summary of factors used to compute* $\lvert V_{us}\rvert$. The mass values are coming from [19] and the other factors values are coming from [21].

Taking the value of $R_{K/\pi}$ shown in Equation (4.33) and the factors detailed in Table 4.16, this yields

$$\lvert V_{us}\rvert = 0.2229 \pm 0.0017. \tag{4.34}$$

The values of both $R_{K/\pi}$ and $\lvert V_{us}\rvert$ are summarised in Table 4.17, along the value obtained by the BaBar collaboration [33].

	Measurement	Central value	Uncertainty	
			Rel.	Abs.
$R_{K/\pi}$	Belle II	0.06438	0.012426	0.00080
	BaBar	0.06531	0.01670	0.00109
$\lvert V_{us}\rvert$	Belle II	0.2229	0.0076	0.0017
	BaBar	0.2255	0.0106	0.0024

Table 4.17: *Summary of* $R_{K/\pi}$ *and* $\lvert V_{us}\rvert$ *values.* The values displayed for the BaBar collaboration come from [33].

4.6 Discussion

Hadronic tau decays with 3×1 prong decay topology have been used to compute the CKM matrix element magnitude $\lvert V_{us}\rvert$ through the ratio of branching fractions $R_{K/\pi}$.

The events were selected in three steps: first, a selection of events relying on rectangular cuts was operated in order to remove regions in the date not well described by the Monte Carlo simulations. A computation of several correction factors then took place is order to adjust the shape of the Monte Carlo simulation to match the data more accurately.

Following these two preliminary steps, a selection of the events was operated using a Boosted Decision Tree (BDT). A set of 24 variables was initially defined, and their correlations and ranking were assessed to reduce this set further in order to use only what is necessary, and simplify the BDT selection.

Several hID selections were explored, comprising two variables and five different working points. The BDT output of these ten configurations were scrupulously analysed to determine the most favorable selection. As the figure of merit only provides an assessment of the best BDT output working point in terms of statistics, and that the measurement is expected to be systematically driven, the choice of the working point was postponed during the evaluation of systematic uncertainties.

A template fit method was put in place to evaluate the ratio $R_{K/\pi}$, which is quadratically proportional to the value of $|V_{us}|$, and has the advantage of not relying on external factors. The fit provided the central value, and the statistical and systematic uncertainties. The latter are implemented in the model as nuisance parameters, through specific modifiers. Not all systematic uncertainties have been implemented yet, but all have been identified, described and will be added in due time.

A preliminary assessment of $R_{K/\pi}$, using the implemented systematic uncertainties, yields $R_{K/\pi} = 0.06438 \pm 0.00080$, which in turn gives $|V_{us}| = 0.2229 \pm 0.0017$.

While these results offer an initial insight into the measurement of $|V_{us}|$, there are already identified areas that can be further refined and improved.

The corrections for the leptonic fake rates were not accessible through the *Systematic Corrections Framework*, and therefore only few tables were available. In particular, it was only possible to obtain the correction tables for the hID selection involving the Global hID variable, with a working point of 0.6. Several efforts are ongoing to make other tables accessible, but as this represents not just one be several major studies, it will take time.

Other corrections should be updated: the π^0 efficiency correction will be assessed for different channels, and will have an associated systematic uncertainty. As for the photon efficiency correction, their central value as well as their total uncertainty should be soon accessible. The measurement will have to be reiterated with these

new central values, and uncertainties.

Once these corrections are in place, the BDT method will be able to undergo several other possible optimisations. As the optimisation of its hyper parameters did not result in an improved performance, this aspect will be necessarily revisited. But one could think for example that another method such as a Neural Network could improve the results.

It will be also possible to assess again the systematic uncertainties related to $R_{K/\pi}$. In particular, the systematic uncertainty associated to π^0 efficiency will be studied more carefully, as its result for the background templates led a value close to 0. As for the systematic uncertainty related to misalignment, it might be revisited as its propagated uncertainty on $R_{K/\pi}$ is significant, and might be more under control after the production of larger samples of misalignment, or by making less conservative assumptions.

On top of that, the other systematic sources, detailed in Sections 4.4.5.5, 4.4.5.6, 4.4.5.7, 4.4.5.14 and 4.4.5.13 will be implemented, in order to give a solid measurement of $R_{K/\pi}$. Their final implementation might slightly change from the initial description, given unknown challenges that can be encountered.

Once the fit validated, it will be possible to use data. At first, a reduced amount of data will be used to study potential issues that could arise at that stage without unblinding the measurement. After several checks to validate the measurement, and iterations on the possible issues, the whole LS1 dataset will be used to finalise the measurement.

Conclusion

"Attendre et espérer!"

- Alexandre Dumas, *Le Comte de Monte-Cristo*, (1846)

Several key elements related to the measurement of $|V_{us}|$ have been explored throughout this thesis.

It began with the study of pion identification corrections, as the understanding of hadron identification limitations was an important aspect of the systematic uncertainties related to $|V_{us}|$. For this, a 3×1 prong tau decay topology was used to calculate correction factors in p and $\cos\theta$ bins for three PID variables, employing a cut-based event selection that achieved a purity of 98.749%. The PID variables included Global, Reweighted and MVA hID variables.

Each performance was evaluated by computing πID efficiency and $\pi \to K$ misID rate across integrated samples and in p and $\cos\theta$ bins. Observations included the MVA hID performing the poorest with increasing cut values, while the Reweighted hID outperformed the Global hID in both performance and data and Monte Carlo consistency. Correction tables were generated for each hID variable across 9 cut values in p and $\cos\theta$ bins. Associated statistical and systematic uncertainties were assessed and indicated on these tables, categorised into modeling of $\tau \to \pi\pi\pi$ decay, kaon fake rates, trigger selection, and $q\bar{q}$ background. Comparisons were made with D^* and K_s samples, showing reasonable agreement in p or $\cos\theta$ bins with D^*, but notable discrepancies with K_s, particularly highlighted by the pulls. Further comparisons between D^* and K_s samples highlighted discrepancies between these channels. While the study results are not yet available via the *Systematic Corrections Framework*, they align with the available corrections.

The CKM matrix element magnitude $|V_{us}|$ was determined using hadronic tau decays with a 3×1 prong decay topology, through the ratio of branching fractions $R_{K/\pi}$. Event selection involved three key steps: initial event selection using rectangular cuts to align data with Monte Carlo simulations, computation of correction factors to better match simulation shapes with data, and a subsequent selection using a

Boosted Decision Tree (BDT) with optimised variable sets.

For hID selections, various configurations were explored and evaluated based on BDT outputs to identify the most effective working point. A template fit method binned in momentum was used to evaluate $R_{K/\pi}$, which is quadratically related to $|V_{us}|$, providing central values and both statistical and systematic uncertainties.

A preliminary assessment yielded $R_{K/\pi} = 0.06438 \pm 0.00045$ (stat.) ± 0.00066 (syst.), corresponding to $|V_{us}| = 0.2229 \pm 0.0017$. Additional systematic uncertainties were identified and will be incorporated into the final analysis in the near future.

References

[1] I. Falconer. “J. J. Thomson and the discovery of the electron”. In: *Physics Education* 32.4 (July 1997), p. 226. DOI: 10.1088/0031-9120/32/4/015. URL: https://dx.doi.org/10.1088/0031-9120/32/4/015.

[2] E. Rutherford. “LXXIX. The scattering of α and β particles by matter and the structure of the atom”. In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 21.125 (1911), pp. 669–688. DOI: 10.1080/14786440508637080. URL: https://doi.org/10.1080/14786440508637080.

[3] A. Einstein. “Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt”. In: *Annalen der Physik* 17.6 (1905), pp. 132–148.

[4] A. H. Compton. “A Quantum Theory of the Scattering of X-rays by Light Elements”. In: *Phys. Rev.* 21 (5 May 1923), pp. 483–502. DOI: 10.1103/PhysRev.21.483. URL: https://link.aps.org/doi/10.1103/PhysRev.21.483.

[5] J. Chadwick. “The existence of a neutron”. In: *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 136.830 (1932), pp. 692–708. DOI: 10.1098/rspa.1932.0112. eprint: https://royalsocietypublishing.org/doi/pdf/10.1098/rspa.1932.0112. URL: https://royalsocietypublishing.org/doi/abs/10.1098/rspa.1932.0112.

[6] C. D. Anderson. “The Positive Electron”. In: *Phys. Rev.* 43 (6 Mar. 1933), pp. 491–494. DOI: 10.1103/PhysRev.43.491. URL: https://link.aps.org/doi/10.1103/PhysRev.43.491.

[7] S. H. Neddermeyer and C. D. Anderson. “Note on the Nature of Cosmic-Ray Particles”. In: *Phys. Rev.* 51 (10 May 1937), pp. 884–886. DOI: 10.1103/PhysRev.51.884. URL: https://link.aps.org/doi/10.1103/PhysRev.51.884.

[8] R. Oerter and B. Holstein. “The Theory of Almost Everything: The Standard Model, the Unsung Triumph of Modern Physics”. In: *Physics Today - PHYS TODAY* 59 (July 2006). DOI: 10.1063/1.2337829.

[9] F. Abe et al. “Observation of Top Quark Production in $\bar{p}p$ Collisions with the Collider Detector at Fermilab”. In: *Physical Review Letters* 74.14 (Apr. 1995), pp. 2626–2631. ISSN: 1079-7114. DOI: 10.1103/physrevlett.74.2626. URL: http://dx.doi.org/10.1103/PhysRevLett.74.2626.

[10] G. Aad et al. “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”. In: *Physics Letters B* 716.1 (Sept. 2012), pp. 1–29. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2012.08.020. URL: http://dx.doi.org/10.1016/j.physletb.2012.08.020.

[11] Y. Fukuda and et al. “Evidence for Oscillation of Atmospheric Neutrinos”. In: *Physical Review Letters* 81.8 (Aug. 1998), pp. 1562–1567. ISSN: 1079-7114. DOI: 10.1103/physrevlett.81.1562. URL: http://dx.doi.org/10.1103/PhysRevLett.81.1562.

[12] V. C. Rubin and W. K. Ford. “Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions”. In: *Astrophysical Journal* 159 (Feb. 1970), p. 379. DOI: 10.1086/150317.

[13] D. P. Aguillard and et al. “Measurement of the Positive Muon Anomalous Magnetic Moment to 0.20 ppm”. In: *Phys. Rev. Lett.* 131 (16 Oct. 2023), p. 161802. DOI: 10.1103/PhysRevLett.131.161802. URL: https://link.aps.org/doi/10.1103/PhysRevLett.131.161802.

[14] I. Adachi and et al. “Search for an Invisibly Decaying Z' Boson at Belle II in $e^+e^- \to \mu^+\mu^-(e^\pm\mu^\mp)$ Plus Missing Energy Final States”. In: *Physical Review Letters* 124.14 (Apr. 2020). ISSN: 1079-7114. DOI: 10.1103/physrevlett.124.141801. URL: http://dx.doi.org/10.1103/PhysRevLett.124.141801.

[15] I. Adachi and et al. “Search for an Invisible Z' in a Final State with Two Muons and Missing Energy at Belle II”. In: *Physical Review Letters* 130.23 (June 2023). ISSN: 1079-7114. DOI: 10.1103/physrevlett.130.231801. URL: http://dx.doi.org/10.1103/PhysRevLett.130.231801.

[16] N. Cabibbo. “Unitary Symmetry and Leptonic Decays”. In: *Phys. Rev. Lett.* 10 (12 June 1963), pp. 531–533. DOI: 10.1103/PhysRevLett.10.531. URL: https://link.aps.org/doi/10.1103/PhysRevLett.10.531.

[17] S. L. Glashow, J. Iliopoulos, and L. Maiani. "Weak Interactions with Lepton-Hadron Symmetry". In: *Phys. Rev. D* 2 (7 Oct. 1970), pp. 1285–1292. DOI: 10.1103/PhysRevD.2.1285. URL: https://link.aps.org/doi/10.1103/PhysRevD.2.1285.

[18] M. Kobayashi and T. Maskawa. "CP-Violation in the Renormalizable Theory of Weak Interaction". In: *Progress of Theoretical Physics* 49.2 (Feb. 1973), pp. 652–657. ISSN: 0033-068X. DOI: 10.1143/PTP.49.652. eprint: https://academic.oup.com/ptp/article-pdf/49/2/652/5257692/49-2-652.pdf. URL: https://doi.org/10.1143/PTP.49.652.

[19] R. L. Workman et al. "Review of Particle Physics". In: *PTEP* 2022 (2022), p. 083C01. DOI: 10.1093/ptep/ptac097.

[20] Lusiani, A. *Status and plans of tau fits for HFLAV/PDG*. https://indico.cern.ch/event/1303630/contributions/5628292/attachments/2768550/4823164/alusiani-tau23-hflav.pdf. Dec. 2023.

[21] Y. Amhis et al. "Averages of b-hadron, c-hadron, and τ-lepton properties as of 2021". In: *Phys. Rev. D* 107 (5 Mar. 2023), p. 052008. DOI: 10.1103/PhysRevD.107.052008. URL: https://link.aps.org/doi/10.1103/PhysRevD.107.052008.

[22] J. C. Hardy and I. S. Towner. "Superallowed $0^+ \to 0^+$ nuclear β decays: 2020 critical survey, with implications for V_{ud} and CKM unitarity". In: *Phys. Rev. C* 102 (4 Oct. 2020), p. 045501. DOI: 10.1103/PhysRevC.102.045501. URL: https://link.aps.org/doi/10.1103/PhysRevC.102.045501.

[23] V. Cirigliano et al. "Scrutinizing CKM unitarity with a new measurement of the Kμ3/Kμ2 branching fraction". In: *Phys. Lett. B* 838 (2023), p. 137748. DOI: 10.1016/j.physletb.2023.137748. arXiv: 2208.11707 [hep-ph].

[24] M. Antonelli et al. "An evaluation of $|V_{us}|$ and precise tests of the Standard Model from world data on leptonic and semileptonic kaon decays". In: *The European Physical Journal C* 69.3–4 (Aug. 2010), pp. 399–424. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-010-1406-3. URL: http://dx.doi.org/10.1140/epjc/s10052-010-1406-3.

[25] W. J. Marciano. "Precise Determination of $|V_{us}|$ from Lattice Calculations of Pseudoscalar Decay Constants". In: *Phys. Rev. Lett.* 93 (23 Dec. 2004), p. 231803. DOI: 10.1103/PhysRevLett.93.231803. URL: https://link.aps.org/doi/10.1103/PhysRevLett.93.231803.

[26] K. Maltman et al. "Status of the hadronic τ determination of V_{us}". In: *International Journal of Modern Physics A* 23.21 (Aug. 2008), pp. 3191–3195. ISSN: 1793-656X. DOI: 10.1142/s0217751x08041803. URL: http://dx.doi.org/10.1142/S0217751X08041803.

[27] F. J. Botella et al. "New physics and evidence for a complex CKM". In: *Nuclear Physics B* 725.1–2 (Oct. 2005), pp. 155–172. ISSN: 0550-3213. DOI: 10.1016/j.nuclphysb.2005.07.006. URL: http://dx.doi.org/10.1016/j.nuclphysb.2005.07.006.

[28] E. Gámiz et al. "Determination of m_s and $|V_{us}|$ from hadronic τ decays". In: *Journal of High Energy Physics* 2003.01 (Feb. 2003), p. 060. DOI: 10.1088/1126-6708/2003/01/060. URL: https://dx.doi.org/10.1088/1126-6708/2003/01/060.

[29] E. Gámiz et al. "V_{us} and m_s from hadronic tau decays". In: *Physical Review Letters* 94.1 (Jan. 2005). ISSN: 1079-7114. DOI: 10.1103/physrevlett.94.011803. URL: http://dx.doi.org/10.1103/PhysRevLett.94.011803.

[30] K. Maltman et al. "Status of the Hadronic τ Decay Determination of V_{us}". In: *Nuclear Physics B - Proceedings Supplements* 189 (2009). Proceedings of the Tenth International Workshop on Tau Lepton Physics, pp. 175–180. ISSN: 0920-5632. DOI: https://doi.org/10.1016/j.nuclphysbps.2009.03.031. URL: https://www.sciencedirect.com/science/article/pii/S0920563209002229.

[31] M. Hernández and J. Alejandro. "Electromagnetic corrections in hadronic tau decays". In: *Suplemento de la Revista Mexicana de Física* 4.2 (Sept. 2023), pp. 021115 1–7. DOI: 10.31349/SuplRevMexFis.4.021115. URL: https://rmf.smf.mx/ojs/index.php/rmf-s/article/view/7116.

[32] S. Banerjee. *Lepton Universality, $|V_{us}|$ and search for second class current in tau decays*. 2008. arXiv: 0811.1429 [hep-ex].

[33] B. Aubert and et al. "Measurements of Charged Current Lepton Universality and $|V_{\mathrm{us}}|$ Using Tau Lepton Decays to $e^-\overline{\nu}_e\nu_\tau$, $\mu^-\overline{\nu}_\mu\nu_\tau$, $\pi^-\nu_\tau$, and $K^-\nu_\tau$". In: *Phys. Rev. Lett.* 105 (5 2010), p. 051602. DOI: 10.1103/PhysRevLett.105.051602.

[34] A. Ishikawa et al. "Observation of $B \to K^*\ell^+\ell^-$". In: *Physical Review Letters* 91.26 (Dec. 2003). ISSN: 1079-7114. DOI: 10.1103/physrevlett.91.261601. URL: http://dx.doi.org/10.1103/PhysRevLett.91.261601.

[35] K. Abe et al. “Observation of Large *CP* Violation in the Neutral *B* Meson System”. In: *Physical Review Letters* 87.9 (Aug. 2001). ISSN: 1079-7114. DOI: 10.1103/physrevlett.87.091802. URL: http://dx.doi.org/10.1103/PhysRevLett.87.091802.

[36] A. J. Bevan et al. “The Physics of the B Factories”. In: *The European Physical Journal C* 74.11 (Nov. 2014). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-014-3026-9. URL: http://dx.doi.org/10.1140/epjc/s10052-014-3026-9.

[37] E. Kou et al. “The Belle II Physics Book”. In: *Progress of Theoretical and Experimental Physics* 2019.12 (Dec. 2019), p. 123C01. ISSN: 2050-3911. DOI: 10.1093/ptep/ptz106. eprint: https://academic.oup.com/ptep/article-pdf/2019/12/123C01/32693980/ptz106.pdf. URL: https://doi.org/10.1093/ptep/ptz106.

[38] K. Akai, K. Furukawa, and H. Koiso. “SuperKEKB collider”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 907 (Aug. 2018). DOI: 10.1016/j.nima.2018.08.017.

[39] D. Besson and T. Skwarnicki. “Upsilon Spectroscopy: Transitions in the Bottomonium System”. In: *Annual Review of Nuclear and Particle Science* 43.1 (1993), pp. 333–378. DOI: 10.1146/annurev.ns.43.120193.002001. eprint: https://doi.org/10.1146/annurev.ns.43.120193.002001. URL: https://doi.org/10.1146/annurev.ns.43.120193.002001.

[40] SuperB Collaboration. *SuperB: A High-Luminosity Asymmetric e^+ e^- Super Flavor Factory. Conceptual Design Report*. 2007. arXiv: 0709.0451 [hep-ex].

[41] P. M. Lewis et al. “First measurements of beam backgrounds at SuperKEKB”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 914 (Jan. 2019), pp. 69–144. ISSN: 0168-9002. DOI: 10.1016/j.nima.2018.05.071. URL: http://dx.doi.org/10.1016/j.nima.2018.05.071.

[42] Bennett, J. *Flavor physics as a tool for discovery at Belle II*. https://indico.fnal.gov/event/58272/contributions/261984/attachments/165183/219405/230323_P5_BelleII_Bennett.pdf. Mar. 2023.

[43] T. Abe and et al. “Belle II technical design report”. In: (2010). arXiv: 1011.0352 [physics.ins-det].

[44] M. Bertemes, G. Inguglia, and C. Schwanda. “Search for dark sector physics and performance of the Belle II detector in final state events with muons and large missing energy”. PhD thesis. Vienna: Vienna, Technische Universität Wien, 2021.

[45] Belle-ECL and et al. “Electromagnetic calorimeter for Belle II”. In: *Journal of Physics: Conference Series* 587.1 (Feb. 2015), p. 012045. DOI: 10.1088/1742-6596/587/1/012045. URL: https://dx.doi.org/10.1088/1742-6596/587/1/012045.

[46] KEK. *A Search for New Physics - The Belle II Experiment*. Youtube. 2014. URL: https://www.youtube.com/watch?v=nGCrrgXSEOk.

[47] J. Kemmer and G. Lutz. “New detector concepts”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 253.3 (1987), pp. 365–377. ISSN: 0168-9002. DOI: https://doi.org/10.1016/0168-9002(87)90518-3. URL: https://www.sciencedirect.com/science/article/pii/0168900287905183.

[48] P. Križan, S. Korpar, and T. Iijima. “Study of a nonhomogeneous aerogel radiator in a proximity focusing RICH detector”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 565.2 (2006), pp. 457–462. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2006.05.233. URL: https://www.sciencedirect.com/science/article/pii/S0168900206010229.

[49] S. Longo et al. “CsI(Tl) pulse shape discrimination with the Belle II electromagnetic calorimeter as a novel method to improve particle identification at electron–positron colliders”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 982 (2020), p. 164562. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2020.164562. URL: https://www.sciencedirect.com/science/article/pii/S0168900220309591.

[50] T. Ogitsu et al. “Superconducting combined function magnet system for J-PARC neutrino experiment”. In: *IEEE Transactions on Applied Superconductivity* 15.2 (2005), pp. 1175–1180. DOI: 10.1109/TASC.2005.849525.

[51] Y. Makida et al. “Development of a Superconducting Solenoid Magnet System for the B-Factory Detector (BELLE)”. In: *Advances in Cryogenic Engineering*. Ed. by Peter Kittel. Boston, MA: Springer US, 1998, pp. 221–228. ISBN: 978-1-4757-9047-4. DOI: 10.1007/978-1-4757-9047-4_25. URL: https://doi.org/10.1007/978-1-4757-9047-4_25.

[52] *Belle II Software Documentation*. https://software.belle2.org/development/sphinx/index.html. Accessed: 2024-02-04.

[53] T. Gesler et al. "The ONSEN Data Reduction System for the Belle II Pixel Detector". In: *IEEE Transactions on Nuclear Science* 62.3 (June 2015), pp. 1149–1154. ISSN: 1558-1578. DOI: 10.1109/tns.2015.2414713. URL: http://dx.doi.org/10.1109/TNS.2015.2414713.

[54] T. Kuhr et al. "The Belle II Core Software". In: *Comput. Softw. Big Sci.* 3.1 (2019), p. 1. DOI: 10.1007/s41781-018-0017-9. arXiv: 1809.04299.

[55] *Belle II Analysis Software Framework (basf2)*. https://doi.org/10.5281/zenodo.5574115. DOI: 10.5281/zenodo.5574115.

[56] S. Jadach, B. F. L. Ward, and Z. Was. "The precision Monte Carlo event generator KK for two-fermion final states in e^+e^- collisions". In: *Comput. Phys. Commun.* 130 (2000), p. 260. DOI: 10.1016/S0010-4655(00)00048-5.

[57] S. Jadach, B. F. L. Ward, and Z. Was. "Coherent exclusive exponentiation for precision Monte Carlo calculations". In: *Phys. Rev.* D 63 (2001), p. 113009. DOI: 10.1103/PhysRevD.63.113009.

[58] T. Sjöstrand et al. "An Introduction to PYTHIA 8.2". In: *Comput. Phys. Commun.* 191 (2015), pp. 159–177. DOI: 10.1016/j.cpc.2015.01.024.

[59] G. Balossini et al. "Matching perturbative and parton shower corrections to Bhabha process at flavour factories". In: *Nucl. Phys.* B 758 (2006), pp. 227–253. DOI: 10.1016/j.nuclphysb.2006.09.022.

[60] G. Balossini et al. "Photon pair production at flavour factories with per mille accuracy". In: *Phys. Lett.* B 663 (2008), pp. 209–213. DOI: 10.1016/j.physletb.2008.04.007.

[61] C. M. Carloni Calame et al. "The BABAYAGA event generator". In: *Nucl. Phys.* **B** *Proc. Suppl.* 131 (2004). Ed. by Marco Incagli and G. Venanzoni, pp. 48–55. DOI: 10.1016/j.nuclphysbps.2004.02.008.

[62] C. M. Carloni Calame. "An improved parton shower algorithm in QED". In: *Phys. Lett.* B 520 (2001), pp. 16–24. DOI: 10.1016/S0370-2693(01)01108-X.

[63] C. M. Carloni Calame et al. "Large angle Bhabha scattering and luminosity at flavor factories". In: *Nucl. Phys.* B 584 (2000), pp. 459–479. DOI: 10.1016/S0550-3213(00)00356-4.

[64] S. Banerjee et al. *Monte Carlo Event Generator updates, for tau pair events at Belle II energies*. 2021. arXiv: 2111.05914 [hep-ph].

[65] O. Shekhovtsova et al. “Resonance chiral Lagrangian currents and τ decay Monte Carlo”. In: *Phys. Rev.* D 86 (2012), p. 113008. DOI: 10.1103/PhysRevD.86.113008.

[66] M. Chrzaszcz et al. “TAUOLA of τ lepton decays—framework for hadronic currents, matrix elements and anomalous decays”. In: *Comput. Phys. Commun.* 232 (2018), pp. 220–236. DOI: 10.1016/j.cpc.2018.05.017.

[67] I. M. Nugent et al. “Resonance chiral Lagrangian currents and experimental data for $\tau^- \to \pi^-\pi^-\pi^+\nu_\tau$”. In: *Phys. Rev.* D 88 (2013), p. 093012. DOI: 10.1103/PhysRevD.88.093012.

[68] S. Jadach, J. H. Kuhn, and Z. Was. “TAUOLA: A library of Monte Carlo programs to simulate decays of polarized tau leptons”. In: *Comput. Phys. Commun.* 64 (1990), p. 275. DOI: 10.1016/0010-4655(91)90038-M.

[69] E. Barberio, B. van Eijk, and Z. Was. “PHOTOS: A universal Monte Carlo for QED radiative corrections in decays”. In: *Comput. Phys. Commun.* 66 (1991), p. 115. DOI: 10.1016/0010-4655(91)90012-A.

[70] F. A. Berends, P. H. Daverveldt, and R. Kleiss. “Radiative corrections to the process $e^+e^- \to e^+e^-\mu^+\mu^-$”. In: *Nucl. Phys.* B 253 (1985), pp. 421–440. DOI: https://doi.org/10.1016/0550-3213(85)90540-1.

[71] F. A. Berends, P. H. Daverveldt, and R. Kleiss. “Complete lowest-order calculations for four-lepton final states in electron-positron collisions”. In: *Nucl. Phys.* B 253 (1985), pp. 441–463. DOI: https://doi.org/10.1016/0550-3213(85)90541-3.

[72] F. A. Berends, P. H. Daverveldt, and R. Kleiss. “Monte Carlo simulation of two-photon processes: II: Complete lowest order calculations for four-lepton production processes in electron-positron collisions”. In: *Comp. Phys. Commun.* 40.2 (1986), pp. 285–307. DOI: https://doi.org/10.1016/0010-4655(86)90115-3.

[73] S. Uehara. “TREPS: A Monte-Carlo Event Generator for Two-photon Processes at e^+e^- Colliders using an Equivalent Photon Approximation”. In: (July 1996). arXiv: 1310.0157 [hep-ph].

[74] D. J. Lange. “The EvtGen particle decay simulation package”. In: *Nucl. Instrum. Meth.* A462 (2001), p. 152. DOI: 10.1016/S0168-9002(01)00089-4.

[75] J. Alwall et al. “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”. In: *Journal of High Energy Physics* 2014.7 (July 2014). ISSN: 1029-8479. DOI: 10.1007/jhep07(2014)079. URL: http://dx.doi.org/10.1007/JHEP07(2014)079.

[76] S. Agostinelli and et al. “GEANT4: A simulation toolkit”. In: *Nucl.Instrum.Meth.* A506 (2003), pp. 250–303. DOI: 10.1016/S0168-9002(03)01368-8.

[77] G. C. Fox and S. Wolfram. “Observables for the Analysis of Event Shapes in e^+e^- Annihilation and Other Processes”. In: *Phys. Rev. Lett.* 41 (23 Dec. 1978), pp. 1581–1585. DOI: 10.1103/PhysRevLett.41.1581. URL: https://link.aps.org/doi/10.1103/PhysRevLett.41.1581.

[78] G. C. Fox and S. Wolfram. “Event shapes in e+e- annihilation”. In: *Nuclear Physics B* 157.3 (1979), pp. 543–544. ISSN: 0550-3213. DOI: https://doi.org/10.1016/0550-3213(79)90120-2. URL: https://www.sciencedirect.com/science/article/pii/0550321379901202.

[79] A. J. Weinstein. “Semi-hadronic tau decays at CLEO”. In: *Nuclear Physics B - Proceedings Supplements* 98.1–3 (Apr. 2001), pp. 261–280. ISSN: 0920-5632. DOI: 10.1016/s0920-5632(01)01235-x. URL: http://dx.doi.org/10.1016/S0920-5632(01)01235-X.

[80] G. Räuber. “Testing lepton flavour universality in $B \to K\pi\pi\ell\ell$ decays using 2017 LHCb data”. Ecole Polytechnique Fédérale, Lausanne, 2021. URL: https://cds.cern.ch/record/2847709.

[81] K. Woodruff. “Introduction to boosted decision trees”. Sept. 2017. URL: https://indico.fnal.gov/event/15356/contributions/31377/attachments/19671/24560/DecisionTrees.pdf.

[82] Y. Freund and R. E. Schapire. “A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting”. In: *Journal of Computer and System Sciences* 55.1 (1997), pp. 119–139. ISSN: 0022-0000. DOI: https://doi.org/10.1006/jcss.1997.1504. URL: http://www.sciencedirect.com/science/article/pii/S002200009791504X.

[83] J. H. Friedman. “Greedy function approximation: A gradient boosting machine.” In: *Ann. Statist.* 29.5 (Oct. 2001), pp. 1189–1232. DOI: 10.1214/aos/1013203451. URL: https://doi.org/10.1214/aos/1013203451.

[84] A. P. Bradley. “The use of the area under the ROC curve in the evaluation of machine learning algorithms”. In: *Pattern Recognition* 30.7 (1997), pp. 1145–1159. ISSN: 0031-3203. DOI: https://doi.org/10.1016/S0031-3203(96)00142-2. URL: https://www.sciencedirect.com/science/article/pii/S0031320396001422.

[85] “Kolmogorov–Smirnov Test”. In: *The Concise Encyclopedia of Statistics*. New York, NY: Springer New York, 2008, pp. 283–287. ISBN: 978-0-387-32833-1. DOI: 10.1007/978-0-387-32833-1_214. URL: https://doi.org/10.1007/978-0-387-32833-1_214.

[86] D. M. Asner and et al. “Search for exclusive charmless hadronicBdecays”. In: *Physical Review D* 53.3 (Feb. 1996), pp. 1039–1050. ISSN: 1089-4918. DOI: 10.1103/physrevd.53.1039. URL: http://dx.doi.org/10.1103/PhysRevD.53.1039.

[87] M. Fujikawa and et al. “High-statistics study of the $\tau \to \pi^- \pi^0 \nu_\tau$ Decay”. In: *Physical Review D* 78.7 (2008). DOI: 10.1103/physrevd.78.072006. URL: https://doi.org/10.1103%2Fphysrevd.78.072006.

[88] G. Ke et al. “LightGBM: A Highly Efficient Gradient Boosting Decision Tree”. In: *Advances in Neural Information Processing Systems*. Ed. by I. Guyon et al. Vol. 30. Curran Associates, Inc., 2017. URL: https://proceedings.neurips.cc/paper_files/paper/2017/file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf.

[89] T. Akiba et al. “Optuna: A Next-generation Hyperparameter Optimization Framework”. In: *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2019.

[90] Microsoft Corporation. *LightGBM*. Version faba8177. URL: https://lightgbm.readthedocs.io/en/latest/index.html.

[91] F. Pedregosa et al. “Scikit-learn: Machine Learning in Python”. In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830.

[92] P. Refaeilzadeh, L. Tang, and H. Liu. “Cross-Validation”. In: *Encyclopedia of Database Systems*. Ed. by LING LIU and M. TAMER ÖZSU. Boston, MA: Springer US, 2009, pp. 532–538. ISBN: 978-0-387-39940-9. DOI: 10.1007/978-0-387-39940-9_565. URL: https://doi.org/10.1007/978-0-387-39940-9_565.

[93] L. Heinrich, M. Feickert, and G. Stark. *pyhf: v0.7.6*. Version 0.7.6. DOI: 10.5281/zenodo.1169739. URL: https://doi.org/10.5281/zenodo.1169739.

[94] L. Heinrich et al. “pyhf: pure-Python implementation of HistFactory statistical models”. In: *Journal of Open Source Software* 6.58 (2021), p. 2823. DOI: 10.21105/joss.02823. URL: https://doi.org/10.21105/joss.02823.

[95] K. Cranmer et al. “HistFactory: A tool for creating statistical models for use with RooFit and RooStats”. In: (June 2012).

[96] V. Blobel. *Smoothing*. https://www.desy.de/~sschmitt/blobel/blobel_smooth.pdf. Accessed: 2024-04-24.

[97] J. H. Friedman. “Data analysis techniques for high energy particle physics”. In: (1974). DOI: 10.5170/CERN-1974-023.271. URL: https://cds.cern.ch/record/695770.

[98] T. Bilka and et al. “Alignment for the first precision measurements at Belle II”. In: *EPJ Web Conf.* 245 (2020). Ed. by C. Doglioni et al., p. 02023. DOI: 10.1051/epjconf/202024502023.

[99] T. Bilka et al. “Simultaneous Global and Local Alignment of the Belle II Tracking Detectors”. In: *EPJ Web Conf.* 251 (2021), p. 03028. DOI: 10.1051/epjconf/202125103028.

[100] *HEPData, Search for an invisible Z' in a final state with two muons and missing energy at Belle II*. https://www.hepdata.net/record/ins2611344. Accessed: 2024-04-22.

[101] ECFA Early-Career Researcher Panel et al. *The ECFA Early Career Researcher's Panel: composition, structure, and activities, 2021 – 2022*. 2022. arXiv: 2212.11238 [physics.acc-ph].

[102] J. Allen et al. *Results of the 2022 ECFA Early-Career Researchers Panel survey on career prospects and diversity*. 2024. arXiv: 2404.02074 [physics.soc-ph].

Acknowledgements

"[...] et que ma reconnaissance reste dans l'ombre comme ton bienfait."
- Alexandre Dumas, *Le Comte de Monte-Cristo*, (1846)

It holds importance to express gratitude to those who have also, in one way or another, contributed to this work.

I'd like to thank first my supervisors *Gianluca Inguglia* and *Christoph Schwanda* for their guidance, trust and endless support. Among many other things, you allowed me to get confidence in my research and to learn so many things from other people by attending several conferences and fruitful meetings. I'm very grateful for the opportunities you gave me during this PhD.

This analysis definitely wouldn't have been the same without *Michel Hernández Villanueva*. I'd like to thank you for your patience, for explaining even the simplest things over and over to me, for your encouragements, and your availability, among all the different things I could thank you for.

I also cannot forget to thank the people from my group at HEPHY. Sharing with you our office, going to Japan, singing karaoke, grabbing lunch, sitting on sticky benches, going to Christmas markets, doing pub quizzes, or having "*just one*" beer will always be fond memories. You were doing much more than you can imagine; thank you for being incredible human beings.

If I think about HEPHY, I have also a special thought for the administration staff. I thank you for your work and for kindly helping me with all sorts of documents, and to always be easy to reach to discuss anything.

Also, during this PhD, I had the opportunity to meet several great people from the Belle II collaboration. I've had the chance of working alongside you, learning from you and sometimes sharing moments with you. I'd like to thank the ones who have contributed to this analysis in one way or another, especially the ones from the τ, *Neutral*/*PID* and of course *lfu-vus* working groups. I'm particularly grateful to people at Charles University in Prague and DESY in Hamburg.

J'ai également une pensée pour mes proches et mes amis. Je les remercie d'avoir compris mon départ pour une ville plus grande, de m'avoir soutenue, motivée, et d'avoir toujours entretenu notre contact. J'ai toujours eu du plaisir à vous voir, que ce soit en Suisse ou lors de vos visites à Vienne, mais aussi simplement de recevoir de vos nouvelles.

Je remercie aussi ma famille, qui a toujours été là, même lorsque que physiquement elle se trouvait ailleurs. Leur soutien a été infaillible du début jusqu'à la fin, et m'a permis de vivre cette aventure inoubliable. À aucun moment je me suis sentie seule, quand bien même je faisais face à des sommets alors jamais explorés.

Je pense finalement à mon amoureux, à qui je dois bien plus que ce que je ne pourrais jamais écrire. Tu m'as tenu la main dans les bons comme dans les plus durs moments. Et pour rien au monde je ne voudrais la lâcher.

Appendix

A Physics Performance: Additional material

Additional material related to Chapter 3 is shown in what follows. The section naming convention used here is such that it repeats the relevant sections from the original chapter.

A.1 Efficiencies

Tables providing the efficiency computed for Monte Carlo, along their propagated statistical and systematic uncertainties obtained for each hID variable with a working point of 0.1, and 0.9 are displayed respectively in Figures A.1 and A.2. Finally, tables providing the ratio of the data and Monte Carlo efficiencies computed for a working point of 0.1, and 0.9 are displayed respectively in Figures A.3 and A.4.

A.2 Comparison with Other Studies

Comparison with the performance resulting from the use of D^* sample, with a working point of 0.1, and 0.9 are displayed in Figures A.5. Comparison with the performance resulting from the use of K_s sample, with a working point of 0.1, and 0.9 are displayed in Figures A.6. Comparison between the D^* and K_s samples, with a working point of 0.1, and 0.9 are displayed in Figures A.7.

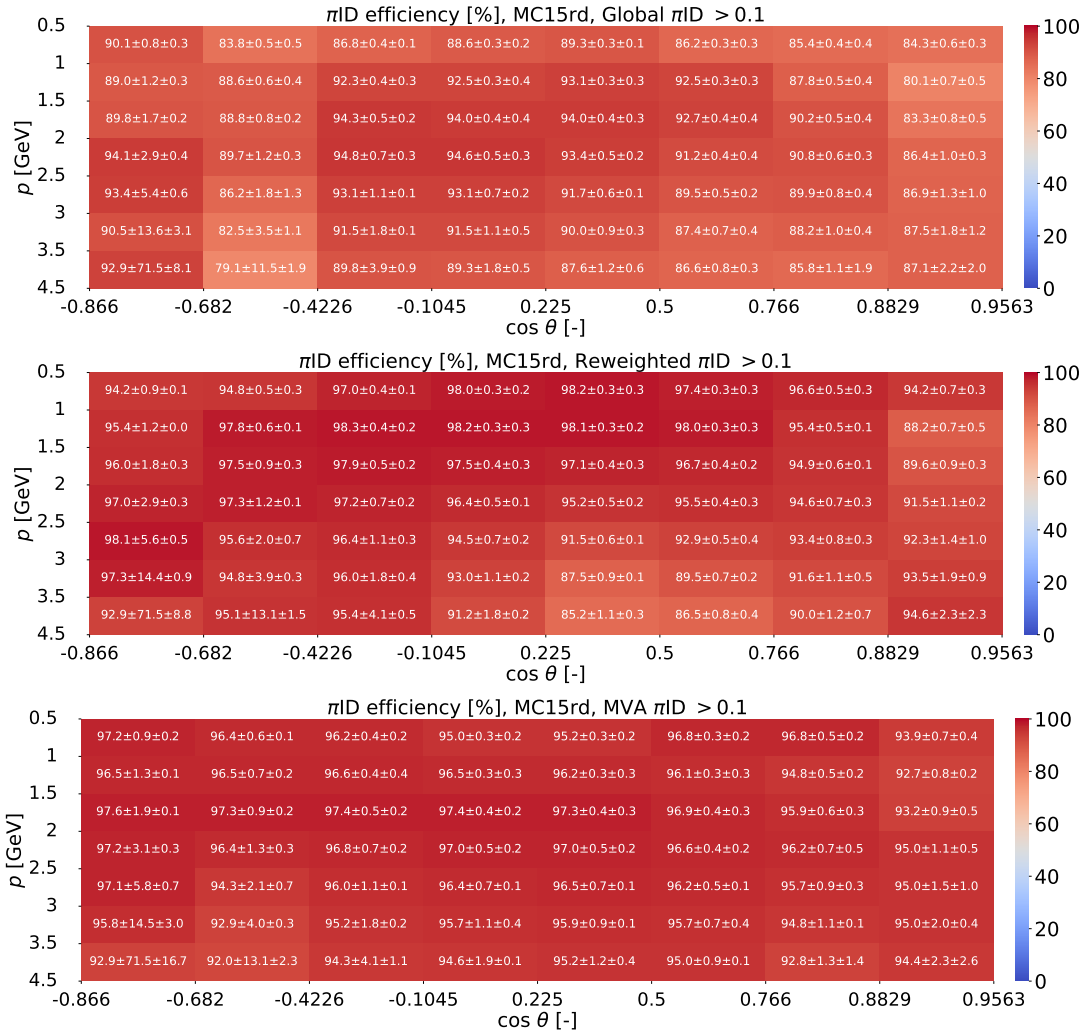


Figure A.1: *Monte Carlo efficiencies.* The PID variables used to get the correction factors binned in momentum and polar angle correspond to the Global (top), the Reweighted (middle) and the MVA (bottom) variables. A working point of 0.1 is selected.

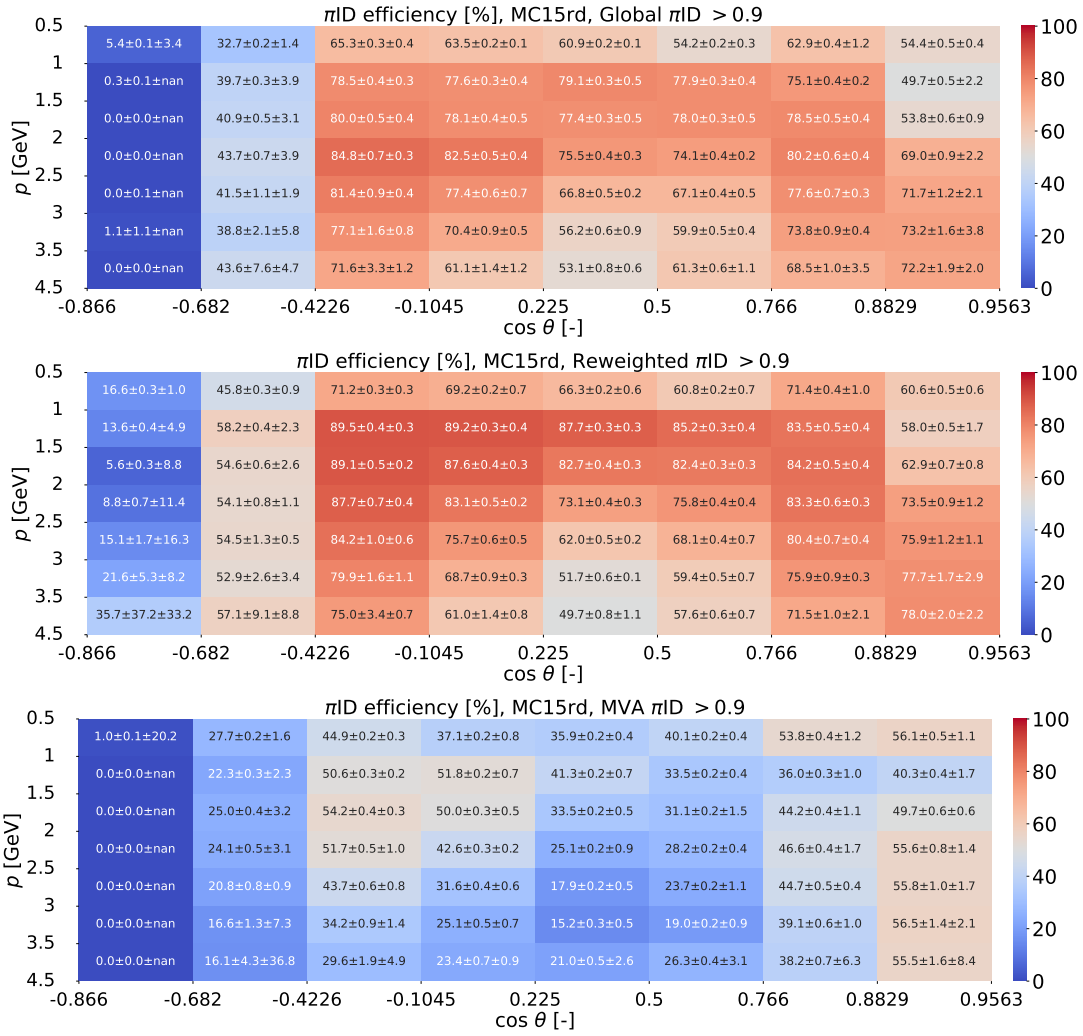


Figure A.2: *Monte Carlo efficiencies.* The PID variables used to get the correction factors binned in momentum and polar angle correspond to the Global (top), the Reweighted (middle) and the MVA (bottom) variables. A working point of 0.9 is selected.

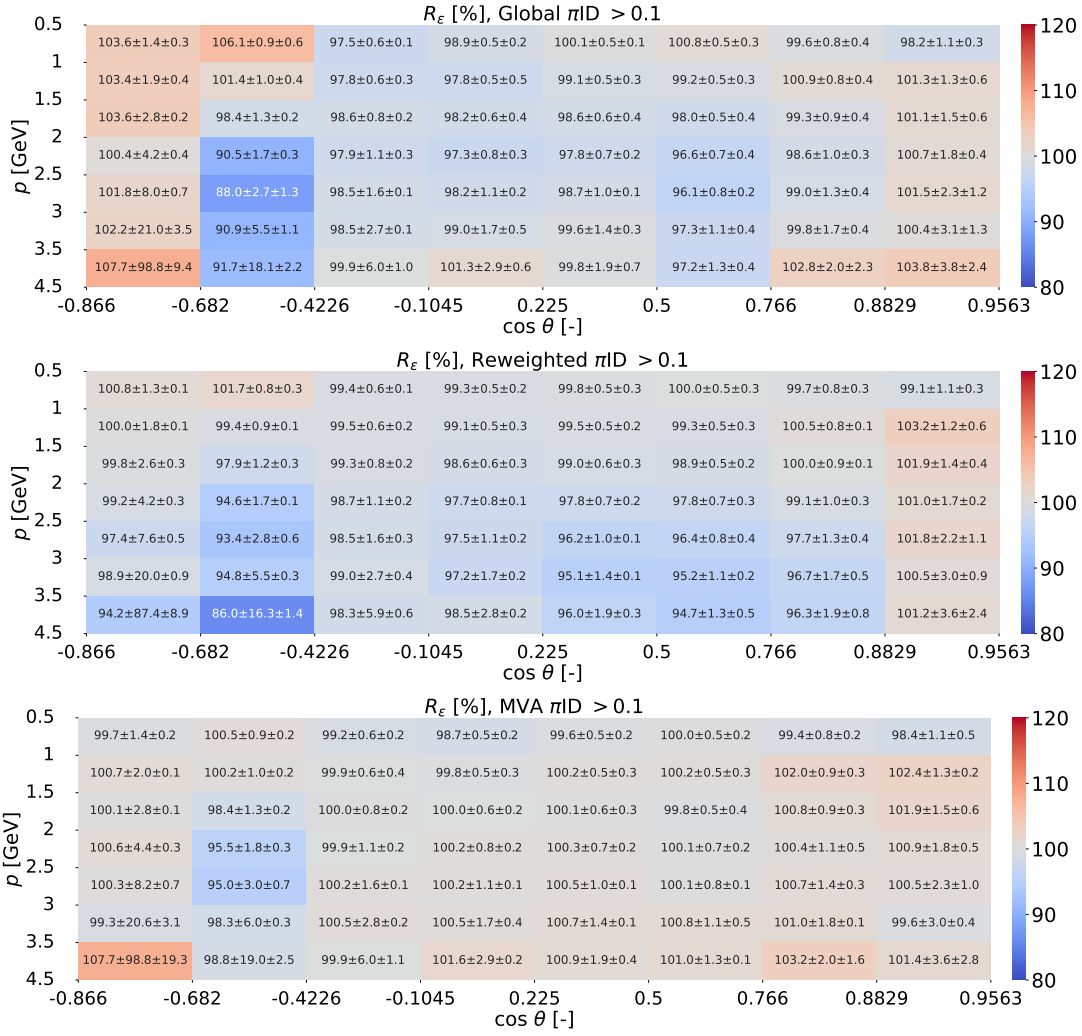


Figure A.3: *Efficiency correction factors.* The PID variables used to get the correction factors binned in momentum and polar angle correspond to the Global (top), the Reweighted (middle) and the MVA (bottom) variables. The correction factors are multiplied by 100 in order to ease their readability. A working point of 0.1 is selected.

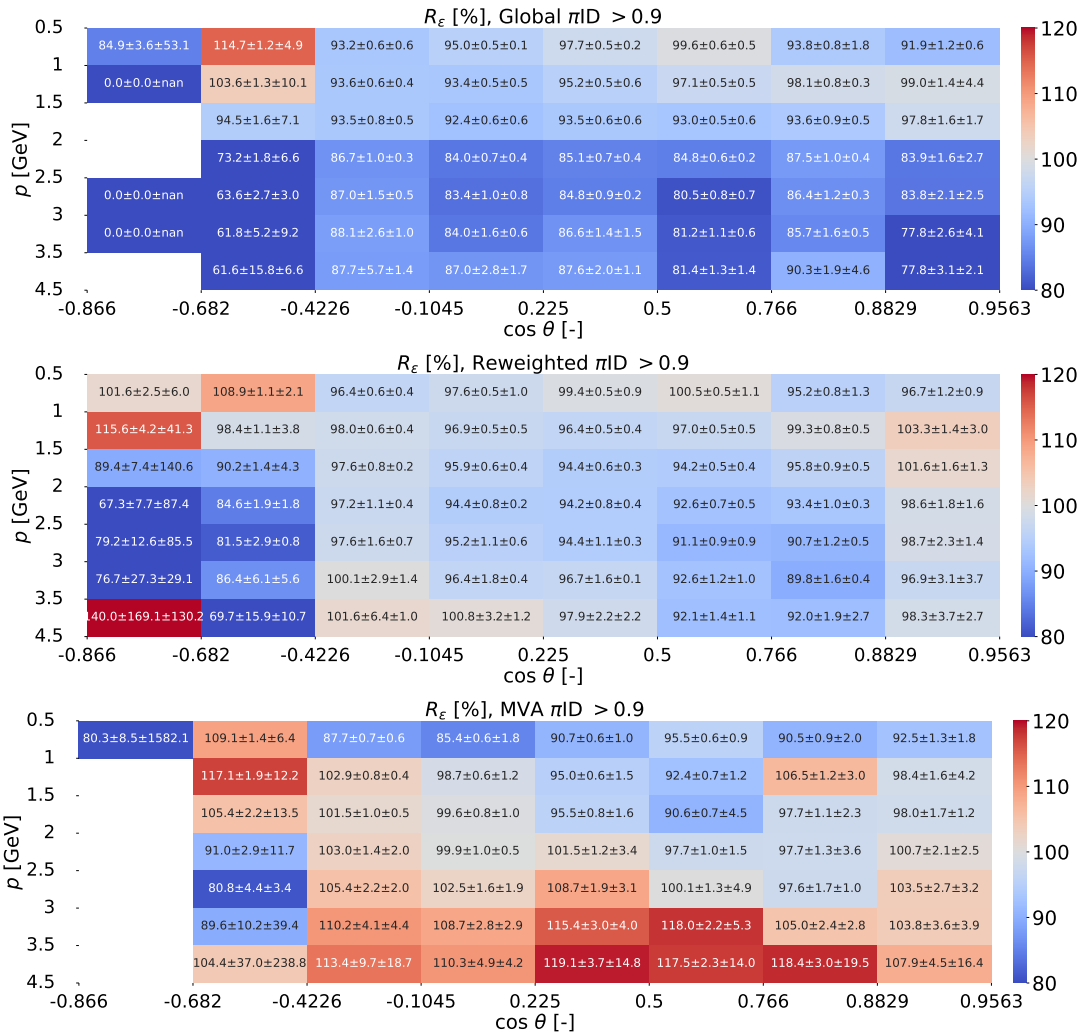


Figure A.4: *Efficiency correction factors.* The PID variables used to get the correction factors binned in momentum and polar angle correspond to the Global (top), the Reweighted (middle) and the MVA (bottom) variables. The correction factors are multiplied by 100 in order to ease their readability. A working point of 0.9 is selected.

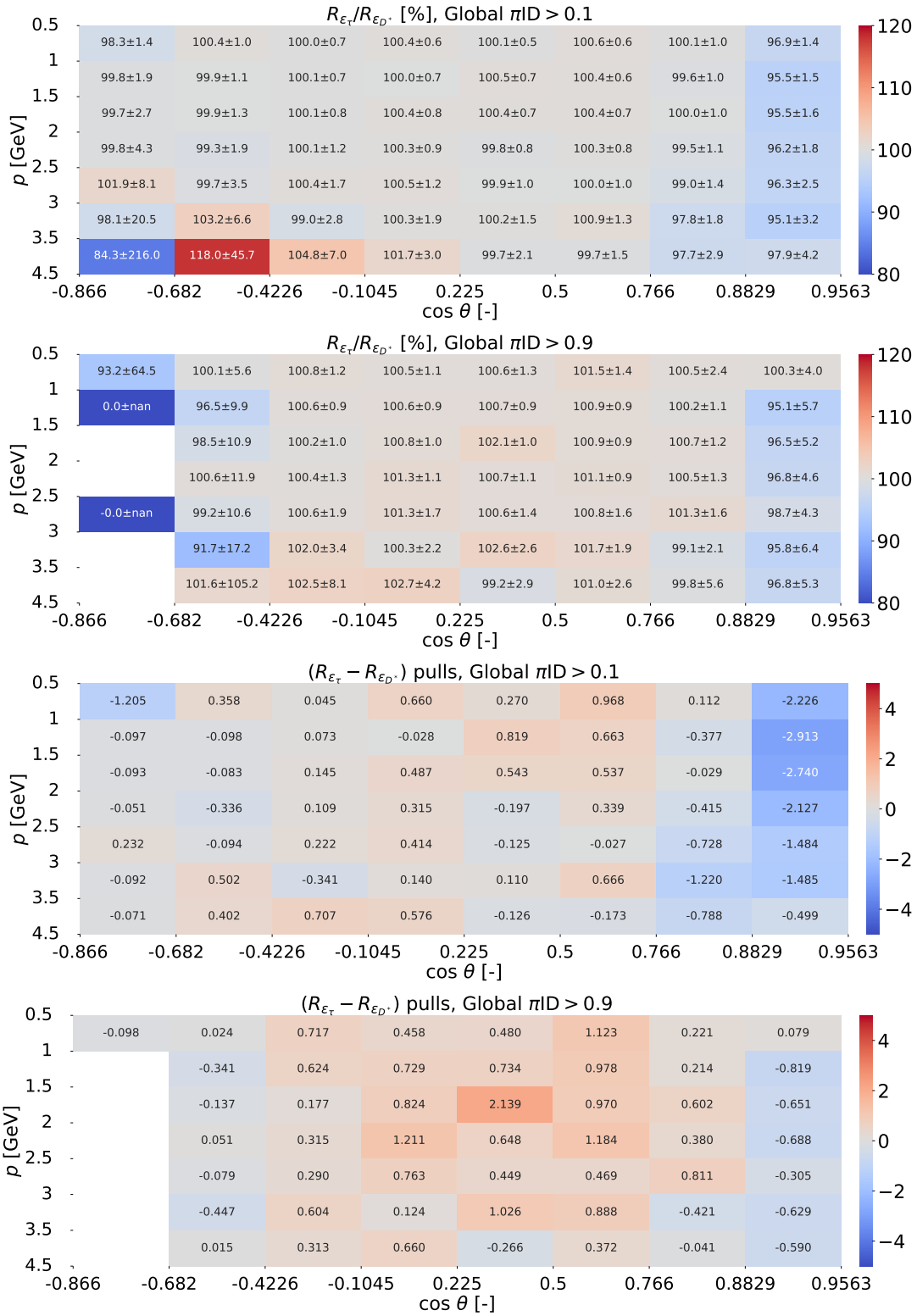


Figure A.5: *Comparison with the D^* sample.* The ratio of correction factors (top part), as well as the associated pulls (bottom part) are displayed. A working point of 0.1 (top in each part) and 0.9 (bottom in each part) are selected.

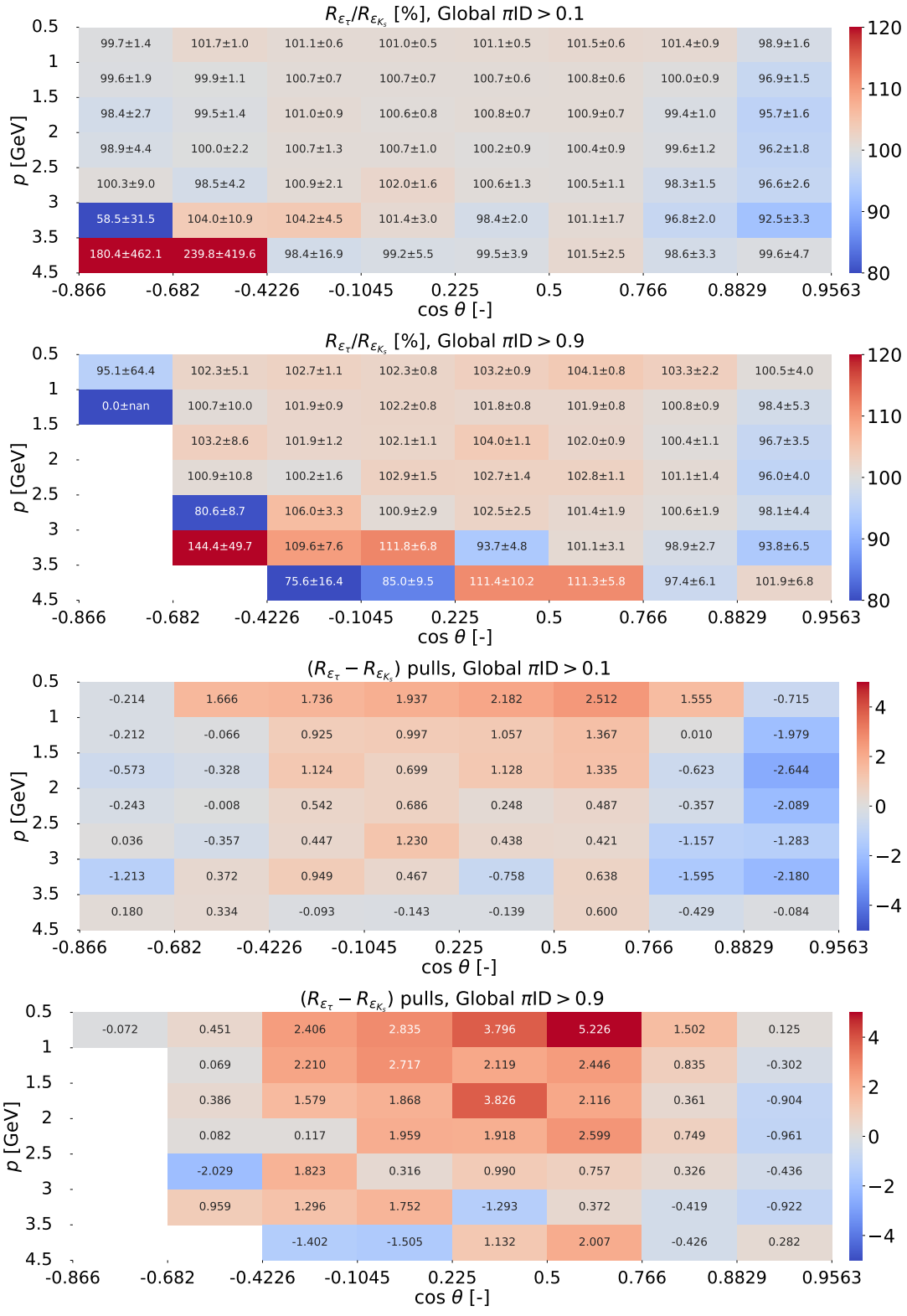


Figure A.6: *Comparison with the K_s sample.* The ratio of correction factors (top part), as well as the associated pulls (bottom part) are displayed. A working point of 0.1 (top in each part) and 0.9 (bottom in each part) are selected.

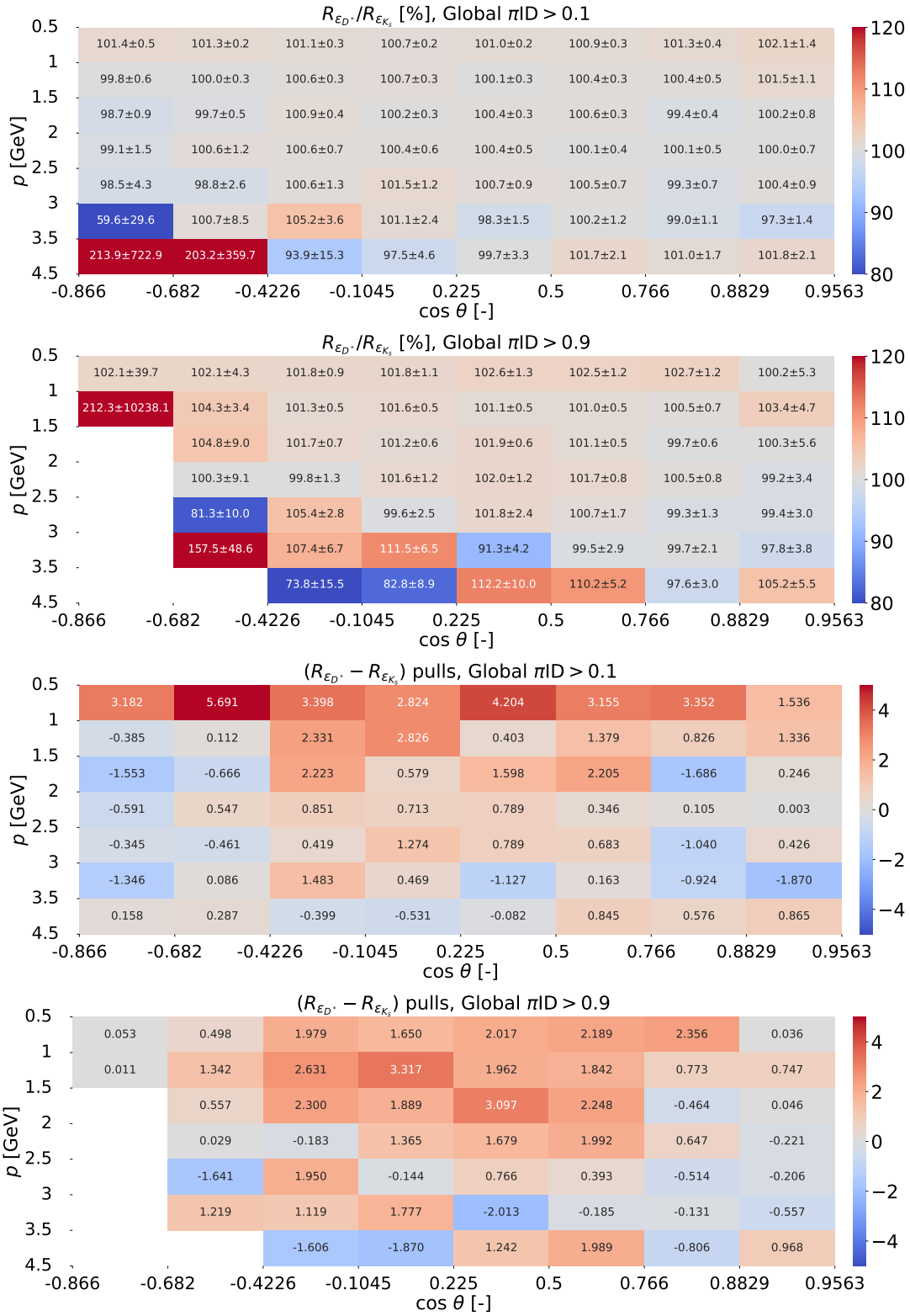


Figure A.7: *Comparison between the D^* and K_s samples.* The ratio of correction factors (top part), as well as the associated pulls (bottom part) are displayed. A working point of 0.1 (top in each part) and 0.9 (bottom in each part) are selected.

B Physics Analysis: Additional material

Additional material related to Chapter 4 is shown in what follows. The section naming convention used here is such that it repeats the relevant sections from the original chapter.

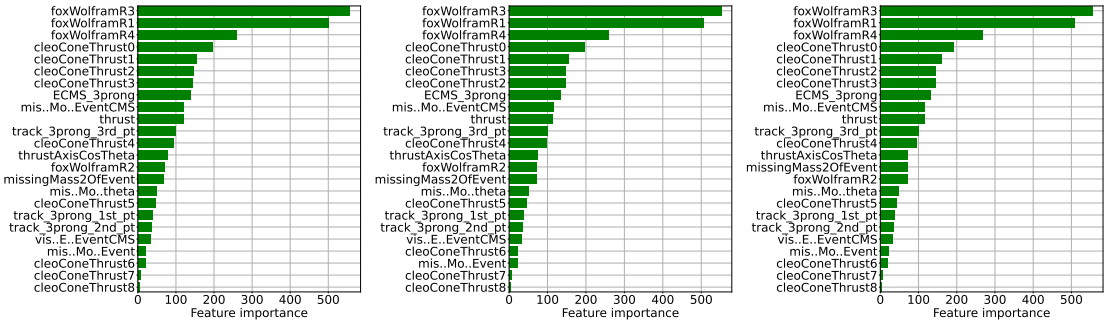
B.1 Event Selection

Figures B.1 displays the feature rankings obtained for various hID selections, and complements the result given in Figure 4.16. Generally, the ranking of these features remains highly consistent across different configurations.

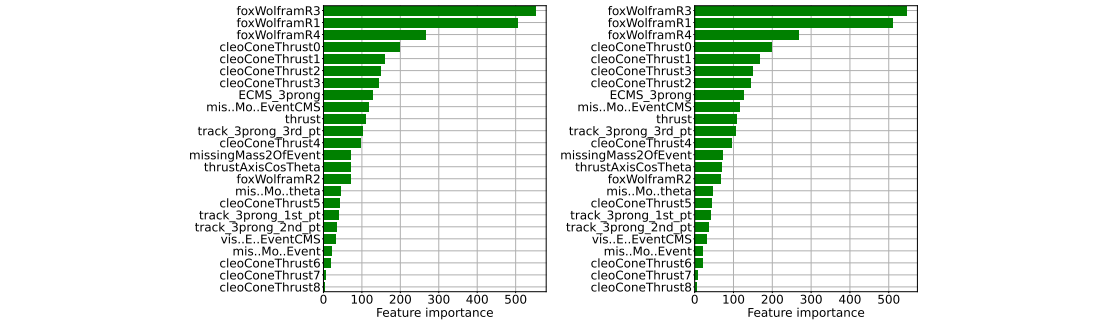
B.2 Measurement of $|V_{us}|$

As demonstrated in Section 4.4.5.2, the Reweighted hID with a working point of 0.9 yields the smallest propagated uncertainty on $R_{K/\pi}$. Figure 4.30 displays the resulting uncertainty on $R_{K/\pi}$ computed considering the statistical, luminosity and hID uncertainties. Figures B.2, B.5, B.4 and B.3 give further information. As in Figure 4.30, the hID selection criteria are provided at the end of each section of the plot, with the variations obtained within this selection corresponding to different BDT output working points. The minimum value is indicated in red.

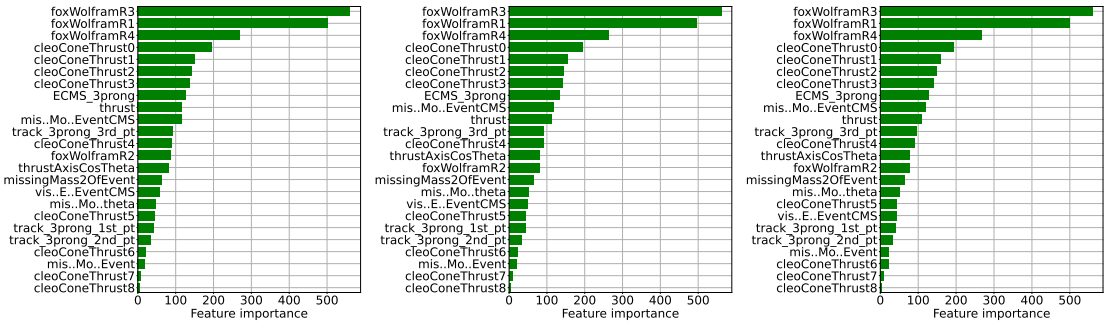
We provide in Figures B.6 the systematic uncertainties assessed with all current sources implemented, computed for different BDT output working points than that discussed in Section 4.4.



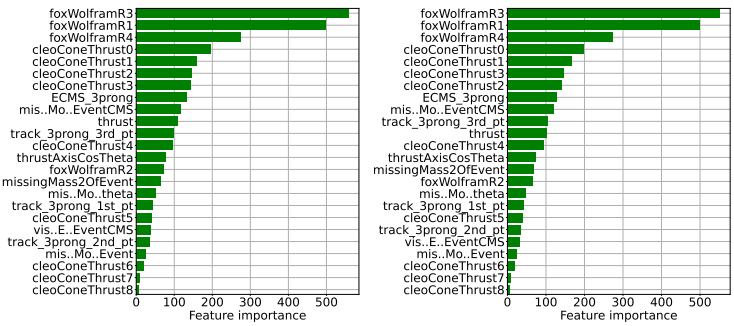
(a) Global 0.5 (b) Global 0.6 (c) Global 0.7



(d) Global 0.8 (e) Global 0.9



(f) Reweighted 0.5 (g) Reweighted 0.6 (h) Reweighted 0.7



(i) Reweighted 0.8 (j) Reweighted 0.9

Figure B.1: *Feature importance.* Computed using a k-fold technique and repetitions. Several hID selections using the Global and Reweighted hID variables were employed.

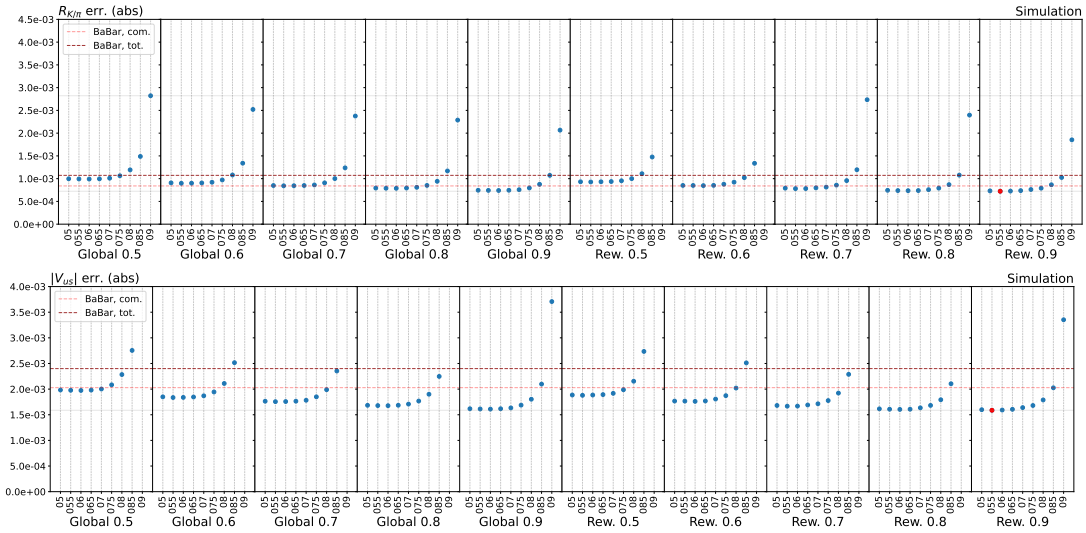


Figure B.2: *Propagation of uncertainties on $R_{K/\pi}$ and $|V_{us}|$ using different hID selections.* Each hID selection is decomposed into several BDT output working points. Two limits provided by the BaBar collaboration [33] are shown: a comparable limit that takes into account the hID and statistical uncertainties, and their total uncertainties. The red dot corresponds to the minimum value achieved across all selections. The grey lines indicate the position of the minimum and maximum values.

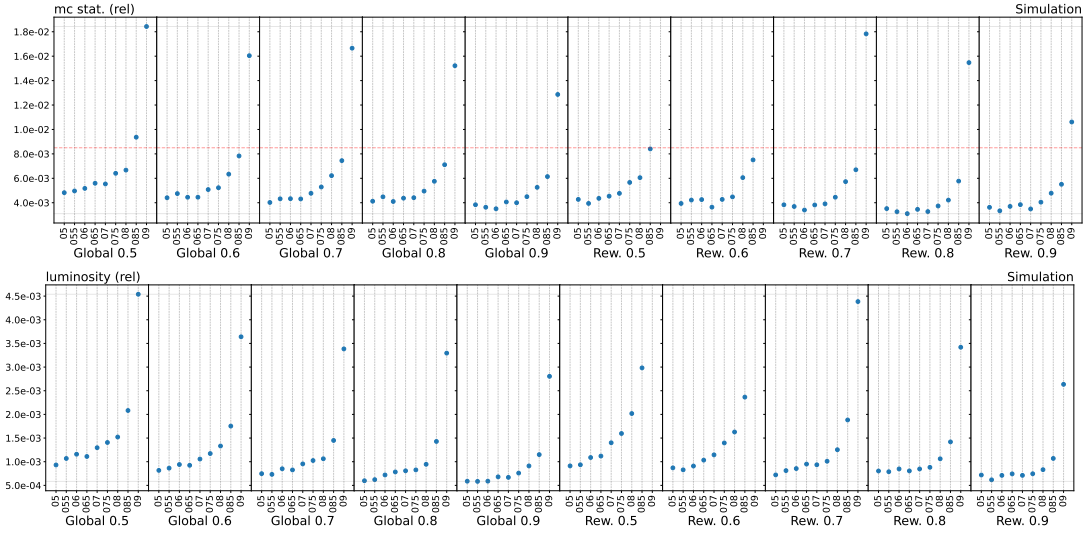


Figure B.3: *Other uncertainties resulting from the propagation of errors on $R_{K/\pi}$ using different hID selections.* Each hID selection is decomposed into several BDT output working points. The limit from the BaBar collaboration [33] is indicated for the statistical uncertainty.

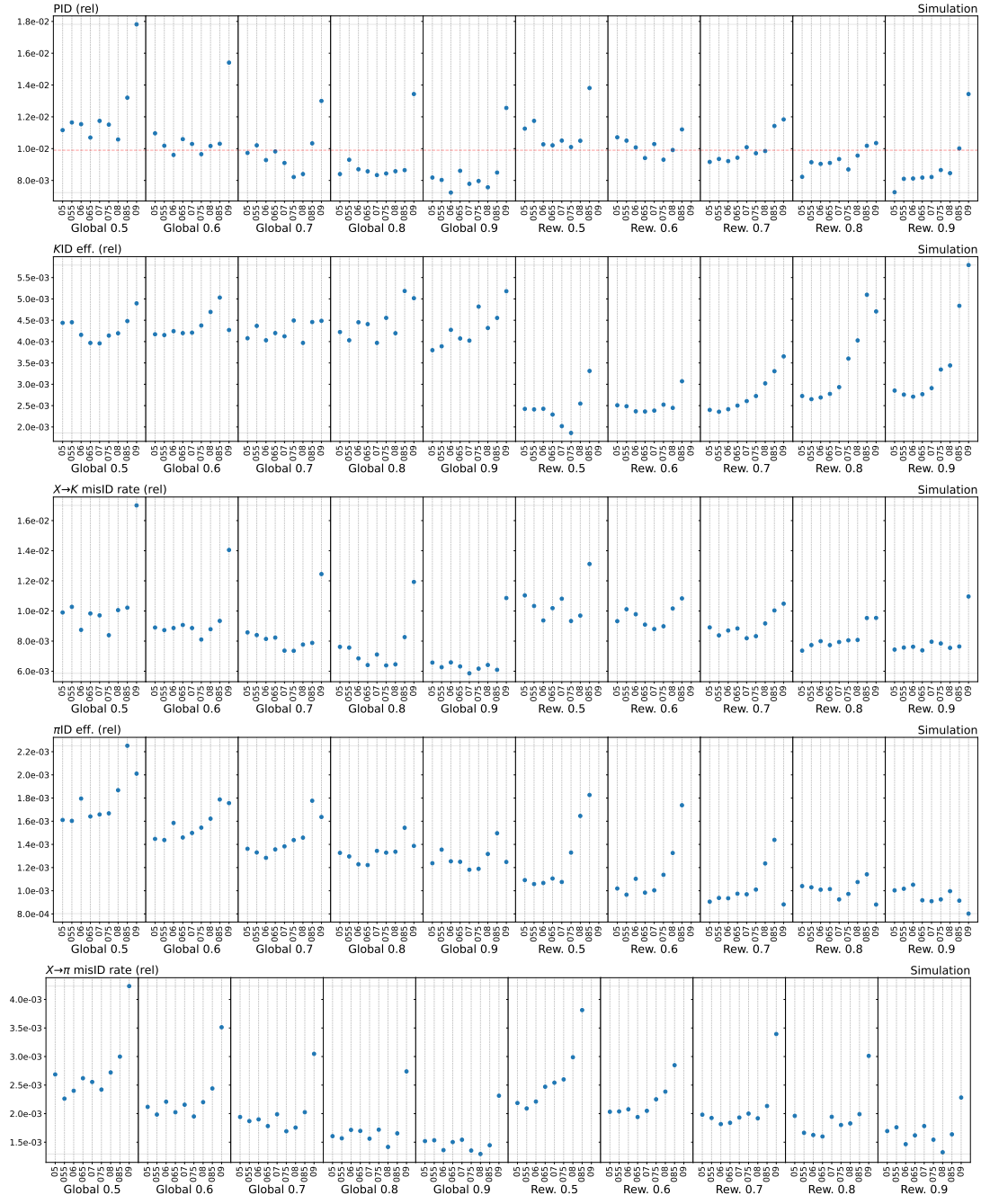


Figure B.4: *hID uncertainties resulting from the propagation of errors on $R_{K/\pi}$ using different hID selections.* Each hID selection is decomposed into several BDT output working points. The limit from the BaBar collaboration [33] is indicated for the combined error.

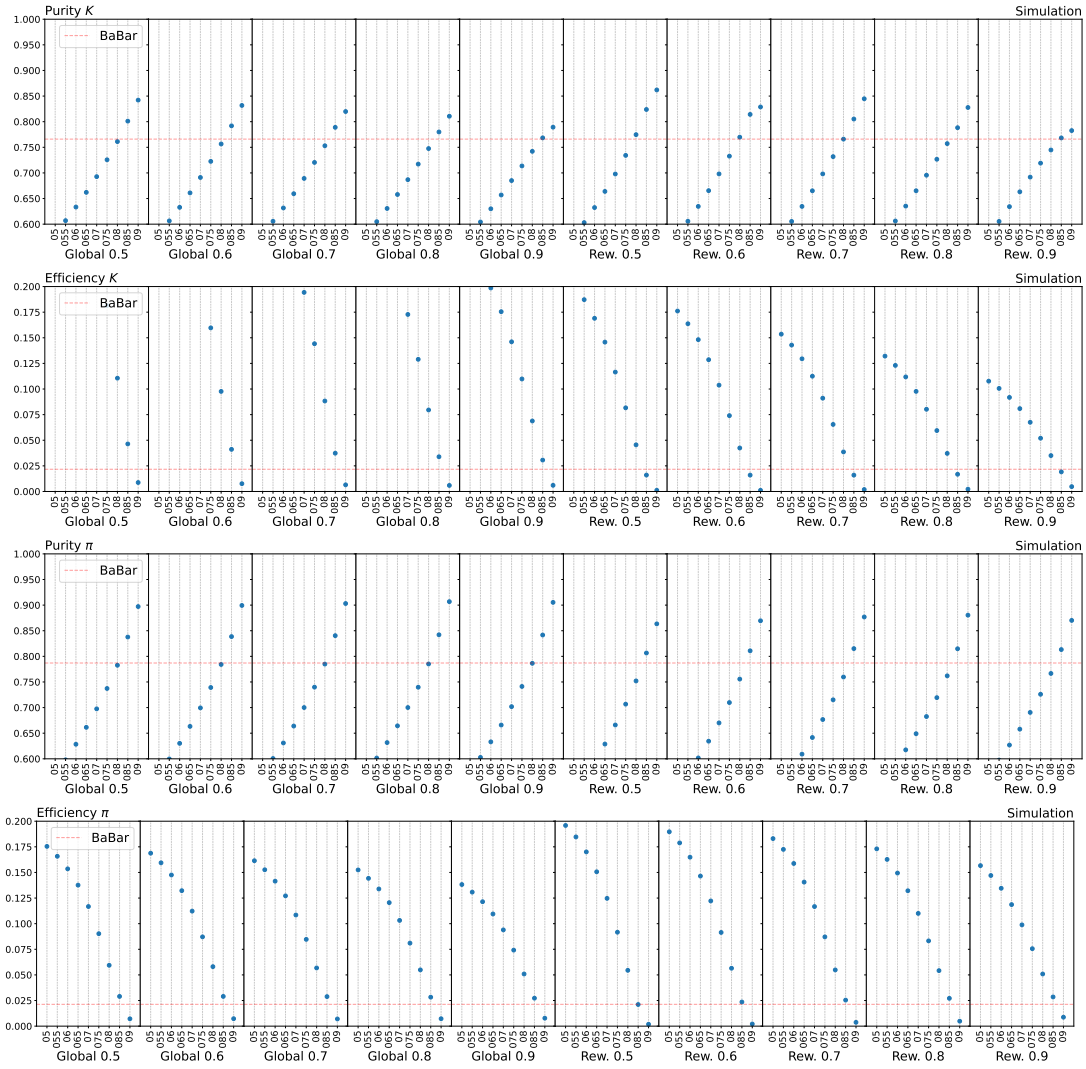


Figure B.5: *Purities and efficiencies resulting from the propagation of errors on $R_{K/\pi}$ using different hID selections.* Each hID selection is decomposed into several BDT output working points. The limit from the BaBar collaboration [33] is indicated.

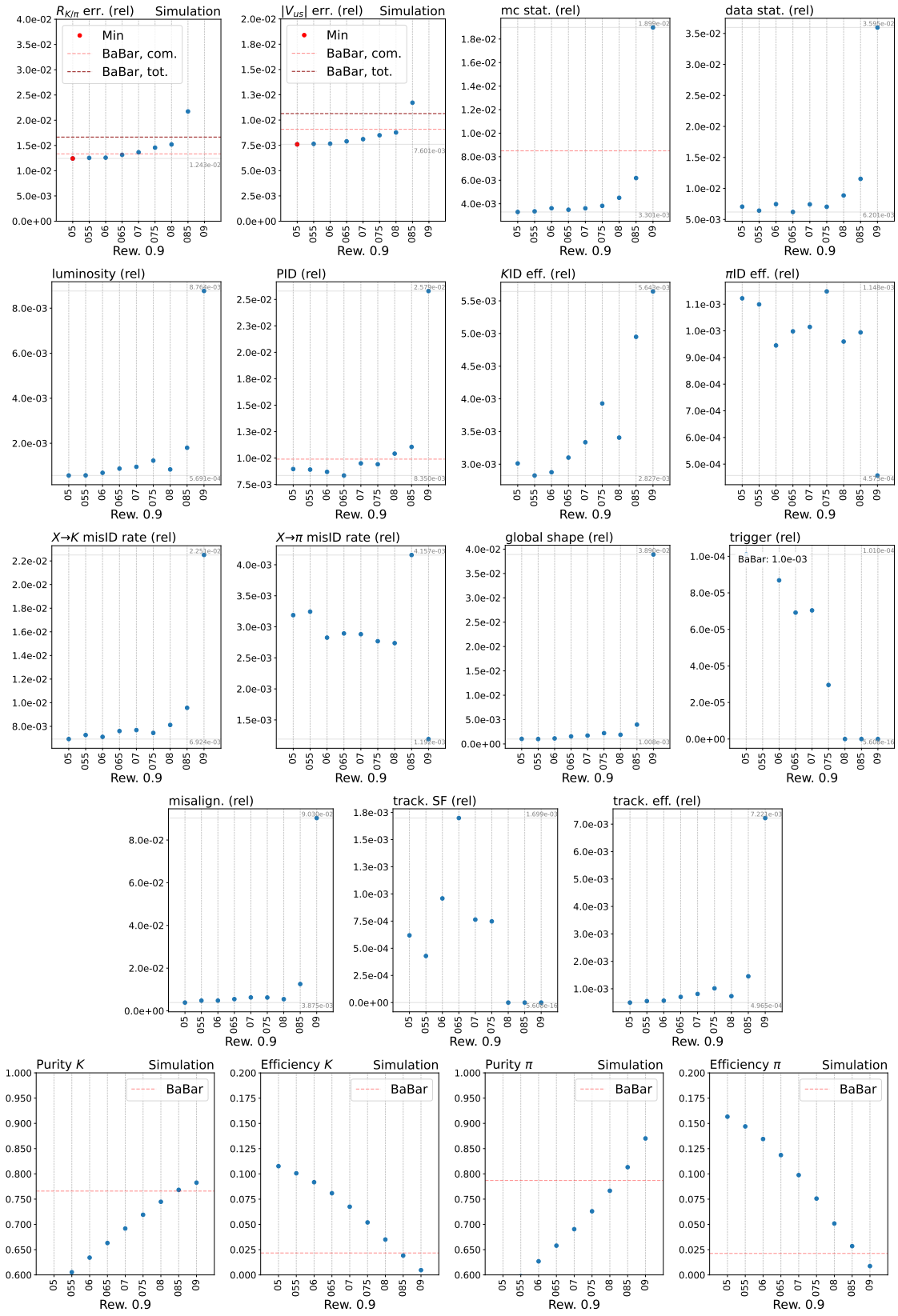


Figure B.6: *Systematic uncertainties assessed with all current sources implemented in the model.* Several BDT output working points are used for the hID selection using the Reweighted hID variable with a working point of 0.9.

C Papers and Additional Activities

Additional material related to additional activities activities carried out during the PhD is shown in the following sections.

C.1 Z' Analysis

As mentioned in Section 1.1.2, rather straightforward possible extensions of the Standard Model involve the addition of an extra $U(1)$ gauge group, which in turns gives rise to an additional gauge boson commonly denoted as Z'. One extension in particular, known as the $L_\mu - L_\tau$ extension, allows the Z' to couple to the Standard Model only through μ, τ, ν_μ and ν_τ with coupling g', and offers explanations for anomalies such as the anomalous magnetic moment of the muon [13]. The search for the Z' has been performed in the process $e^+e^- \to \mu^+\mu^-(Z' \to \text{inv.})$, where the Z' is radiated off a muon, as illustrated in Figure C.1.

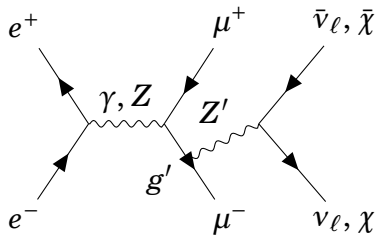


Figure C.1: *Feynman diagram of the process* $e^+e^- \to \mu^+\mu^-(Z' \to inv.)$. The invisible final states allowed are composed of either ν_ℓ for $\ell = e, \mu$ or χ.

Two different scenarios are considered: in the first one, called "*Vanilla* $L_\mu - L_\tau$ *model*", the invisible decay happens only through Standard Model particles, which are in this case a neutrino pair, whereas in the second scenario, referred to as "*Fully invisible* $L_\mu - L_\tau$ *model*", the Z' decays into a pair of dark matter particles $\chi\bar{\chi}$. Samples of e^+e^- collisions collected by Belle II at the CMS energy of the $\Upsilon(4S)$ resonance (10.58 GeV) between 2019 and 2020 has been used, and correspond to a total integrated luminosity of 79.1 fb^{-1}. Results are provided for each of the two scenarios for $M_{Z'} < 9$ GeV/c^2, and are reported in [15], whose abstract is displayed in Figure C.2.

In the analysis, my task involved, on top of following discussions, creating an entry in HEPData and filling it with all the relevant information from the published paper. This effort aimed to make the data easily accessible to any analyst interested in consulting it. The entry can be consulted in [100].

PHYSICAL REVIEW LETTERS **130,** 231801 (2023)

Search for an Invisible Z' in a Final State with Two Muons and Missing Energy at Belle II

(Belle II Collaboration)

(Received 7 December 2022; accepted 19 April 2023; published 7 June 2023)

The $L_\mu - L_\tau$ extension of the standard model predicts the existence of a lepton-flavor-universality-violating Z' boson that couples only to the heavier lepton families. We search for such a Z' through its invisible decay in the process $e^+e^- \to \mu^+\mu^- Z'$. We use a sample of electron-positron collisions at a center-of-mass energy of 10.58 GeV collected by the Belle II experiment in 2019–2020, corresponding to an integrated luminosity of 79.7 fb^{-1}. We find no excess over the expected standard-model background. We set 90%-confidence-level upper limits on the cross section for this process as well as on the coupling of the model, which ranges from 3×10^{-3} at low Z' masses to 1 at Z' masses of 8 GeV/c^2.

DOI: 10.1103/PhysRevLett.130.231801

Figure C.2: *Abstract of Belle II second search for an invisible Z' particle paper.* The full paper is available in [15].

C.2 ECFA ECR activities

The European Committee for Future Accelerators (ECFA) Early Career Researcher (ECR) Panel discusses key aspects impacting the future of particle physics and advises ECFA through regular reports and annual presentations. Panel members, typically PhD students or postdocs within 8 years of their PhD, are nominated by ECFA countries and major laboratories for 2-year terms. They represent the views of early-career researchers and appoint observers for ECFA meetings. The panel organizes meetings, working groups, and studies to support the early-career physics community within ECFA countries. The ECFA ECR Panel activities between 2021 and 2022 are reported in [101], whose abstract is provided in Figure C.3.

In this panel, my role involved collaborating with a team to study the career outlook and experiences of ECRs, as well as investigating diversity and sociological aspects within particle physics research. In particular, results of a survey distributed to ECRs in academic particle physics between September 2022 and March 2023 have been studied, and are presented in [102], whose abstract is made available in Figure C.4.

The ECFA Early Career Researcher's Panel: composition, structure, and activities, 2021 – 2022

The ECFA Early Career Researcher's (ECR) Panel

December 22, 2022

The European Committee for Future Accelerators (ECFA) Early Career Researcher's (ECR) panel, which represents the interests of the ECR community to ECFA, officially began its activities in January 2021. In the first two years, the panel has defined its own internal structure, responded to ECFA requests for feedback, and launched its own initiatives to better understand and support the diverse interests of early career researchers. This report summarises the panel composition and structure, as well as the different activities the panel has been involved with during the first two years of its existence.

Figure C.3: *Abstract of ECFA ECR Panel description of activities paper.* The full paper is available in [101].

Results of the 2022 ECFA Early-Career Researchers Panel survey on career prospects and diversity

The ECFA Early-Career Researchers Panel: Career Prospects and Diversity in Physics Programmes Working Groups

April 3, 2024

This document presents the outcomes of a comprehensive survey conducted among early career researchers (ECRs) in academic particle physics. Running from September 24, 2022, to March 3, 2023, the survey gathered responses from 759 ECRs employed in 39 countries. The study aimed to gain insights into the career prospects and experiences of ECRs while also delving into diversity and sociological aspects within particle physics research. The survey results are presented in a manner consistent with the survey choices. The document offers insights for the particle physics community, and provides a set of recommendations for enhancing career prospects, fostering diversity, and addressing sociological dimensions within this field.

Figure C.4: *Abstract of ECFA ECR Panel analysis of survey on career prospects and diversity paper.* The full paper is available in [102].