Technical University of Munich
School of Natural Sciences

Max Planck Gesellschaft
Max Planck Institut für Physik
Belle II Group

Bachelor's Thesis

Study of $_s\mathcal{P}$lot Background Subtraction for Particle Identification Efficiency Determination at Belle II

Lars Rasmus von der Werth
September 11, 2024

Supervisor:	Prof. Dr. Stephan Paul
Second Corrector:	Prof. Dr. Sherry Suyu
Advisor:	Dr. Stefan Wallner

Technical University of Munich

Abstract

Particle identification is a central experimental challenge in the precise measurement of Standard Model parameters and in the search for new physics beyond the standard model at the Belle II experiment. This also includes an accurate knowledge of the imperfections in particle identification efficiencies. This efficiency is typically determined from Monte Carlo simulations, but crosschecking and validating these simulations, requires to determine efficiencies directly from the real data.

In this thesis, we present the determination of muon identification efficiencies using $J/\Psi \to \mu^+\mu^-$ decays. To remove background events, that remain after event selection, we do a background subtraction using the $_s\mathcal{P}$lot method. In the existing implementation of the efficiency calculation, we observe in simulated $J/\Psi \to \mu^+\mu^-$ decays, that the efficiencies computed using background subtraction disagree with the Monte Carlo truth efficiencies in the momentum range below 1.5 GeV/c. The $_s\mathcal{P}$lot method can be impaired by correlations between the kinematic variables of the sample. Therefore, we investigate the effect of correlations to understand and mitigate a possible bias introduced by them to the determination of particle identification efficiencies.

We find out, that it is possible to improve the efficiency estimates by taking into account correlations between kinematic variables in the background subtraction. However, these improvements alone are not sufficient to reach agreement with the true efficiencies in the full momentum range. A possible explanation for the remaining discrepancies are observed imperfections of the parametrization of the probability density function used to model the invariant mass distribution of the two particle system, which is needed for the $_s\mathcal{P}$lot method.

Contents

1   Introduction

The Belle II experiment is at the forefront of high precision tests of the Standard Model of particle physics and the search for new physics beyond the Standard Model [1]. The key to that search is knowing the species of a particle, which determines how that particle can interact, and allows us to reconstruct the processes and decays, that have occurred in the experiment. For that we need particle identification (PID), which means assigning the species of a particle to a track measured in the detector.

However the particle identification of a detector is not always perfect, which can lead to errors in the experimental results. These imperfections need to be accounted for in corrections. Therefore it is also very important to test the performance of PID. This is quantified by the particle identification efficiency, which can be tested for each species. The efficiency is the number of particles correctly identified as belonging to a given species, divided by the total number of particles of that species tested [2]. The efficiency depends on the kinematics of a particle, so for example its momentum [1]. These PID efficiencies can either be obtained directly from experimental data or from so called Monte Carlo simulations of the physics process in the experiment and of the detector. However, the Monte Carlo simulations may not perfectly resemble the real physics processes of the experiment. Therefore we need to crosscheck and correct the efficiency results obtained from simulations with those obtained from real data. This also yields valuable information, that can be used to better understand the underlying processes and improve simulations.

To obtain PID efficiencies directly from real data, we need to employ analysis methods on a known decay process. In this analysis we will study the muon identification efficiency of Belle II. To obtain efficiencies from real data, we are first required to obtain a pure sample, that only consists of muons.

We need to do this on a decay, where the species of the decay daughter particles are known. For this study we use the decay mode $J/\Psi \to \mu^+\mu^-$. However, since the J/Ψ can also decay into pairs of particles, that are not muons, we need to use the tag and probe approach. Identifying the species of one of the daughter particles, which is called the tag particle, implies the species of the other so called probe particle. This means we can first use PID to identify the tag particle and, given that both particles come from a J/Ψ decay, we can use the probe particle to determine the muon identification efficiency. However, since not all events, where one particle is identified as a muon stem from a J/Ψ decay, we also need to do background subtraction. For this we use the $_s\mathcal{P}$lot method [3]. This method uses the known distribution functions of signal and background in one so called discrimination variable to assign weights to each event. These are called sWeights [4]. As discrimination variable for the $_s\mathcal{P}$lot method we use the invariant mass $m_{\mu\mu}$ of the two particle system, in which the distribution of signal and background is very distinctive. The sWeights can be used to reconstruct

histograms of only signal events in kinematic variables, that are uncorrelated with the discrimination variable $m_{\mu\mu}$. From these histograms we can then finally compute the PID efficiencies in kinematic variable bins.

This approach is already implemented in the current analysis framework, which is called the systematic corrections framework [3, 5]. To develop and test the analysis methods, we can again use the Monte Carlo simulation, where the true efficiencies are known. In the muon identification efficiencies computed using the systematic corrections framework, we observe a mismatch between the efficiencies, which are computed using the $_s\mathcal{P}$lot method, and the efficiencies computed using Monte Carlo truth information. In particular in the momentum range below 1.5 GeV/c the two efficiencies do not agree. Here the muon identification efficiencies computed using $_s\mathcal{P}$lot background subtraction are wrong for the simulation and probably also for real data.

In this thesis, we will study the $_s\mathcal{P}$lot method for background subtraction for the J/$\Psi \to \mu^+\mu^-$ sample. We will try to find the reasons for the muon identification efficiencies using the $_s\mathcal{P}$lot background subtraction method. We will investigate correlations between the event distributions in kinematic variables of the two muons to understand their effect on the sWeight background subtraction and efficiency calculation. In the end, we want to understand and mitigate the mismatch between Monte Carlo and sWeight efficiencies, to be able to obtain valid PID efficiencies directly from data.

In section 2, we will briefly introduce the Standard Model and the particles relevant to understand PID at Belle II, as well as introduce some concepts of collider experiments. In section 3, an introduction into the Belle II experiment will be given. The methods used in this analysis, including the $_s\mathcal{P}$lot method, will be discussed in section 4. In section 5, the data set, that was used for the analysis, will be introduced and initial cuts as well as the probability density function model will be explained. We will discuss the results of different approaches to achieve better agreement of the muon identification efficiencies in section 6. Finally, we will come to a conclusion and outlook in section 7.

2 Standard Model and Concepts of Collider Experiments

In this section, general terms and concepts necessary for understanding the following sections will be introduced. Therefore the Standard Model will be briefly introduced, to give a definition for the necessary particles and to understand why we need particle identification. After that some concepts of particle collider experiments will be introduced.

2.1 Standard Model

The Standard Model is our best theory to describe the phenomena of three out of the four known fundamental forces, being the electromagnetic, weak and strong force. It lays a theoretical framework for the processes observed in particle physics experiments. It has made multiple predictions that were later proved true, some of the most prominent being the existence of the top quark and the Higgs boson [6].

The Standard Model is a quantum field theory. It describes physics using fields, whose excitations are quantized and occur as particles. These fundamental quantum fields, the Standard Model deals with, can either be scalar fields or spinor fields. The excitation of a spinor fields resembles a fermion, a particle with a half integer spin. These are the particles most of the matter we know is made up of. Each of those particles has a corresponding anti particle with opposite charge. The excitation of a scalar fields resembles a boson, a particle with an integer spin [6, 7]. The bosons act as exchange particles of the forces of the standard model. The eight gluons are the exchange particles of the strong interaction, the weak interaction is carried by the $W^\pm$ and Z^0 and for the electromagnetic interaction the exchange particle is the photon γ. The coupling between an exchange boson and a type of fermion decides, if that fermion takes part in the respective force conveyed by the boson. This offers a way to group the fermions of the standard model. A common distinction is that between quarks and leptons. Those fermions that couple with gluons, i.e. interact via the strong interaction, are called quarks. Those fermions that do not couple to gluons are called leptons. The Standard Model includes three generations of both quarks and leptons [6, 8].

An important aspect of a theory is its behavior under discrete symmetry. For that it is considered, how the physics of a theory changes, if a symmetry operation is applied. For the Standard Model there are three discrete symmetry operations whose effect can be investigated. The charge conjugation C replaces particles with their antiparticles. The parity P flips the sign of all spacial coordinates, whereas the time reversal symmetry T flips the sign of the time coordinate. Also combinations of these symmetry operations like CP can be considered. If the physics changes under a symmetry operation, this is called a violation of the symmetry. Measuring, if a symmetry holds or how much it is violated grants valuable insights to better understand the processes of nature [6].

2.1.1 Leptons

For each of the three generations of leptons there exist two types of particles, which are referred to as being from the same lepton family. These are the electromagnetically charged lepton which carries an electromagnetic charge $Q = -e$ and the charge less neutrino with $Q = 0$. The leptons are called electron e and electron neutrino ν_e for the first generation, muon μ and muon neutrino ν_μ for the second generation and tau τ and tau neutrino ν_τ for the third generation [6, 8]. The charges as well as the currently known masses in the Standard Model of all these leptons are listed in table 1.

Table 1: List of leptons including the current world average for their mass m [9] as well as their quantum numbers in the Standard Model.

Lepton	m [MeV/c^2]	Q [e]
ν_e	0	0
e	$0.51099895000 \pm 0.00000000015$	-1
ν_μ	0	0
μ	$105.6583755 \pm 0.0000023$	-1
ν_τ	0	0
τ	1776.93 ± 0.09	-1

On interaction with a $W^\pm$ boson a left handed charged lepton turns into its corresponding neutrino and vice versa. No transitions between a charged lepton and a neutrino of a different family have yet been observed. These transitions are therefore not included in the Standard Model. This is known as the lepton family number conservation [8].

2.1.2 Quarks

There exist two differently charged quarks for each generation. Up type quarks carry an electromagnetic charge of $+2/3$ e while down type quarks have an electromagnetic charge of -1/3 e. The six resulting quarks are called up (u) and down (d) for the first generation, charm (c) and strange (s) for the second generation and top (t) and bottom (b) for the third generation [6, 8].

The defining property of quarks is that they carry a color charge and therefore interact via the strong interaction. There exist three colors and three anti colors. Each quark can have each color and each antiquark each anticolor. The potential of the strong interaction and thus the energy needed to separate two quarks rises approximately linearly with the distance. This so called confinement results in very high energies necessary to separate two quarks bound by the strong interaction. This is the reason why quarks never appear in a state that externally carries a color charge, but always in neutral colored bound states [8].

2.1.3 Hadrons

In general the bound states of quarks via the strong interaction are called hadrons. The strong interaction bound state of a quark that carries a color and an anti quark that carries the corresponding anti color is called a meson. Some important mesons for this analysis are listed in table 2. The neutral kaon, which consists of a down and a strange quark, appears in states that are mixed from the pure quark states d$\bar{\text{s}}$ and s$\bar{\text{d}}$. The two mixed states have different lifetimes and are therefore called K_S for the short living kaon and K_L for the long living one. Furthermore mesons have a ground state and can be exited. This is denoted similar to the hydrogen atom with the quantum numbers of the state [6, 8].

Table 2: List of important mesons including their quark content and the current world average for their mass m [9].

Meson name	Symbol	Quark content	Q [e]	m [MeV/c^2]
Charged Pion	π^+	u$\bar{\text{d}}$	1	139.57039 $\pm$0.00018
	π^-	d$\bar{\text{u}}$	-1	
Neutral Pion	π^0	$\frac{\text{u}\bar{\text{u}}+\text{d}\bar{\text{d}}}{\sqrt{2}}$	0	134.9768 $\pm$0.0005
Charged Kaon	K^+	u$\bar{\text{s}}$	1	493.677 $\pm$ 0.016
	K^-	s$\bar{\text{u}}$	-1	
Neutral Kaon	K_S	$\frac{\text{d}\bar{\text{s}}-\text{s}\bar{\text{d}}}{\sqrt{2}}$	0	497.611 $\pm$ 0.013
	K_L	$\frac{\text{d}\bar{\text{s}}+\text{s}\bar{\text{d}}}{\sqrt{2}}$		
J/Ψ - Meson	J/Ψ	c$\bar{\text{c}}$	0	3096.900 $\pm$0.006
Charged B-meson	B^+	u$\bar{\text{b}}$	1	5279.41 $\pm$ 0.07
	B^-	b$\bar{\text{u}}$	-1	
Neutral B-meson	B^0	$\frac{\text{d}\bar{\text{b}}-\text{b}\bar{\text{d}}}{\sqrt{2}}$	0	5279.72 $\pm$ 0.08
	$\overline{B}^0$	$\frac{\text{d}\bar{\text{b}}+\text{b}\bar{\text{d}}}{\sqrt{2}}$		
Υ(4S) - Meson	Υ(4S)	b$\bar{\text{b}}$	0	10579.4 $\pm$ 1.2

Another kind of strong interaction quark bound state is the baryon, which is a particle that has an odd number of valence quarks. The most well known baryons are the proton p with uud valence quarks and the neutron n with udd valence quarks [8]. Another important particle that can occur in particle collisions in the Belle II experiment is the deuteron, which is the bound state of a proton and a neutron. In table 3 their masses are listed.

Table 3: List of some baryons as well as the deuteron with the current world average of their masses m [9].

Name	Symbol	m
Proton	p	938.27208816 $\pm$ 0.00000029 MeV/c^2
Neutron	n	939.5654205 $\pm$ 0.0000005 MeV/c^2
Deuteron	d	1875.61294500 $\pm$ 0.00000058 MeV/c^2

2.2 Basic Concepts in Particle Collider Experiments

Most of today's leading particle physics experiments are particle colliders. In this section some common terminology of particle colliders will be introduced.

Luminosity

To characterize the performance of a particle accelerator, a useful quantity is the luminosity L. It is a measure for the number of collision events per time. Together with the cross section σ, which is a measure for the probability that a given process may take place, it can be used to calculate the transition rate $\frac{dR}{dt}$ [2]:

$$\frac{dR}{dt} = L \cdot \sigma. \tag{1}$$

To quantify the overall data a collider is able to produce, we can also define the integrated luminosity L_{int}, by integrating over the luminosity over time [2]:

$$L_{int} = \int L(t)dt. \tag{2}$$

Invariant Mass

The center of mass energy or invariant mass E_{cm} is the energy a system of particles has in its rest frame. It can be calculated by contracting the sum of the four-momenta p_i^μ of each particle i with itself. Using Einsteins sum convention this can be written as [6]:

$$E_{cm}^2 = \left(\sum_i p_i^\mu \right) \left(\sum_i p_{i\mu} \right). \tag{3}$$

Since the invariant mass is a Lorentz scalar, it is invariant under Lorentz transformations. This implies, that it can be calculated using equation 3 with the four-momenta in any system. It is especially useful to check if a given set of particles comes from the decay of a single particle. This is possible since the only quantity needed to calculate the invariant mass, the total four momentum, is conserved and therefore the invariant mass itself is conserved. Therefore the center of mass energy of the system of particles has to be identical to the mass of the decaying particle. Actually, because of the uncertainty principle, it is possible for the invariant mass of the particle system to deviate from the exact mass of the decaying particle because of its limited lifetime. The same principle also applies for particles colliding. The produced particle, for which the production cross section at a given invariant mass is maximal, is called a resonance [6].

Particle Identification

Particle identification is the assignment of a hypothesis for a particle species to a measured track. The particle is a real physical object whereas the species in this context is its type [2]. For collider experiments it is often assumed, that only a certain set of species with long enough lifetimes can actually appear in the detector. For charged particles at the Belle II experiment these species are electrons, muons, pions, kaons, protons and deuterons [1].

3   Belle II Experiment

The Belle II experiment is located at the High Energy Accelerator Research Organization (japanese: kō-enerugī kasokuki kenkyū kikō) KEK in Tsukuba, Japan. It is a leading experiment of the high precision frontier of particle physics. That means it aims to accurately measure the quantities of the Standard Model and search for new physics in processes deviating from the Standard Model and forbidden decay processes, in contrast to experiments at the Large Hadron Collider (LHC) at Cern searching for new particles at high energies. The Belle II experiment has been taking data since 2018 and will be up for service for some years to come. This means the experiment is still producing data and is being optimized.

Belle II is designed to generate a large data set of B-mesons. These are produced from the collision of electrons and positrons at the resonance energy of the $\Upsilon(4S)$ resonance [1]. Its energy is sufficient to decay into a state of two neutral or charged B-mesons with a branching fraction of over 96% [9]. The same production mechanism was already used in previous similar experiments, namely Babar at the Stanford Linear Accelerator Center SLAC and Belle IIs direct predecessor Belle. Such experiments producing lots of B-mesons are called B-factories [10].

3.1   Goals of the Belle II Experiment

The Belle II experiment as a B-factory is set to produce a large number of B-mesons. It aims to precisely measure the size of CP violation in the $B\overline{B}$ system and its decay. This is especially important to better understand the matter anti-matter asymmetry of nature. Any deviations from the Standard Model found in precision measurements of CP violating decays could be a hint of new physics.

Another main product of e^+e^- collisions at $E_{cm} = 10.58\text{GeV}$ is the production of $\mu^+\mu^-$ and $\tau^+\tau^-$ pairs. The cross section for these processes is about the same size as that for the $\Upsilon(4S)$ production. The high number of τ pairs to be produced allows Belle II to contribute a lot to our understanding of τ physics. Particularly it is being searched for lepton family number violation in the decay of τ leptons in processes, like $\tau^\pm \to \mu^\pm\gamma$ and $\tau^\pm \to \mu^\pm + \mu^+\mu^-$. These decay modes are forbidden in the Standard Model but appear in extensions like many neutrino mass mechanisms. For this search particle identification in the lepton sector is especially important to differentiate these exclusively leptonic decays from allowed decays in the Standard Model [1, 10].

Belle II is also searching for potential dark sector physics, for which it has the right prerequisites, namely its precisely known total center of mass energy, Lorentz boost and the large amount of data at hand. Therefore it is feasible to search for missing energy in Standard Model processes. These may be caused by production of dark matter particles, that we can not observe in the detector [1].

3.2   SuperKEKB Accelerator Facility

The high energy e^+e^- beams at Belle II are provided by the SuperKEKB accelerator facility, which was upgraded from its predecessor KEKB used for Belle. SuperKEKB aims to achieve 40 times the luminosity of its KEKB. This is mainly done by decreasing the beam size at the collision point by a factor of 20. The target luminosity will be $8 \times 10^{35} \frac{1}{\text{cm}\cdot\text{s}}$ [1].
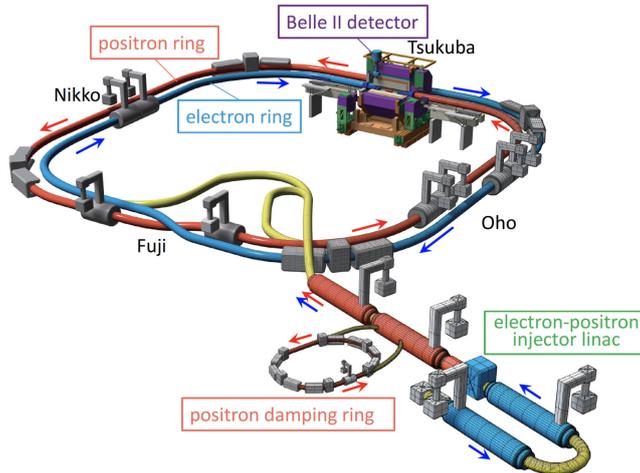

Figure 1: Structure of the SuperKEKB accelerator complex [11].

Since SuperKEKB collides electrons and positrons, two separate storage rings are needed. The structure of the accelerator complex is shown in figure 1. The collider mainly operates at the center of mass energy of 10.58 GeV corresponding to the $\Upsilon(4S)$ resonance. The center of mass system is Lorentz boosted to the lab system, so that the decay products travel some measurable distance in the lab system before decaying. This makes it possible to gather precise information about the B-meson lifetime and CP violation. To achieve the boosted system, the beam energies are asymmetric with the high energy electron beam having an energy of 7 GeV and the low energy positron beam chosen to be 4 GeV [1].

3.3   Detectors

The Belle II detector was upgraded from Belle significantly, to be able to handle the higher background levels caused by the increased luminosity. The detector is in general barrel shaped around the beam axis. The top and bottom of the barrel are called end caps. The structure of the experiment is shown in figure 2. The coordinate system of Belle II has its origin in the collision point. The z axis is parallel to the bisector of the two beams. Its positive direction is in the flying direction of the high energy beam. From the z axis the polar angle θ can be defined with $\theta = 0$ being along the z axis [12].

The collision takes place in a homogeneous magnetic field of around 1.5T at the center in z axis direction. The field is produced by a superconducting solenoid magnet and an iron yoke. This causes all charged particles to fly on curved trajectories dependant on their charge and momentum. The detector is
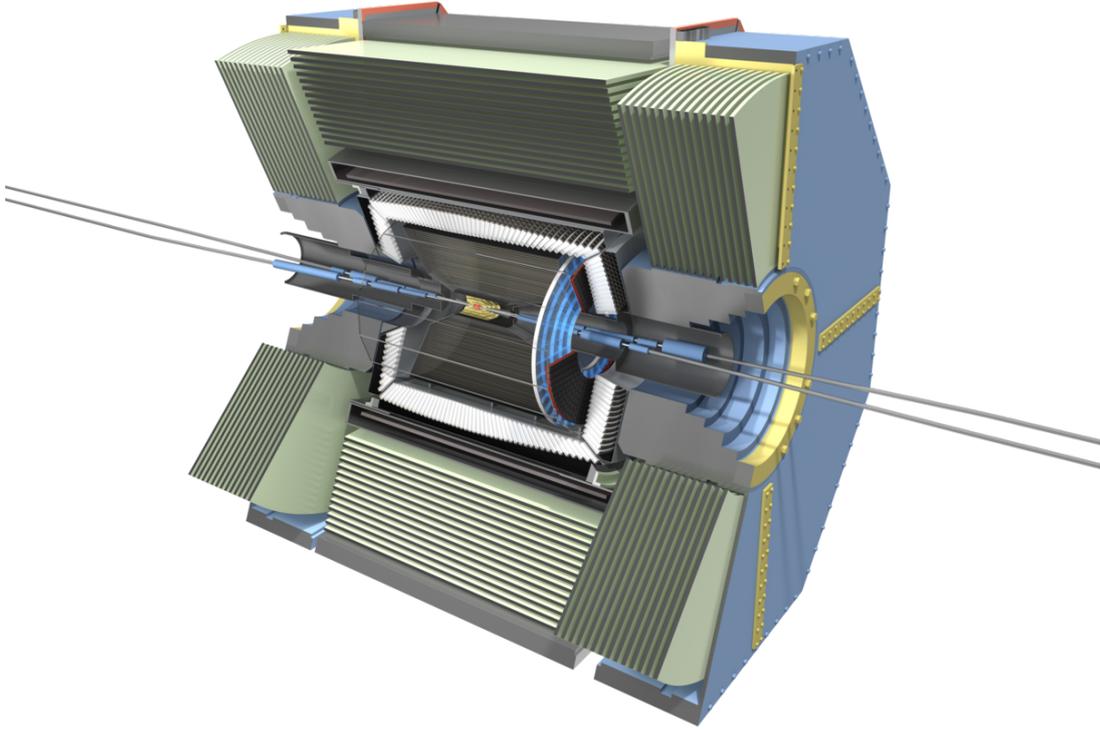
Figure 2: Structure of the Belle II detector [13].

build in such a way that a particles track can be reconstructed, its energy, momentum and velocity determined and thus a species of particle assigned. For that Belle II makes use of several sub detectors, each specialized for some of those tasks. The sub detectors are mounted in layers around the collision point along the barrel and end caps. In the following sub sections the sub detectors will be described in detail.

3.3.1 Tracking and Ionization Detectors

A very important task in the Belle II experiment is the reconstruction of particle tracks. This is done using three subdetectors: the silicon pixel detector PXD, silicon vertex detector SVD and the central drift chamber CDC. The PXD and SVD are often grouped together as the vertex detector VXD. All of those detectors make use of the ionization caused by a charged particle passing [1].

When a charged particle passes through matter, it ionizes the atoms and loses energy. This ionization causes an avalanche of electrons that conduct an electric current in an applied electric field in the detectors. This causes a signal in readout wires, that is used to reconstruct the particles track [14]. Since the charge of all long living charged particles expected to be observed is $\pm e$, the particle momentum can be inferred from the curvature of its track in the magnetic field [1].

Additionally, from the ionization the energy loss per unit length dE/dx of the particle can be measured [15]. The energy loss is governed by the Bethe-Bloch formula and depends on the particle's velocity. Therefore at the same particle momentum, particles of different masses will have different energy losses [8]. This allows to calculate a likelihood for different particle species hypothesis. For values of $\beta\gamma$ in a range of 1 to 10, the Bethe-Bloch curve reaches a flat minimum [14]. This makes it harder to separate between particle species since for two different species of particles with same momentum and thus different speed the energy loss is nearly the same [14, 15]. The expected dE/dx in the CDC for several particle species of different momenta is shown in figure 3.
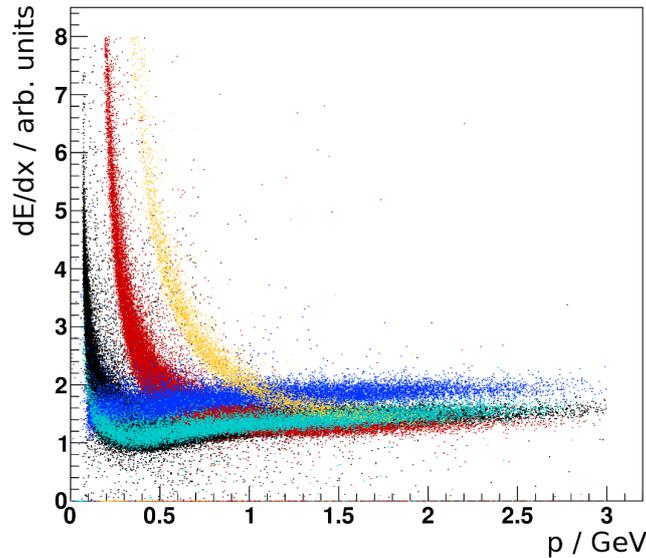
Figure 3: dE/dx in the CDC for electrons (blue), muons (turquoise), pions (black), kaons (red) and protons (yellow) as a scatter plot of the particle momentum [16].

Vertex Detector VXD

The Vertex detector VXD is the innermost detector of Belle II. It consists of the silicon pixel detector PXD and the silicon vertex detector SVD, both located in cylindrical layers around the beam lines and collision point. In forward direction of the experiment the detector wedges in for greater angle coverage. Overall the range of the covered polar angle of both detector types is 17° to 150° in the lab frame. The PXD consists of so called DEPFET type sensors arranged in two layers at radii of 14mm and 22mm. The SVD consists of double-sided silicon strip sensors in layers at 9mm, 80mm, 104mm and 135mm [1, 17]. Both detector types work on the same general principle. A particle passing through ionizes the silicon atoms produces electron hole pairs. These are then detected as currents in the applied electric field [18]. The VXD is mainly used for track reconstruction. It has to be of very high resolution for a precise reconstruction of decay vertices of short lived particles like the K_s^0 [1].

Central Drift Chamber CDC

The core tracking device of Belle II is the central drift chamber CDC. It is a large volume drift chamber filled with a mixture of He and C_2H_6 gas in a ratio of 50:50 that extends out to a radius of 1130mm from the beam axis. The CDC contains 14,336 sense wires arranged in 56 layers. These wires are either aligned with the magnetic field (axial) or skewed (stereo) [1]. The electron avalanche resulting from an ionizing particle passing causes a current between the wires and the cathodes lying in between. This results in an electric signal in the wires. Since the drift velocity is known, one can deduce the distance between the wire and the passing particle from the time it takes for the ionization electrons to reach the wire [14, 15]. By combining the information from both the axial and stereo wires, the particle tracks can be reconstructed [1].

3.3.2 Cherenkov Detectors

In addition to the information about the momentum of a particle, one also needs to know its speed to obtain the mass for a valid particle identification. A common effect used for this measurement is the emission of cherenkov radiation by a particle traveling faster than the speed of light in a medium. The cherenkov radiation is emitted in a cone pointed in the flying direction of the particle. Its opening angle is given by the speed of the particle βc and the speed of light $\frac{c}{n}$ with n being the refractive index in the surrounding medium via the formula [8, 19]:

$$cos\theta = \frac{1}{n\beta}. \qquad (4)$$

For high velocities the angle θ approaches an asymptotic maximum at $\theta_{max} = arccos(\frac{1}{nc})$. In this high velocity limit, the opening angles hardly change with β. Consequently for a particle with known momentum, the separation between other particle species becomes harder, the higher its velocity is [19].

Time of Propagation TOP

The cherenkov detector in the cylindrical barrel with polar angles of f 31° to 128° is called the time of propagation (TOP) detector. It consists of 45 cm wide and 2 cm thick quartz bars with photo anodes mounted at the end. It measures the time dependant projection including reflections of the cherenkov cone onto the photon detector. Since a cherenkov photon is reflected at the borders of the detector volume, its path and thus the time to propagate through the detector is dependant on the opening angle θ of the cherenkov light cone [1, 20].

At the TOP detector the particle identification is in practice done by comparing the detector outcome wit that theoretically predicted for each particle species hypothesis passing through with given track and momentum. This is especially important for separation of kaons and pions, but lacks for pions and muons because of their similar mass [1, 21].

Aerogel Ring Imaging Cherenkov Detector ARICH

The forward endcap region of 14° to 30° is equiped with an aerogel ring imaging cherenkov detector (ARICH). The ARICH measures the opening angle of the cherenkov lightcone directly. To enhance the intensity of the cherenkov radiation, it makes use of two layers of different refractive indices, so that their cherenkov cones are overlapping at the photon detector. It consists of two layers of aerogel of 2cm thickness each and refractive indices n = 1.045 and n = 1.055. The opening angles of the cherenkov cones are different in the two layers according to equation 4. The cherenkov photons are then detected by a very sensitive hybrid avalanche photon detector (HAPD). Again a likelihood for each particle species hypothesis can be determined [1].

3.3.3 Electromagnetic Calorimeter ECL

Outside of the central drift chamber lies the electromagnetic calorimeter (ECL). It consists of caesium iodide crystals doped with thallium CsI(Tl) which act as a scintillator. The detector covers around 90% of the solid angle in the center of mass system of the collision [1]. A charged particle hitting a crystal produces a particle shower and dispenses a lot of energy. The exited scintillator crystal emits light on relaxation to lower energy states, which can be measured to obtain the energy [8]. In general three types of interactions with the detector crystals may take place. These are electromagnetic and hadronic showers as well as ionization [22]. Electrons hitting the ECL mostly emit bremsstrahlung, which is suppressed for other particles due to their higher masses. As electrons have a low mass they deposit nearly all of their energy in the ECL and are therefore easy to detect using this device. Hadrons on the other hand mostly interact via the strong interaction [8]. These events can be separated from electromagnetic interactions via the cluster shape, which can be obtained from the measurements of the individual crystals [1, 22].

3.3.4 K_L and Muon Detector KLM

The last subdetector of Belle II is the K_L and muon detector KLM, which lies outside the solenoid magnet. It consists of altering layers of 4.7cm thick iron plates and sensitive detector elements. Because of its positioning at the outer edge of the detector, most particles produced in a reaction will be shielded by the ECL and solenoid magnet and do not reach the KLM. Particles that can reach the KLM are muons as well as long lived hadrons like K_L and some charged pions. The muons reach the KLM since in contrast to electrons they are much heavier and therefore hardly emit bremsstrahlung. They instead mostly loose energy via ionization. Each track from the inner detectors is extrapolated outwards with the muon hypothesis and then compared to hits in the KLM to discriminate against other particle species [1, 23].

4 Methods of Data Analysis

In this section a brief overview of analysis methods important for this thesis is given. This includes some underlying concepts and methods like Monte Carlo simulations and likelihood, that are not explicitly used for this thesis analysis, but that are required to understand the data sample and the concept of particle identification. Also the tag and probe method and calculation of sWeights will be explained.

4.1 Monte Carlo Simulation

A common tool for studying the performance of a detector and its evaluation methods are Monte Carlo simulations. In these simulation the physics processes are simulated by a event generator. The detector information is produced in detail by a detector simulation. These simulations are of course also dependant on the detectors software. In case of Belle II the development of the Belle II analysis software framework basf2 is still ongoing and the evaluation of old data is reproduced by new versions of the software. Therefore also the Monte Carlo simulation data is produced for each version of basf2. The simulations are labeled as MC followed by the number that indicates the production software version. Also the condition of the detector and the beam can change over time. Therefore the simulation can be run dependent noted as "rd", if the condition of the detector for a specific data taking run is taken into account, or run independent referred to as "ri" [1].

The big advantage of Monte Carlo data is the fact, that the truth, for example the true species of a particle, is known. In the following quantities calculated using this truth information are referred to as using MC truth. This truth knowledge can also sometimes be ambiguous. The reconstructed tracks in the detector from the detector simulation have to be matched to the trajectories of a particle with known species from the physics simulation. This can be sometimes ambiguous, but still provides information about the true identity of most events.

The data sample used in the analysis of this thesis stems from MC15ri simulation of the detector. A simulation is used instead of real data, since the truth information offered is required as a benchmark to compare the performance of the background subtraction to.

4.2 Likelihood

The likelihood is a measure of how well observations fit to the predictions of a model. To define the likelihood one first needs to introduce probability density functions (PDF). A PDF $f(x; p)$ describes the probability distributions for each observable x and parameter set p. Since its a probability density the PDF is positive and normalized, so that its integral is equal to one over the allowed range of the observable x for any set of p values. From the PDF one can define the Likelihood of an event e with given observables x_e as [10]:

$$\mathcal{L}_e(p) = f(x_e; p). \tag{5}$$

Based on this one can define the negative log likelihood for later ease of use. The total likelihood of a data set of multiple events is the product of the likelihoods of the single events or alternatively the total log likelihood is the sum of the single event log likelihoods [10]:

$$\mathcal{L}(p) = \prod_e \mathcal{L}_e = \prod_e f(x_e; p). \tag{6}$$

$$-\mathrm{log}\mathcal{L}(p) = -\sum_e \log(\mathcal{L}_e). \tag{7}$$

4.3 Charged Particle Identification

The Belle II detector allows us to distinguish between all types of particle species that can occur after the collision process. This includes all proposed charged particle species hypothesis which are electrons, muons, pions, kaons, protons and deuterons. Each sub detector is particularly suitable for identifying certain types of particles and has different separating powers [1].

For a given track, each sub detector the track has passed, can make a likelihood statement for each particle species hypothesis. This indicates, how likely it is, that a given track was of a given particle species. However as explained, it is not feasible to use the likelihood of just one detector to receive a valid result, since each detector is especially suited for the separation of certain particles. Therefore one can combine the likelihoods for a particle species hypothesis h from multiple detectors [1]:

$$\mathcal{L}(h) = \prod_{det} \mathcal{L}_{det}(h). \tag{8}$$

Often one wants to separate only two species. For that the likelihood ratio can be calculated to quantify the separation power between two types of particle hypothesis h_1 and h_2. Using the corrected equation found in [1], we can write:

$$L(h_1 : h_2) = \frac{\mathcal{L}(h_1)}{\mathcal{L}(h_1) + \mathcal{L}(h_2)} = \frac{1}{1 + e^{\mathrm{log}\mathcal{L}(h_2) - \mathrm{log}\mathcal{L}(h_1)}}. \tag{9}$$

If $L(h_1 : h_2)$ is greater than 0.5 the particle is more likely to be of species h_1 whereas otherwise the track resembles type h_2 more.

In the same way, one can define a global likelihood, which will also be referred to as the particle ID, for the species h by calculating the ratio between the likelihood $\mathcal{L}(h)$ and the sum over the likelihoods of all species hypothesis h_i [24]:

$$L_{global}(h) = \frac{\mathcal{L}(h)}{\sum_{h_i} \mathcal{L}(h_i)}. \tag{10}$$

4.4 Particle Identification Efficiencies and the Tag and Probe Method

Efficiencies measure, how good a detector is in assigning a particle its species. We want to measure efficiencies to determine the effect of particles that are wrongly identified and to crosscheck the simulated data. The particle identification efficiencies are dependant on kinematic variables like momentum p and polar angle θ [1]. The efficiency ϵ can be computed as the number of particles correctly identified as belonging to a given species $N_{id}(p, \theta)$ divided by the total number of all particles of that species $N_{tot}(p, \theta)$ accepted in the detector in a given momentum and polar angle bin [2]:

$$\epsilon = \frac{N_{id}(p, \theta)}{N_{tot}(p, \theta)}. \tag{11}$$

However to compute efficiencies in that way, we need a pure sample of tracks, of whom the particle species is known. One method to obtain such a sample is the tag and probe method. It is performed on known decay modes into two particles. The underlying data set therefore consists of events of two particles. It is assumed, that the species of both particles can be tested independently. If two particles stem from the same decay and the species of one of the particles is known, the species of the other one can be automatically tagged. By performing particle identification on the second particle, the so called probe particle, the particle identification efficiency can be obtained [3].

To do the tag and probe approach, some cuts have to be applied to the data. On the initial data we make a tight selection on the tag particles particle ID for the particle species we want to test the efficiency of. On that set of so called tag cut data background subtraction utilizing the methods specified below is performed. After that only events remain, that are believed to stem from the decay, we want to use for the efficiency calculation. If the particle ID selection on the tag particle was right and the background subtraction was done correctly, the remaining events should only be such, where the probe particle is truly of the species, we want to investigate. Therefore the particle ID efficiency ϵ can be calculated as the fraction the selected events, where the probe particles species is correctly identified, make up of all events [3]:

$$\epsilon = \frac{\text{Number of passing probe particle events}}{\text{Number of all events}}. \tag{12}$$

The denominator corresponds to the number of all events after the tag cut and background subtraction. The numerator corresponds to the number of those events, where the probe particle passes the particle ID criterion and is thus correctly identified. Those numerator events will be referred to as after the probe cut.

As we can see from equation 12, the efficiency calculation only requires the knowledge of the number of numerator and denominator events, not which exact events were selected. To be able to calculate the efficiency in bins of momentum and angle, we also need to reconstruct the distribution of the signal events in those variables. This can be achieved using the $_s\mathcal{P}$lot background subtraction.

4.5 $_s\mathcal{P}$lot Method

The $_s\mathcal{P}$lot technique was developed to extract the distribution of signal type events from an event sample made up of multiple contributions. So effectively it is a type of background subtraction technique. The $_s\mathcal{P}$lot method can be used on a data sample consisting of contributions of several sources, that we want to distinguish. These are in our case signal and background. For the data sample there exists a set of observables, that we can determine for each data point and therefore construct histograms. These histograms count the combined events all sources. A set of observables, for which we know the distribution functions for both signal and background, can be defined as the so called discriminating variables. The distribution function in the discriminating variables of signal and background has to be different. The remaining observables, that are not defined as discriminating variables, are called control variables [4].

We now want to obtain the histograms in the control variables for only signal type events. The $_s\mathcal{P}$lot method does that by assigning a weight to each event, that is constructed from the known distribution functions $f_s(m)$ and $f_b(m)$ for signal and background in the discriminating variable m. This is called the sWeight. From the weighted sample one can extract the count distributions for control variables that are uncorrelated to the discriminating variables. In case of a correlation existing between discrimination and control variable the resulting distributions from the sWeights may not be valid [4].

To calculate the sWeights, one has to first define some more quantities. For each event e its value of the discriminating variable is denoted as m_e. The total number of events of signal and background are referred to as N_s and N_b. These can be introduced as so called yield parameters for the PDF model in the discriminating variable, to describe the counts with the otherwise normalized PDF. Finally we need the covariance matrix V_{ij} of the likelihood of the data using the given PDF model. The likelihood $\mathcal{L}$ is calculated using equation 6. The covariance matrix element V_{ij} is then computed for the signal and background yield. For i, j being s, b for signal and background it is defined as [4]:

$$V_{ij}^{-1} = \frac{\partial^2(-\mathcal{L})}{\partial N_i \partial N_j}. \tag{13}$$

From all these quantities one can define the sWeights $_s\mathcal{P}(m_e)$. They can be computed for each event using the following formula [3, 4]:

$$_s\mathcal{P}(m_e) = \frac{V_{ss}f_s(m_e) + V_{sb}f_b(m_e)}{N_s f_s(m_e) + N_b f_b(m_e)}.$$

(14)

In case the covariant matrix element $V_{sb} = 0$ and $V_{ss} = 1$, the sWeight $_s\mathcal{P}(m_e)$ is just the probability of an event with discriminating variable values m_e being a signal event.

From the sWeights, one can reconstruct histograms of only signal events in the control variable. The number of signal events $N_s(B)$ in the control variable bin B can be obtained by summing over the sWeights $_s\mathcal{P}(m_e)$ of all events e_B that are in the given control variable bin B [4]:

$$N_s(B) = \sum_{e_B} {}_s\mathcal{P}(m_{e_B}).$$

(15)

The standard deviation σ_s of the number of events in that bin can be calculated as the square root of the sum of squares of the sWeights $_s\mathcal{P}(m_e)$ [4]:

$$\sigma_s = \sqrt{\sum_{e_B} {}_s\mathcal{P}^2(m_{e_B})}.$$

(16)

5 The J/Ψ Data Set

5.1 Decay Modes of J/Ψ

The decay modes of J/Ψ can be distinguished into hadronic decays and leptonic decays. Hadronic decays are such decays, that have only hadrons and exchange bosons in the final state and therefore involve predominantly the strong interaction. On the other hand for leptonic decays there are only leptons and exchange bosons produced, which predominantly happens via the electromagnetic interaction [6]. In the decay of J/Ψ those two decay types occur roughly at equal rates. The hadronic decays have a branching fraction of $(87.7 \pm 0.5)\%$ whereas leptonic decays have a branching fraction of $(11.9 \pm 0.1)\%$ [9]. The hadronic decays to a large extend are decays involving three gluons or two gluons and a photon, which then produce pions, kaons and other light resonances. On the other hand the leptonic decays involve a photon [6]. The leptonic decays have about equal branching fractions into pairs of electrons e^+e^- with $(5.971 \pm 0.032)\%$ and muons $\mu^+\mu^-$ with (5.961 ± 0.033) % [9].

5.2 Data Sample

In this analysis the muon identification efficiency is ought to be tested using the decay $\text{J}/\Psi \to \mu^+\mu^-$. In this section we will introduce the data sample used for that analysis and explain the cuts that are initially applied.

The data sample consists of 4,925,436 events from the MC15ri Monte Carlo simulation of the detector. Each event contains two muon candidate particles which will be denoted as particle 0 and particle 1. For each of these particles the total momentum p, polar angle θ and muon ID are stored along with some other variables unused in the analysis. The muon ID is the likelihood of a given particle being a muon determined from the detector results as in equation 10 [24]. Also, for each event we calculate the invariant mass $m_{\mu\mu}$ of the two particle system. Since the events stem from a Monte Carlo simulation, for each event it is also known, if it comes from a true J/Ψ decay. Such events, that contain two true muons that stem from a J/Ψ decay are called signal, whereas the rest of the events will be referred to as background.

The data sample used has undergone some initial selection. These cuts will just be briefly described here. More detailed information can be found in [3]. Firstly, since we are reconstructing the decay of a J/Ψ, the two tracks have to originate from the same interaction point. In the lab frame the difference in coordinates for this point is required to fulfill: $|\Delta z| < 5$ cm and $|\Delta r| < 2$ cm. Moreover it is required for the invariant mass $m_{\mu\mu}$ to be within 2.8 GeV/c^2 to 3.3 GeV/c^2. Both particles are required to have clusters in the ECL calorimeter which are matched to tracks in the inner detectors. Furthermore, to reduce

beam induced background, it is required for the momentum p of the particles that $p > 0.1$ GeV/c. To suppress $e^+e^- \to q\bar{q}$ events, it is required for the ratio of the second order and zeroth order Fox-Wolfram moments to be below 0.4. To further suppress e^+e^- collision background processes, for each event at least five tracks are required in the detector [3].

We use the tag and probe approach as explained in section 4.4 to calculate efficiencies. To fully use the available data, the data sample can be symmetrized. This is done by using particle 0 as tag muon and particle 1 as probe muon as well as particle 0 as probe and particle 1 as tag muon. This effectively doubles the data set. After that we make the tag cut. In the analysis of the systematic corrections framework the muon ID is required to be greater than 0.95 [3]. However for all of the following analysis the tag cut is only done for a muon ID value of 0.9. This does not effect the results of the analysis. After the probe cut a total of 5,410,384 events remain. Note that this number is higher than the initial number of events in the data set, which is possible since the sample was symmetrized.

The resulting data set can now be plotted in several variables to get an idea of its structure. In figure 4 the 2d histogram of the absolute momentum of the two particles is plotted. One can see that the data set is not populated for most of the two particle momentum space. However, a band can be identified, which mostly stems from the J/Ψ decay. This peak ranges from a minimal momentum of about 0.6 GeV/c to a maximum tail that extends to about 4 GeV/c for each of the two particles. Most of the sample events are therefore contained in this momentum range. In figure 5 one can see the 2d histogram of the invariant mass $m_{\mu\mu}$ of the two muon system and the probe muon momentum. The peak at the J/Ψ mass of 3.1 GeV/c^2 is clearly visible. We also observe that the number of signal and background events is highest between momenta of 1.0 and 2.5 GeV/c and drops of for higher and lower momenta.
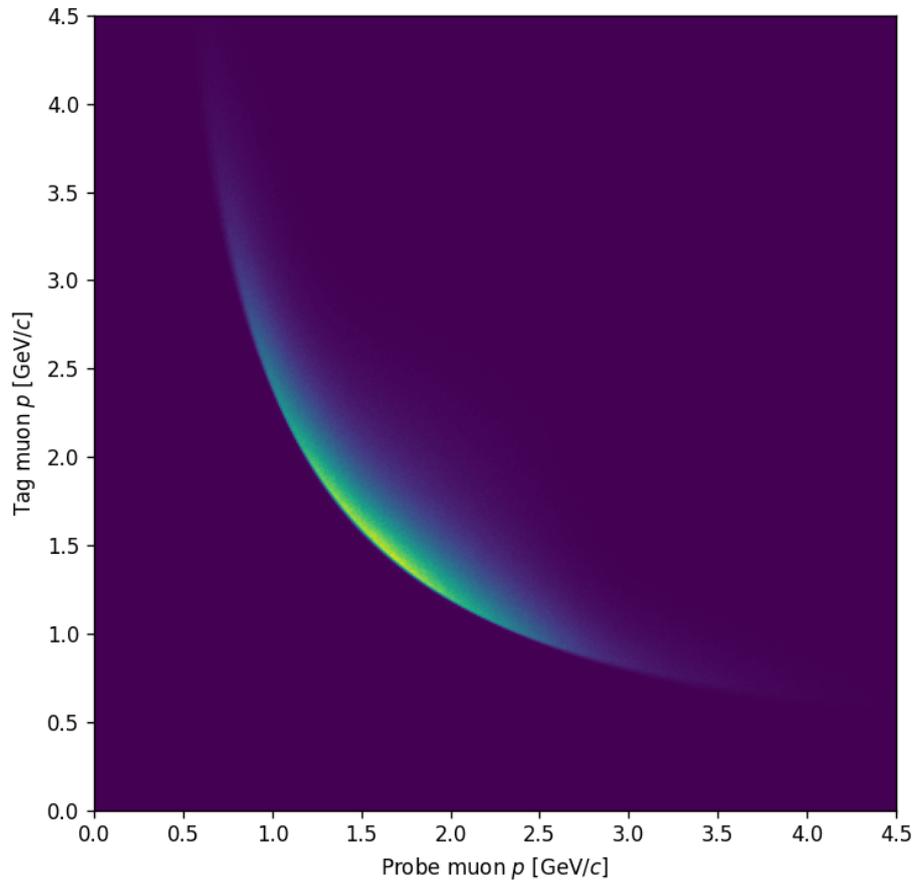
Figure 4: 2d histogram in the momenta of the two particles for the data set after initial cut to tag particle muon ID > 0.9.
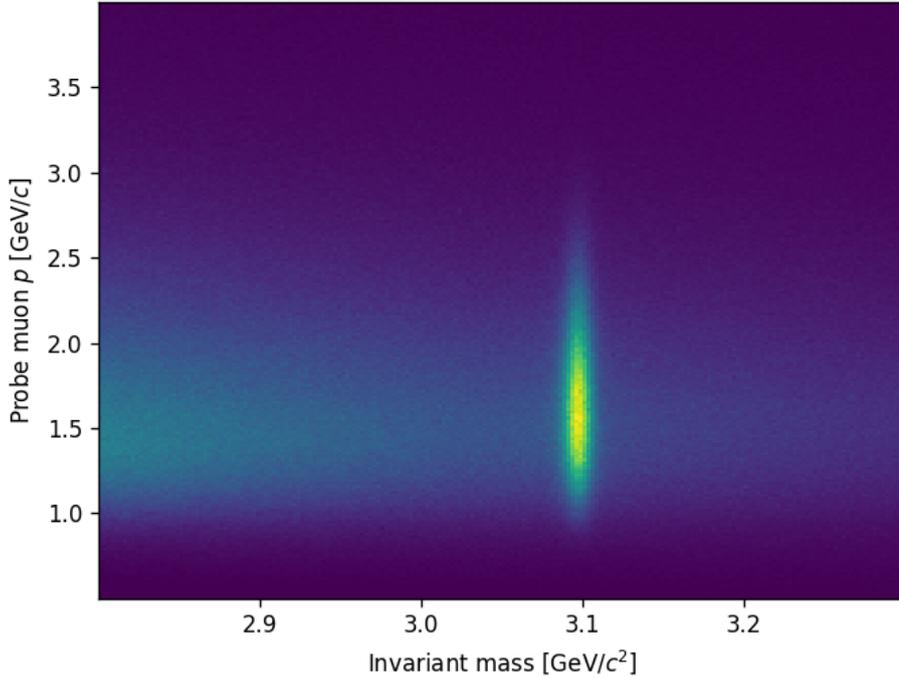
Figure 5: 2d histogram in probe particle momentum p and invariant mass $m_{\mu\mu}$ for the data set after initial cut to tag particle muon ID > 0.9.

In figure 6 one can see a 2d histogram of $\cos\theta$ of the two particles. Note that the data is not symmetric in exchanging the two particles, since the tag cut was applied, which cuts on the muon ID of one of those particles. Apart from that the distribution has higher values in forward direction of the particle collision, which corresponds to the upper right corner of the plot. This makes sense considering that the center of mass frame of the e^+e^- collision is Lorentz boosted. Therefore statistically the decay products of any reaction will also be boosted forward. The distribution has its maximum for one particle flying forwards and the other one backwards. The horizontal and vertical stripes at $\cos\theta$ = -0.63 and $\cos\theta$ = 0.85 are the effect of gaps between the barrel and end cap region of the ECL [3]. In figure 7 the 2d histogram of invariant mass $m_{\mu\mu}$ and the cosine of the probe particle polar angle θ is plotted. Again the J/Ψ signal peak lies at an invariant mass of around 3.1 GeV/c^2. In the forward and backward direction the peak becomes blurred out.

21
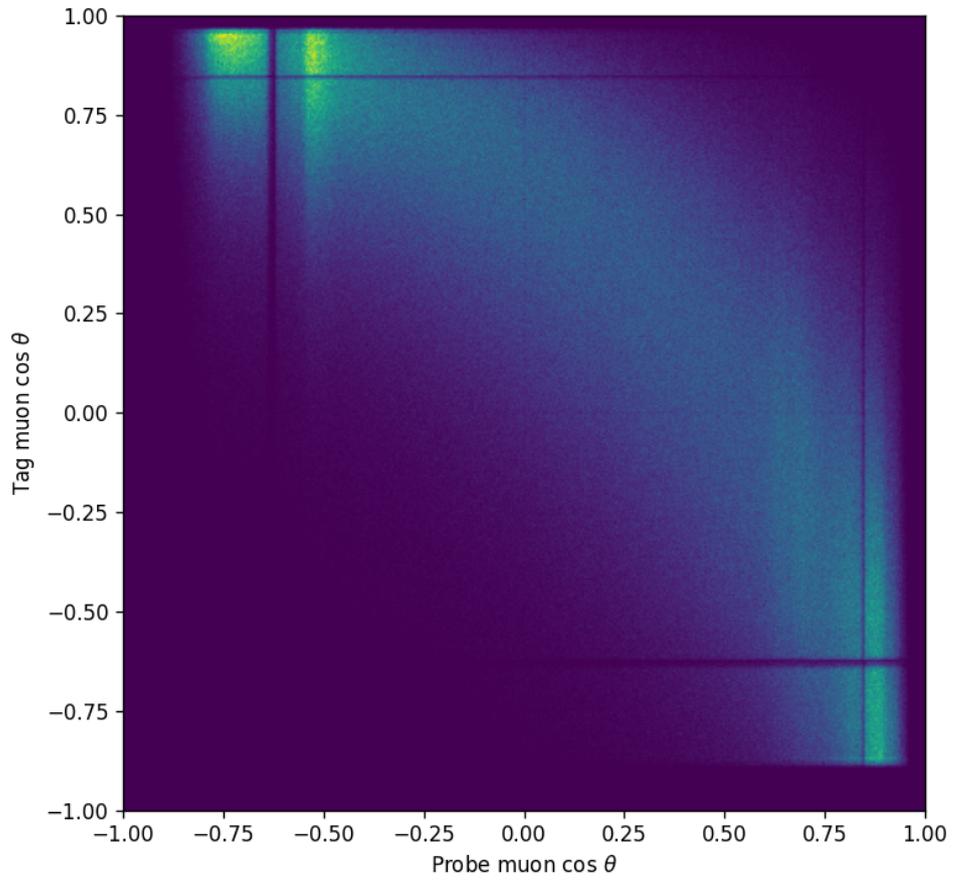
Figure 6: 2d histogram in the cosine of the polar angle of the two particles for the data set after initial cut to tag particle muon ID > 0.9.
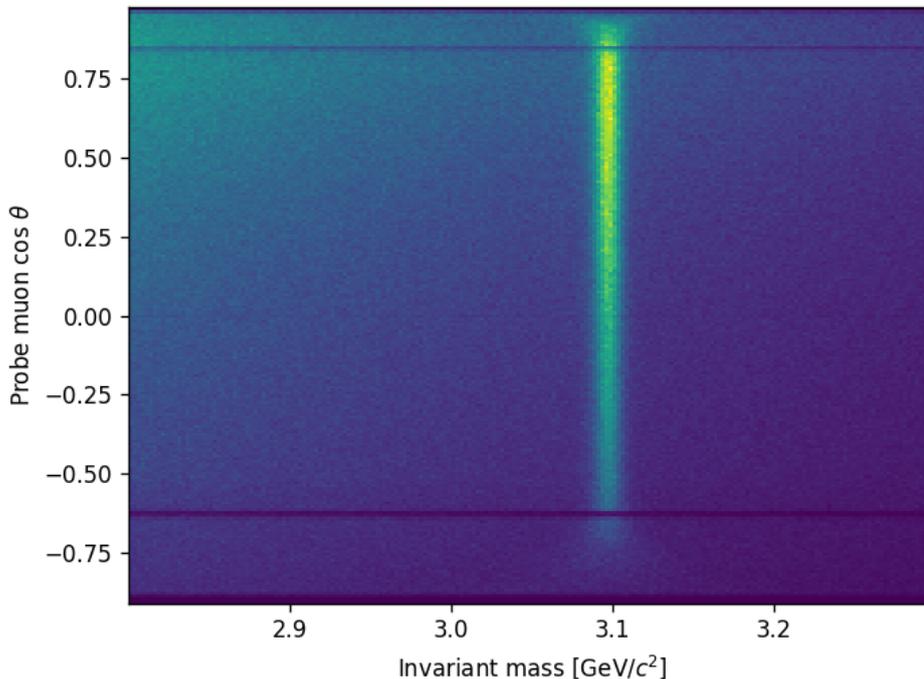
Figure 7: 2d histogram in the cosine of the probe particle polar angle θ and invariant mass $m_{\mu\mu}$ for the data set after initial cut to tag particle muon ID > 0.9.

The discriminating variable for the sWeight calculation is the invariant mass $m_{\mu\mu}$. To obtain a distribution in this variable the data sample was binned into a total number of 300 invariant mass bins of equal width ranging from 2.8 GeV/c^2 to 3.3 GeV/c^2. After binning the number of events in each bin follows a Poisson distribution. The Poisson distribution only has one parameter, the expectation value of the rate. Therefore the standard deviation can be obtained only from knowing the expectation value. If we assume the number of events in the bin to be at the expectation value, the standard deviation can be calculated as the square root of the number of events in the bin. This is used for plotting and propagating uncertainties as done in appendix A.

5.3   Muon Background

The data sample described in section 5.2, that will be investigated in this analysis, is composed of two contributions. The J/$\Psi \rightarrow \mu^+\mu^-$ signal and the background. This background is comprised of events of many reactions and decays, that occur after the high energy collisions at Belle II. Because of the initial event selection these events are in an invariant mass range between 2.8 GeV/c^2 and 3.3 GeV/c^2 and stem from the same interaction point. After the tag cut, described in section 5.2, only some of these events remain, that can not be removed only by these cuts. These are on the one hand the events, where the tag particle is no muon, but is wrongly identified as such, and on the other hand the events, that contain

a muon and that do not stem from a $J/\Psi \to \mu^+\mu^-$ decay but from other muon producing decay modes. These combined backgrounds do not form distinctive peaks but a uniform background distribution, that we need to adjust for in the efficiency calculation.

5.4 Model for Probability Density Function in $m_{\mu\mu}$

In this section the model for the probability density function used in this analysis will be discussed. It describes the distribution of events involving two particles after the cuts described in section 5.2 including the tag cut dependant on invariant mass $m_{\mu\mu}$ of the two particle system. The PDF and its parameterization used in this work originates from the systematic corrections framework, the current framework for such analyses. All functions, values and parameter ranges described in this section were taken over unchanged from this framework [3, 5, 25].

The PDF is a function of the invariant mass $m_{\mu\mu}$ of the $\mu^+\mu^-$ system and is defined and normalized in an invariant mass range from 2.8 GeV/c^2 to 3.3 GeV/c^2. The parameters of this functin will later be determined by a fit to the data. It is constructed as the sum of the signal and the background PDFs. Each of those PDFs is also assigned a yield variable, that modulate what fraction of the distribution is made up of signal and background events. The signal yield N_s describes the number of signal events, whereas the background yield N_b describes the number of background events [3]. Since the total PDF needs to be normalized, only the ratio of signal yield and background yield matters for its overall shape. The individual values of the yields provide additional information in form of the number of events of signal or background [26]. This is especially useful for later sWeight calculation as in equation 14, as both the PDFs $f_s(m)$ and $f_b(m)$ as well as the event numbers N_s and N_b for signal and background are needed.

The PDF is a density function with the units counts per mass step. To be able to compare the PDF to absolute counts in form of a histogram binned in the invariant mass $m_{\mu\mu}$, the PDF can be multiplied by a factor of bin width times the total number of events, which is sum of signal yield N_s and background yield N_b. This reproduces the distribution for a given total number of events.

The signal part of the PDF is again constructed as the sum of two functions. A Gaussian function describes the symmetrical peak and the asymmetrical part of the distribution is modeled by a bifurcated Gaussian. The Gaussian has two parameters, the mean μ and its standard deviation σ_{gaus} also called width. The bifurcated Gaussian is a distribution function with two parameters σ_L and σ_R for the standard deviation on the left and rights side of its peak. The invariant mass, where the bifurcated Gaussian reaches its maximum, is fixed at the same value as the symmetric Gaussian distribution. The two σ_L and σ_R parameters of the bifurcated gaussian as well as the ratio of heights of the Gaussian and bifurcated Gaussian, R_{gaus} were determined by an initial fit and then fixed to their value [3].

In total the signal PDF has up to five independent parameters, but only two of them are free for the final fit. The range of the fit parameters is constricted to a relative variation of 0.02 for the mean μ and 0.4 for the standard deviation σ_{gaus}. The starting values for the fit and which parameters are fixed is listed in table 4.

Table 4: Parameters of the signal PDF and their starting values as well as information if they are free for the fit.

Parameter	Starting value	Status for the fit
μ [GeV/c^2]	3.096	free
σ_{gaus} [GeV/c^2]	$7.19222 \cdot 10^{-3}$	free
σ_L [GeV/c^2]	0.027986	fixed
σ_R [GeV/c^2]	0.0188336	fixed
R_{gaus}	3.7157	fixed

The background PDF is a second order polynomial. It is constructed as sum of chebychev polynomials of the first kind, which serve as a basis. The zeroth order coefficient is set to one, whereas the other two parameters C_0 and C_1 have an allowed range for the fit of [-1, 1] each [25, 26]. Their starting values are listed in table 5. The resulting PDF in terms of C_0 and C_1 is:

$$f_{\text{Background}}(m) = 2C_1 m^2 + (C_0 - C_1)m + 1. \tag{17}$$

Table 5: Parameters of the background PDF and their starting values as well as information if they are free for the fit.

Parameter	Starting value	Status for the fit
C_0 [c^2/GeV]	-0.4	free
C_1 [c^2/GeV]	0	free

The fitting is done using a binned maximum likelihood fit [3]. In this method the likelihood $\mathcal{L}(p)$ depending on the parameter set p, which is calculated using equation 6, is maximized. This means the resulting parameters are chosen in such a way, that the data is the most probable outcome using the resulting PDF model [2]. The fit also determines the covariance matrix of the likelihood as in equation 13, which is needed for sWeight calculation.

6 Results

6.1 Problem

To check, if the background subtraction and all calculations necessary for the efficiencies of the tag and probe approach are valid, one can compare the efficiency results to those utilizing MC truth data. For the MC truth it is known, whether the decayed particle for tag and probe is the desired J/Ψ or not. Therefore the MC truth efficiencies are the ideal efficiency result for the simulation data and should be reached for flawless background subtraction.

The analysis software used for this background subtraction is called the systematic corrections framework [5, 25]. It also includes code to produce particle identification efficiencies in bins of kinematic variables like momentum or polar angle. In figure 8 the muon identification efficiencies are plotted using the systematic corrections framework for MC truth as well as the background subtraction using sWeights in bins of the probe particle momentum. In this figure one can see, that for momenta smaller than about 1.5 GeV/c the efficiency of the sWeights is greater than the MC truth efficiency and even greater than one, which should in theory not be possible for efficiencies. This suggests that in the background subtraction there may be some correlations causing this issue. Therefore in the following sections an attempt is made to understand and fix those correlations in the background subtraction. Different approaches will be made to fix particular deviations and correlations between kinematic variables in the efficiency calculation.
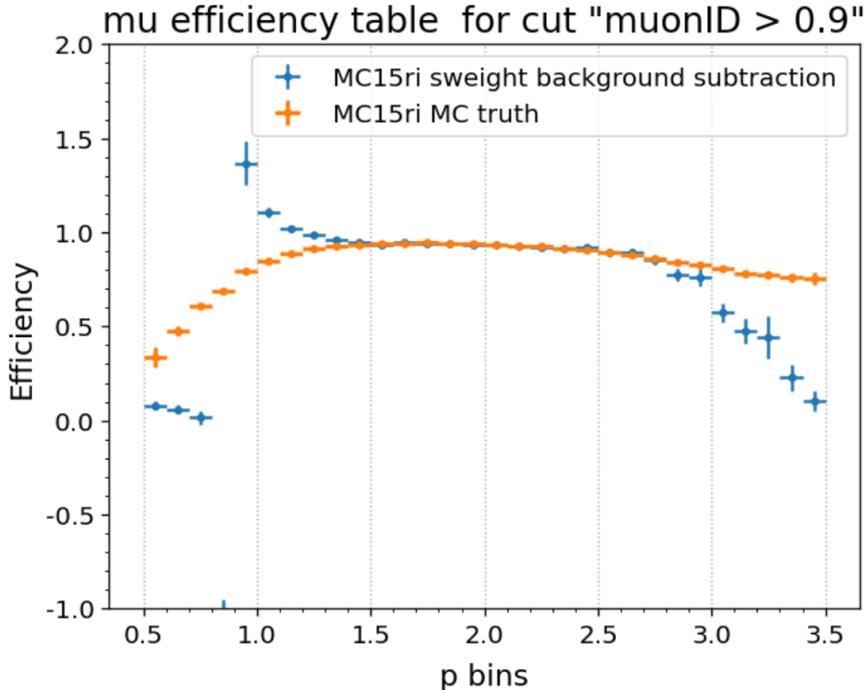
Figure 8: Tag and probe efficiencies in bins of the probe particle momentum in GeV/c from MC truth data and using background subtraction with sWeight from the systematic corrections framework.

6.2   Fitting the PDF and Calculating sWeights

The first step is to produce sWeights and efficiencies to have a benchmark to see how much the different approaches improve the efficiency estimates. This will be in part redone independently form the systematic corrections framework, to better understand the exact procedure of obtaining sWeights and efficiencies.

To calculate sWeights, we first need to fit the PDF, which was described in section 5.4, to the data sample after some initial cuts as described in section 5.2. The resulting parameters are listed in table 6. In figure 9 the fitted PDF and the underlying MC data are plotted. The MC data distribution has a peak at an invariant mass of around 3.09 GeV/c^2, which corresponds to the J/Ψ signal. The background decreases with the invariant mass and is curved slightly upwards. The subplot shows the deviation in form of a pull calculated as $(\text{MC\_counts} - \text{PDF\_counts})/\sqrt{\text{MC\_counts}}$. The pull is fluctuating around zero for most of the invariant mass range $m_{\mu\mu}$, which exhibits that the PDF generally fits the data well. Only directly at the peak, the absolute values of the pull between MC and PDF are higher, which hints, that the PDF may not fit the data perfectly for the signal peak.
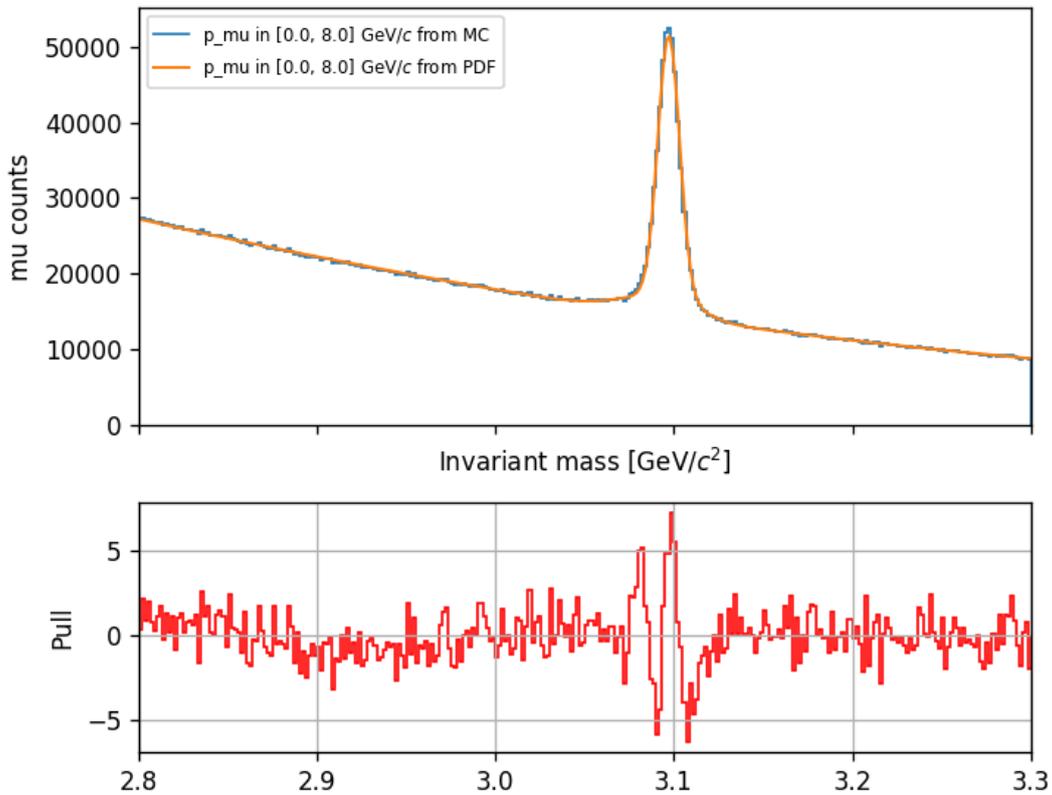
Figure 9: Invariant mass histogram of MC data for probe particle momenta in [0, 8] GeV/c and fitted PDF as well as their pull in the subplot.

Table 6: Resulting PDF parameters after fit to event histogram.

Parameter	Fit result
N_s	397107
N_b	5013270
μ [GeV/c^2]	3.09711
σ_{gaus} [GeV/c^2]	$6.07465 \cdot 10^{-3}$
C_0 [c^2/GeV]	-0.539195
C_1 [c^2/GeV]	0.0590958

From this PDF the sWeights can be calculated for each event using equation 14. By summing up the weights for each event, one can reconstruct the background subtracted counts in bins of the momentum. This can be used to obtain the counts for the denominator and numerator for the efficiency calculation. For the denominator all events are used, whereas for the numerator only the weights of events remaining after the probe cut are summed. In figure 10 the denominator and in figure 11 the numerator counts are plotted in momentum bins for the probe muon momentum p in a range between 0.5 GeV/c and 3.5 GeV/c. We observe that the counts for both denominator and numerator drop to almost zero for momenta close to the borders of the plotting range at 0.5 GeV/c and 3.5 GeV/c, where there are low numbers of events. Both distributions reach their maximum at around 1.6 GeV/c. What stands out is, that the denominators and numerators reconstructed from sWeights are systematically lower than the MC truth values. It turns out, that this is most likely not an effect originating from the count reconstruction using sWeights but from the yields of the PDF fit. The total number of events in the data set almost perfectly agrees with the the sum of signal and background yield. The signal yield however is only 397,107, which is about 8% smaller than the 433,024 true signal events. The mismatch will be discussed in more detail in section 6.5. At this point it seems reasonable to assume, that this PDF related deviation has the same relative effect on denominators and numerators. The mismatch should then cancel out and it should still be possible to obtain valid efficiencies.

The efficiencies and their uncertainties are plotted in momentum bins in figure 12. The error calculation for the efficiencies is specified in appendix A. In the plot we observe that the efficiencies from the MC simulation and that calculated from sWeight background subtraction do not match. The efficiencies determined using sWeights are systematically too high. This could be an effect of the cut applied to the numerator, which will be considered in more detail in the next section. Another thing to notice is, that for momenta greater than 3 GeV/c the uncertainties of the sWeight efficiencies become quite large and are in the same order of magnitude as the efficiency itself, which is a result of the small amount of data points in that range.
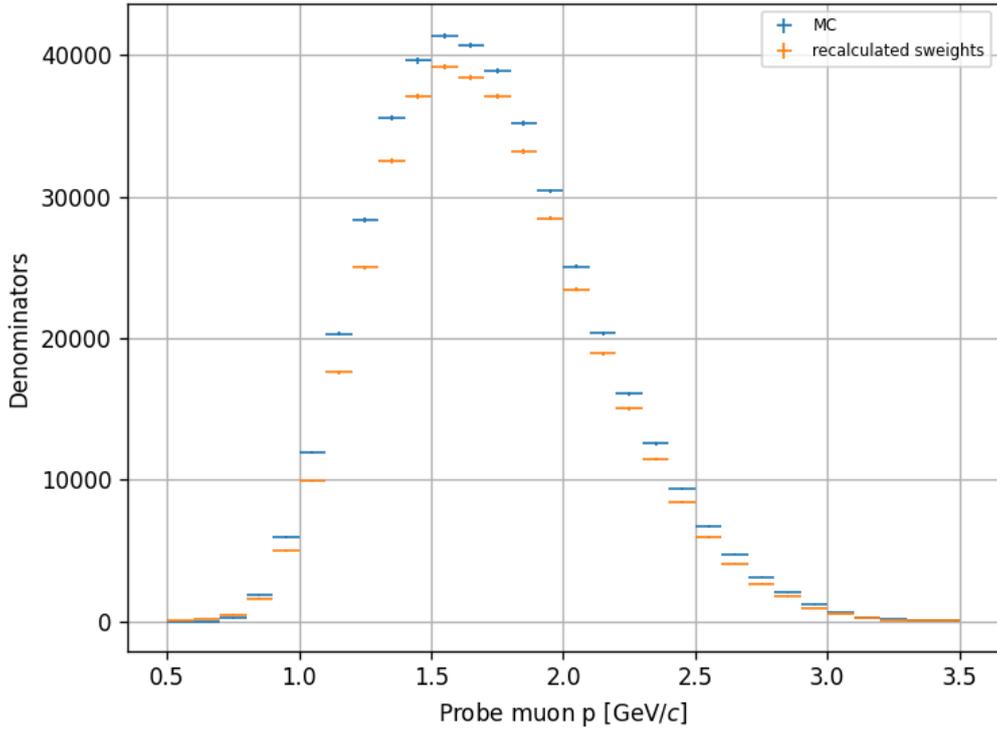
Figure 10: Denominators in momentum bins of the probe particle for the calculation of the efficiencies from MC truth and using sWeight background subtraction.
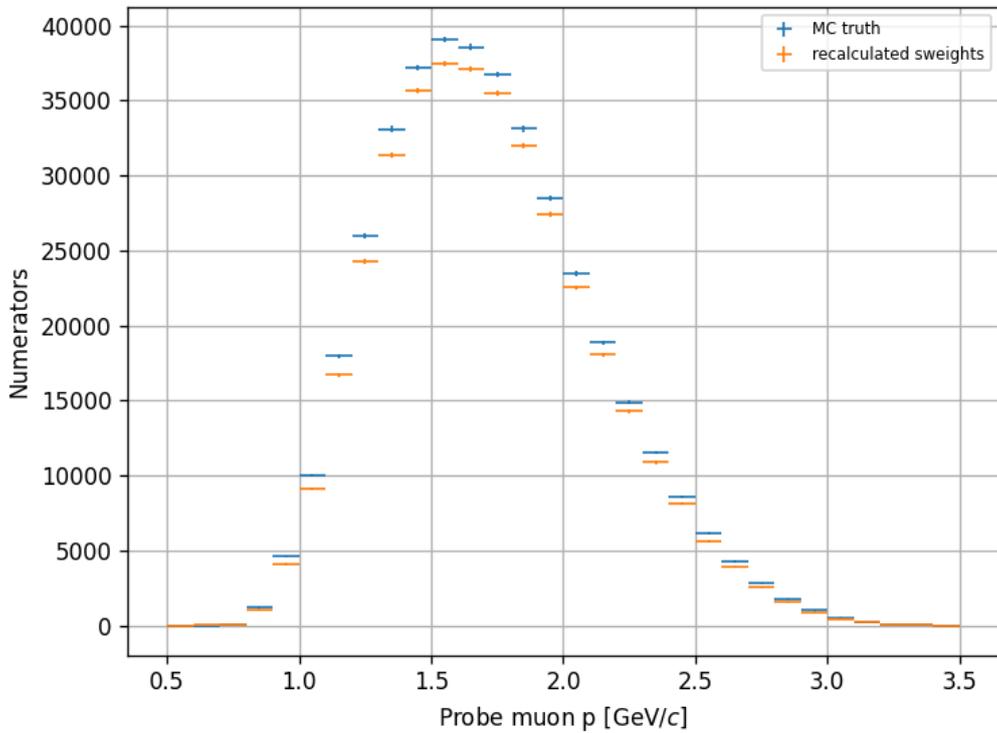

Figure 11: Numerators in momentum bins of the probe particle for the calculation of the efficiencies from MC truth and using sWeight background subtraction.
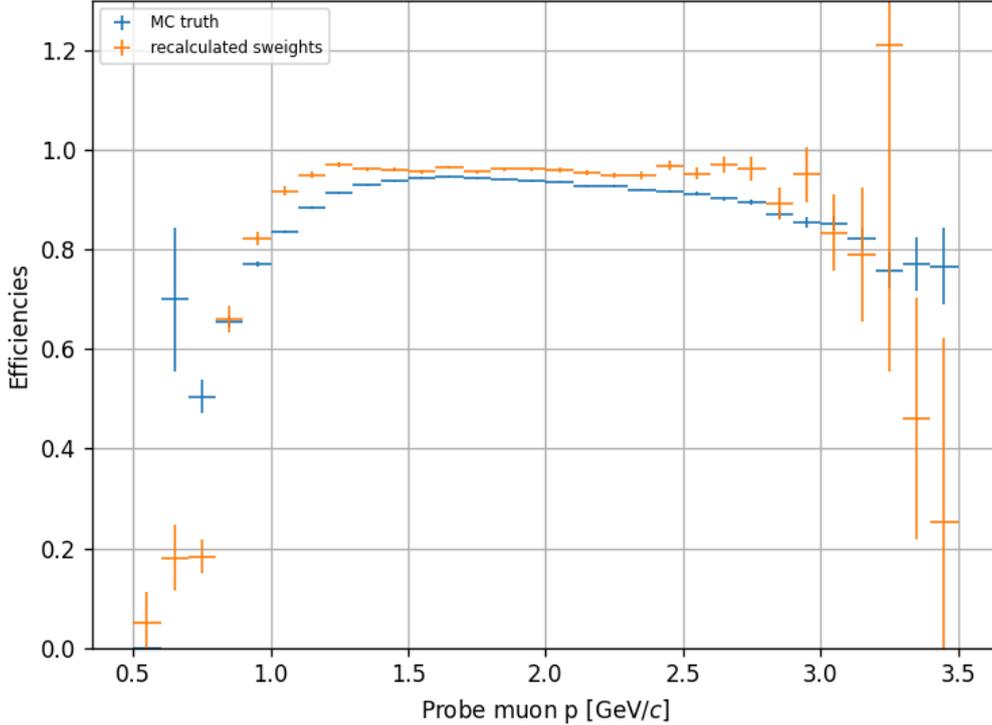
Figure 12: Efficiencies in momentum bins of the probe particle from MC truth and using sWeights background subtraction.

6.3   sWeights Before and After Probe Cut

In this section we will try to resolve systematic deviations of numerator and denominator in respect to each other to the improve the efficiency estimates. Systematically too high efficiencies as seen in the previous section could either be the result of a too high numerator or a too low denominator in relation to each other. This seems very valid when choosing the sWeight approach. Events in the signal mass range have positive weights, whereas events in the background tail region have negative weights. These weights are constructed in such a way, that the total number of events is equal to the number of signal events as estimated by the PDF fit. Now however a cut is applied on the events before calculating the number of events, as illustrated in figure 13. This figure shows the invariant mass distributions for all events and for only signal type events before and after the probe cut. The plot illustrates, that the probe cut mostly reduces background, since it is rare to observe two muons or particles identified as such, that do not stem from a J/Ψ decay in the given invariant mass range. On the other hand the signal peak is hardly reduced, since signal events contain mostly true muons of which most are again classified as such. This means, the probe cut affects the background region much more than the signal. If one now uses the sWeights constructed for a higher background level before the cut to reconstruct the number of signal events after the cut, this may cause discrepancies.
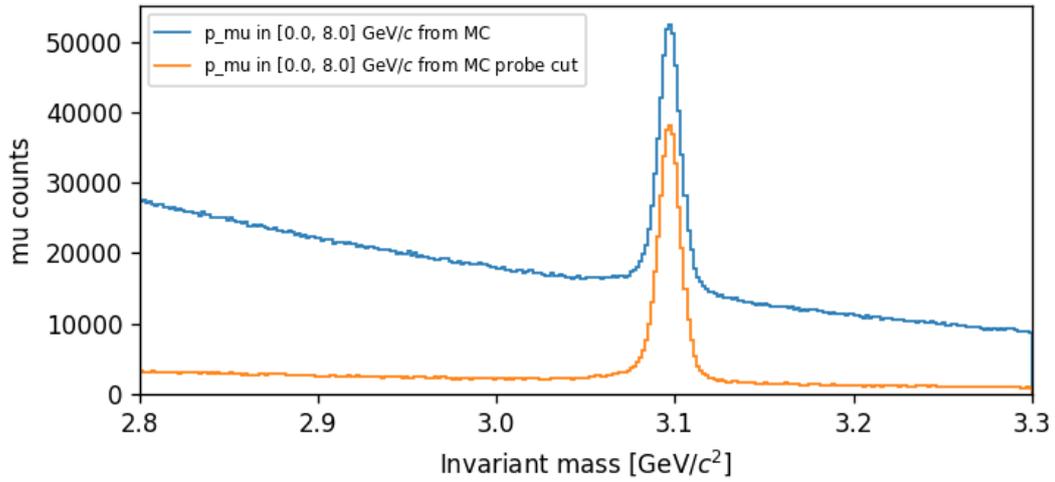
(a) Histogram of all events
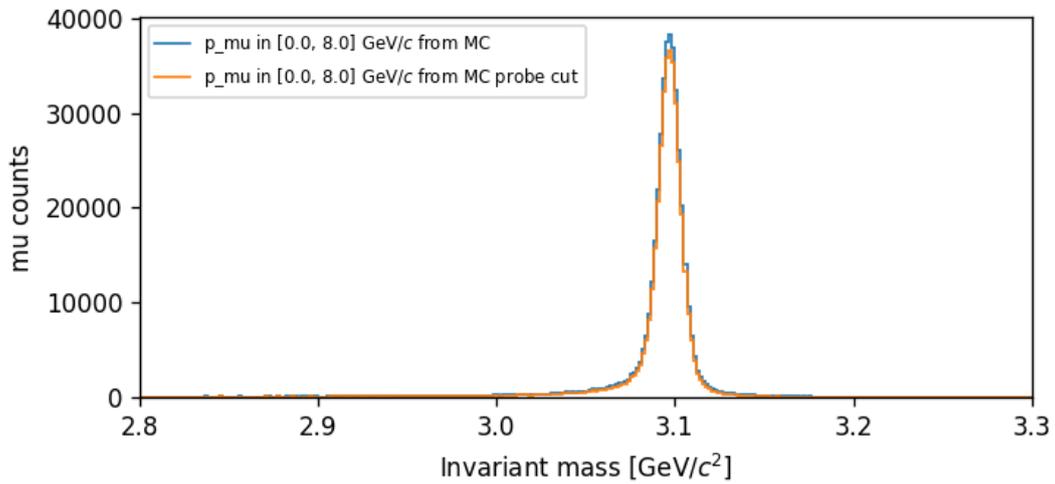
(b) Histogram of signal events using MC truth information

Figure 13: Invariant mass histograms of MC data distribution before and after the probe cut for all events (a) and for the signal events using the MC truth information (b).
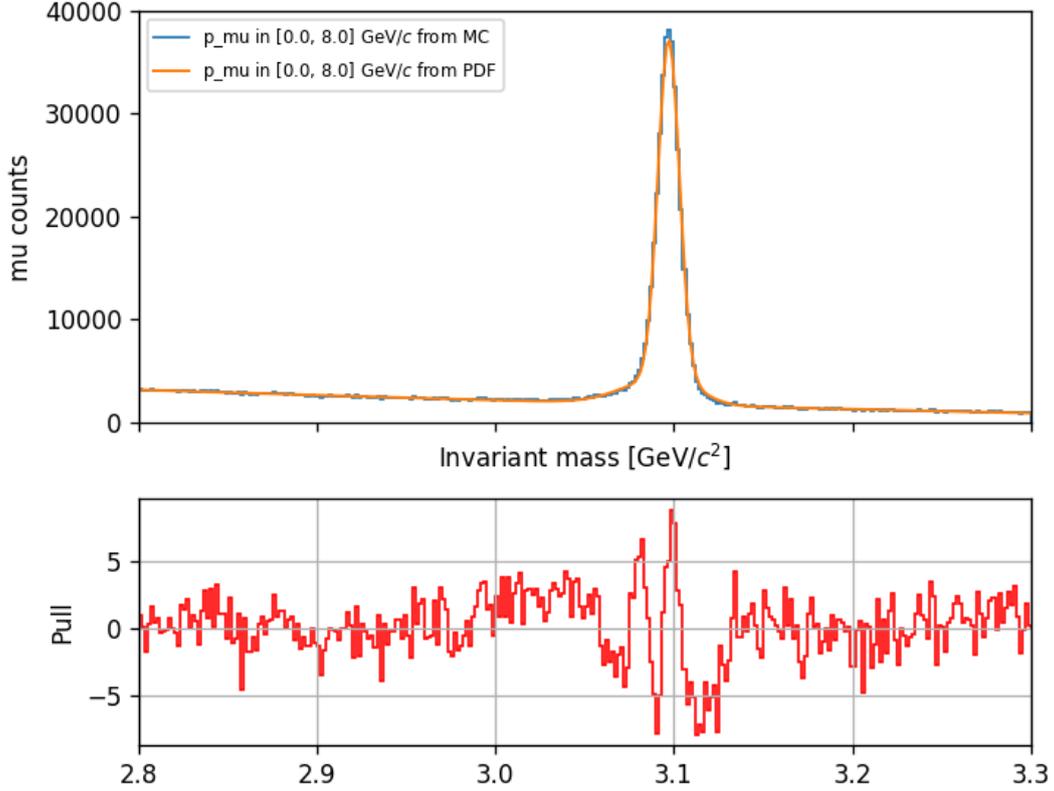
Figure 14: Invariant mass histogram of MC data after the probe cut for probe particle momenta in $[0, 8]$ GeV/c and fitted PDF as well as their pull in the subplot.

Table 7: Resulting PDF parameters after fit to probe cut event histogram.

Parameter	Fit result
N_s	375580
N_b	590041
μ [GeV/c^2]	3.09707
σ_{gaus} [GeV/c^2]	$5.99748 \cdot 10^{-03}$
C_0 [c^2/GeV]	-0.570929
C_1 [c^2/GeV]	0.0278453

An approach to quantify this potential problem is doing an individual calculation of sWeights for the events before and after the probe cut. This results in two sWeights for each event, one to calculate the numerator of the efficiency and one for the denominator. The denominator sWeights are the same as in section 6.2, whereas the numerator sWeights are calculated individually. The numerator sWeights are obtained from an individual PDF fit to the data after the probe cut. Therefore the numerator sWeights only make sense for such events, that remain after the probe cut. The same PDF parametrization as for the uncut data as specified in section 5.4 was used for the fit. The fit result is shown in figure 14 and the resulting parameters are listed in table 7. The PDF fits well to the data. Only in the signal region some discrepancies can be seen in the pull.

The resulting numerators are shown in figure 15. We observe, that the numerator values have indeed lowered for event numbers reconstructed using sWeights from this new fit. The efficiencies obtained for the new numerators are plotted in figure 16. It exhibits, that the efficiencies now fit better to the MC truth in a broad range of momentum from 1.4 GeV/c to 2.3 GeV/c. However outside of the 1.4 GeV/c to 2.3 GeV/c momentum range there is still no agreement between MC truth and sWeight efficiencies. As in the systematic corrections framework from figure 8 the sWeight efficiencies are still to high outside that range.

Therefore there exists a systematic deviation of the numerator counts, that was caused by using the wrong sWeights for the probe cut event numbers and effects the efficiency. This deviation can be resolved by computing individual sWeights for both the numerator and denominator, but does not lead to full agreement between MC truth and sWeight efficiency.
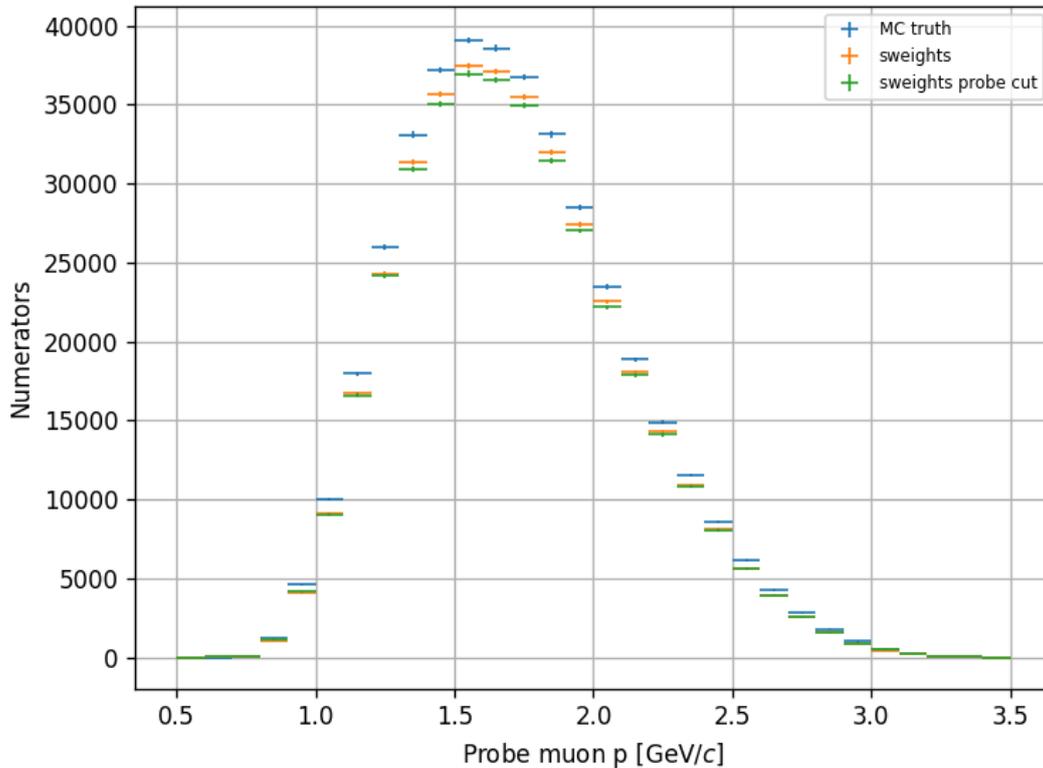
Figure 15: Numerators in momentum bins of the probe particle for the calculation of efficiencies from MC truth and using sWeights for background subtraction with sWeights calculated from the invariant mass distributions before and after the probe cut.

33
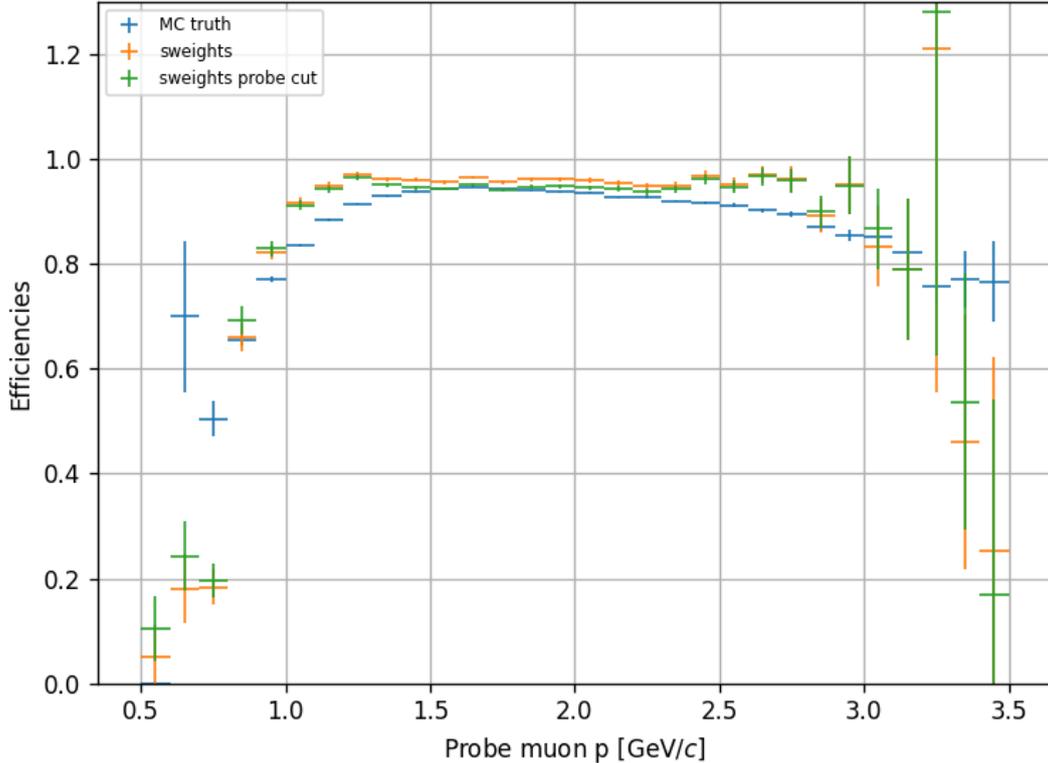
Figure 16: Efficiencies as in figure 12 but including the efficiencies calculated using numerators from probe cut sWeights.

6.4   sWeights in Momentum Bins

A possible explanation of why the efficiencies as seen in figure 8 do not fit the MC truth well in certain momentum ranges is a correlation of the the invariant mass with the momentum of the probe muon. As stated in section 4.5 a correlation between discriminating variable $m_{\mu\mu}$ and control variable probe muon p can cause the distribution reconstructed from sWeights to not be valid. The linear correlation between $m_{\mu\mu}$ and p can be calculated from the data easily. The linear correlation is only 0.04 and will therefore have a negligibly small effect. To visualize possible correlations one can look at invariant mass histograms of the data in bins of the probe muon momentum p. Such histograms are shown in figure 17. The plot exhibits, that the background shape changes in particular for the low momentum range of 0.5 GeV/c to 1.4 GeV/c. Especially the second plot of figure 17, where the normalized distribution is shown, exhibits, that the background for momenta of 0.5 GeV/c to 0.8 GeV/c has a significantly different gradient. This could be a possible higher order correlation that causes the lacking efficiencies.
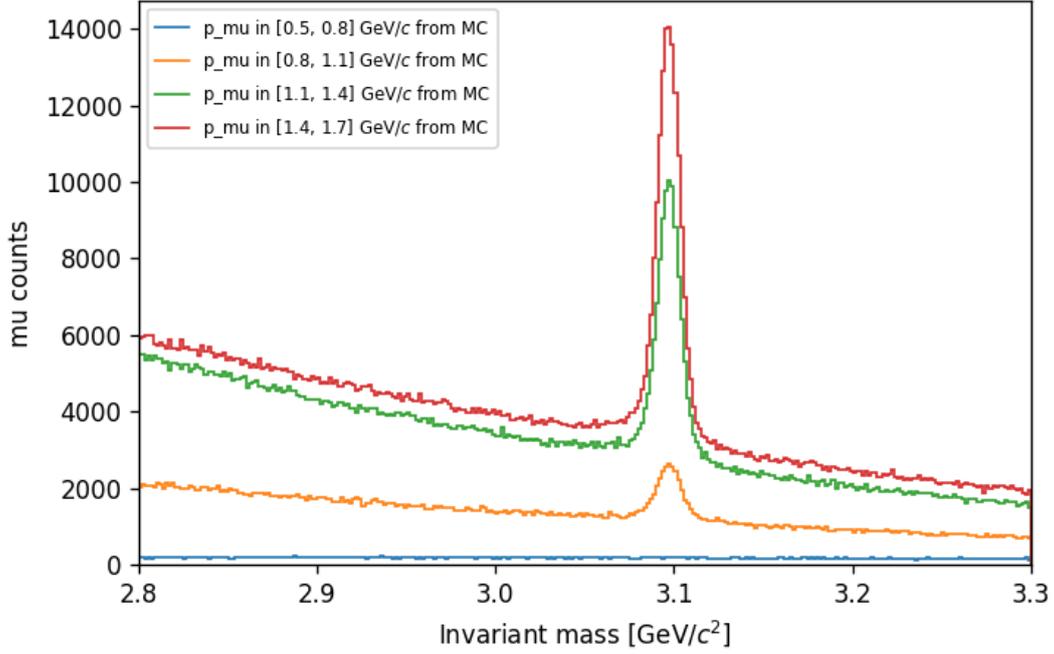
(a) Unnormed histogram
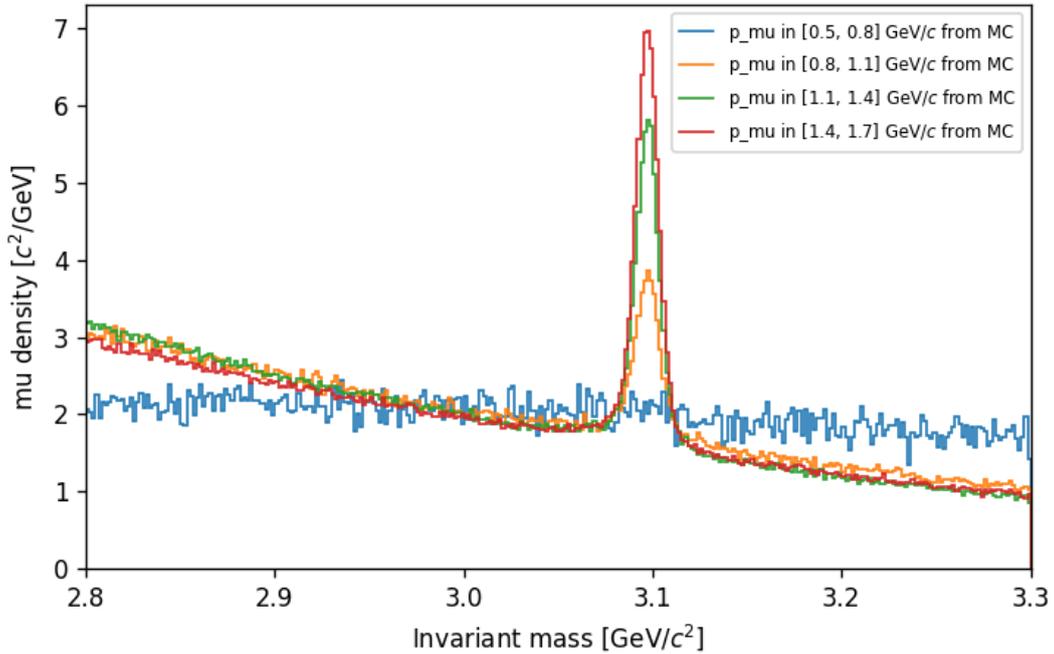
(b) Normed histogram

Figure 17: Invariant mass histogram of MC for the four lowest momentum bins of the probe particle from 0.5 GeV/c to 1.7 GeV/c as absolute counts (a) and normalized to integral one (b) for better comparison of the shapes.

An approach to solve this issue is to calculate the sWeights from an individual PDF for each momentum range. To do this, the data sample was first binned into bins of the momentum of the probe particle. Since for momenta below 0.5 GeV/c and above 3.5 GeV/c there are hardly any data points as well as no J/Ψ

signal peak in the invariant mass variable of the data, only the range in between was considered. A total of ten equidistant bins were used resulting in a bin width of 0.3 GeV/c. The PDF has to be fitted for each bin individually, again twice before and after the probe cut. Since the shape was still roughly the same, the same PDF parametrization with the same starting parameters as described in section 5.4 was used. The resulting fitted PDF, MC data and their pull are shown for four selected of the ten total bins in figure 18 and figure 19. We observe, that the fitted PDF describes the MC data well in all bins. Especially for the low momentum bins the pull does fluctuate around zero and does not exhibit any abnormal rises around the signal peak. Therefore the PDF resembles the data even better than for the fit to the total data set. The fit parameters for all momentum bins are listed in appendix B.
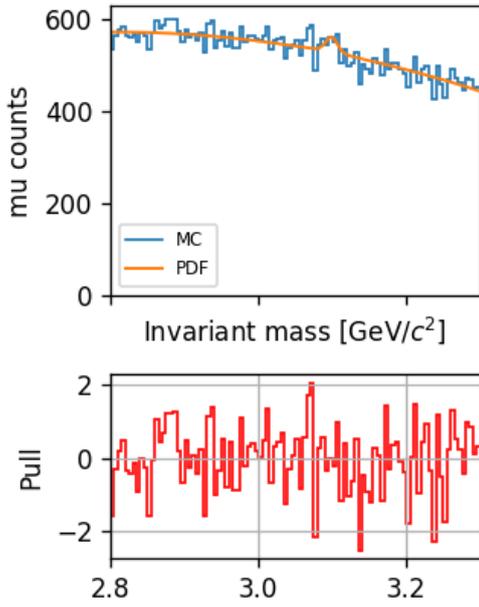
(a) No probe cut, p in $[0.5, 0.8]$ GeV/c
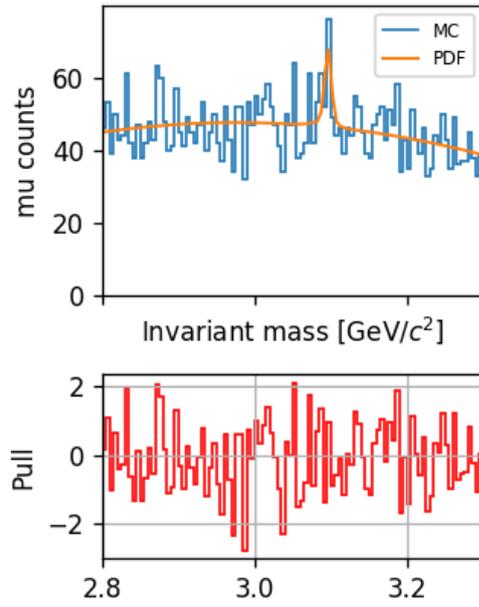
(b) Probe cut, p in $[0.5, 0.8]$ GeV/c
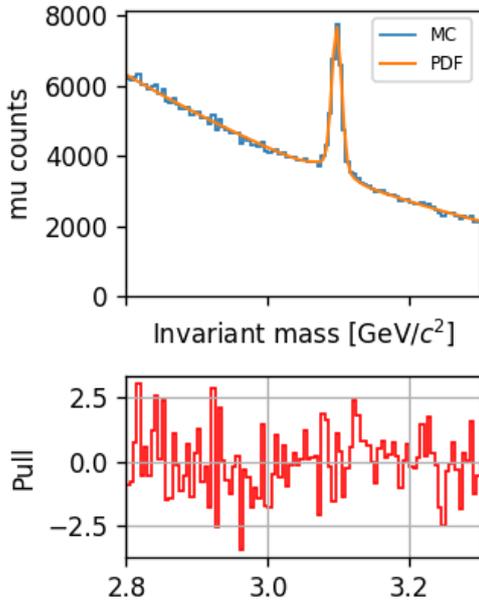
(c) No probe cut, p in $[0.8, 1.1]$ GeV/c
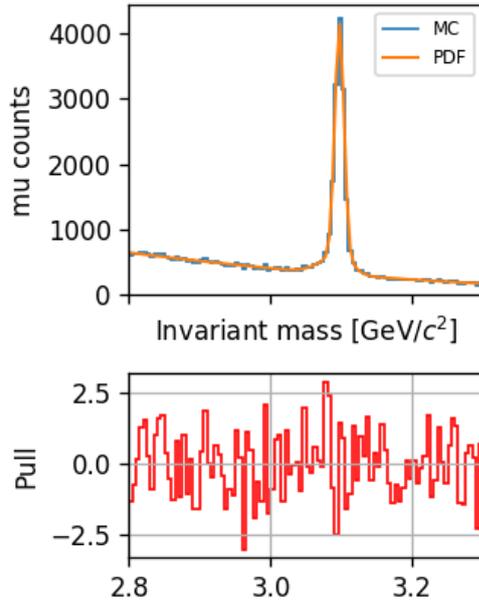
(d) Probe cut, p in $[0.8, 1.1]$ GeV/c

Figure 18: Invariant mass histogram of MC for probe particle momenta in stated bin range with or without probe cut and corresponding fitted PDF as well as their pull in the subplot.
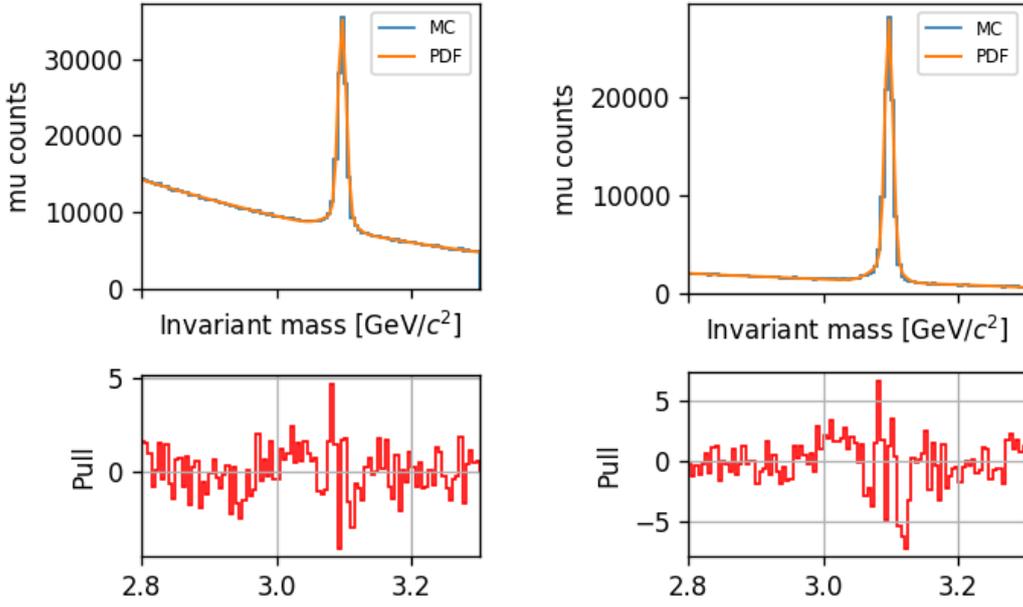
(a) No probe cut, p in $[1.7, 2.0]$ GeV/c

(b) Probe cut, p in $[1.7, 2.0]$ GeV/c
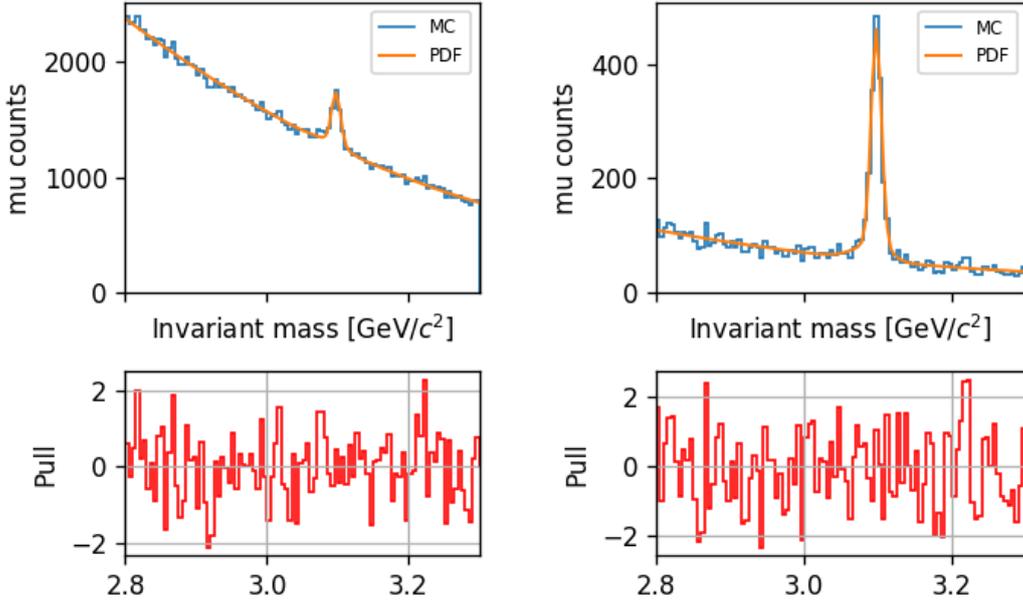
(c) No probe cut, p in $[2.6, 2.9]$ GeV/c

(d) Probe cut, p in $[2.6, 2.9]$ GeV/c

Figure 19: Invariant mass histogram of MC for probe particle momenta in stated bin range with or without probe cut and corresponding fitted PDF as well as their pull in the subplot.

These PDFs are again utilized to calculate sWeights using equation 14. This has to be done by using the fitted PDF from that momentum bin, an event falls into. Again the denominators, numerators and efficiencies can be calculated in probe muon momentum bins. A comparison with the results of the previous approach from section 6.3 is shown in figure 20, figure 21 and figure 22. In figure 20 we observe, that for low momenta between 0.9 GeV/c and 1.4 GeV/c, the

denominators have increased, whereas for higher momenta between 1.4 GeV/c and 2.0 GeV/c, the denominators dropped slightly. The numerators calculated using sWeights after the probe cut as shown in figure 21 on the other hand have decreased for a broad momentum range of 1.0 GeV/c to 2.3 GeV/c. Both numerators and denominators do still not agree with the Monte Carlo truth. The efficiencies are shown in figure 22. The plot exhibits, that the efficiencies hardly change in a momentum range of 1.4 GeV/c to 2.0 GeV/c, but have decreased for momenta between 0.9 GeV/c and 1.4 GeV/c. The efficiencies from sWeights are therefore now closer to the MC truth efficiencies for this low momentum range, but still do not resemble the MC truth.

We see, that it is indeed possible to improve the sWeight efficiency estimates by resolving correlations between the invariant mass of the two particles $m_{\mu\mu}$ and the probe particle momentum p. However the result still does not fully match the MC truth efficiencies. The reason must therefore not just lie in the effect of numerator cuts and correlations of kinematic variables.
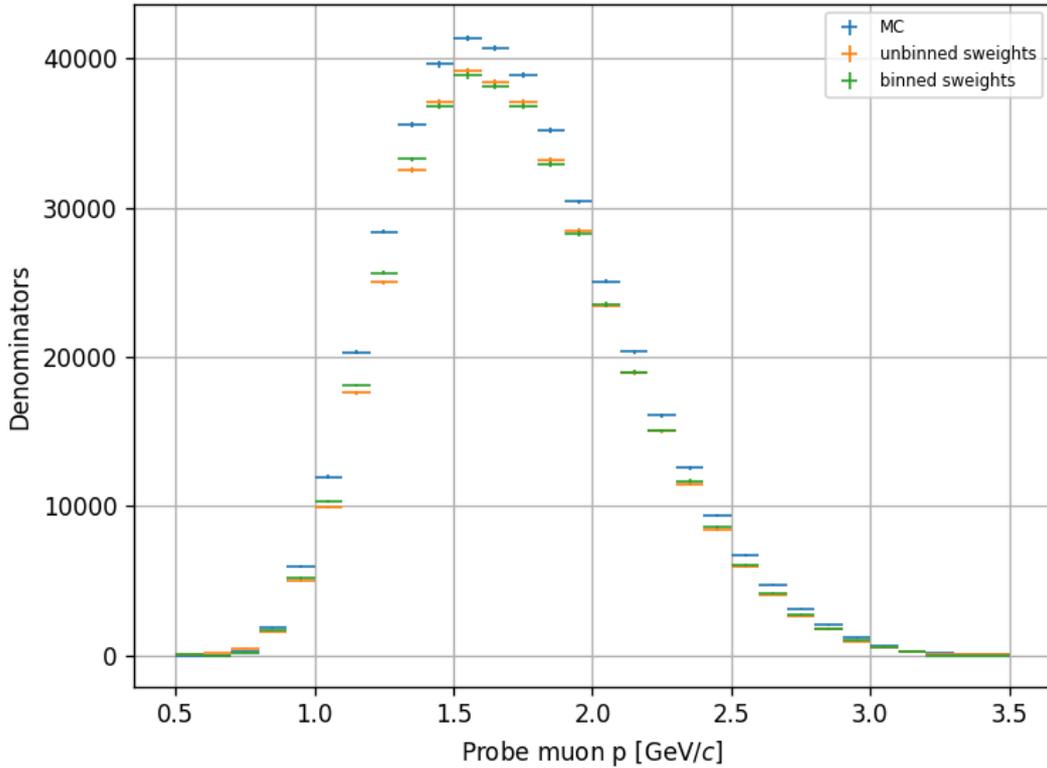
Figure 20: Denominators as in figure 10 including denominators calculated using sWeights calculated in bins of the probe particle momentum.
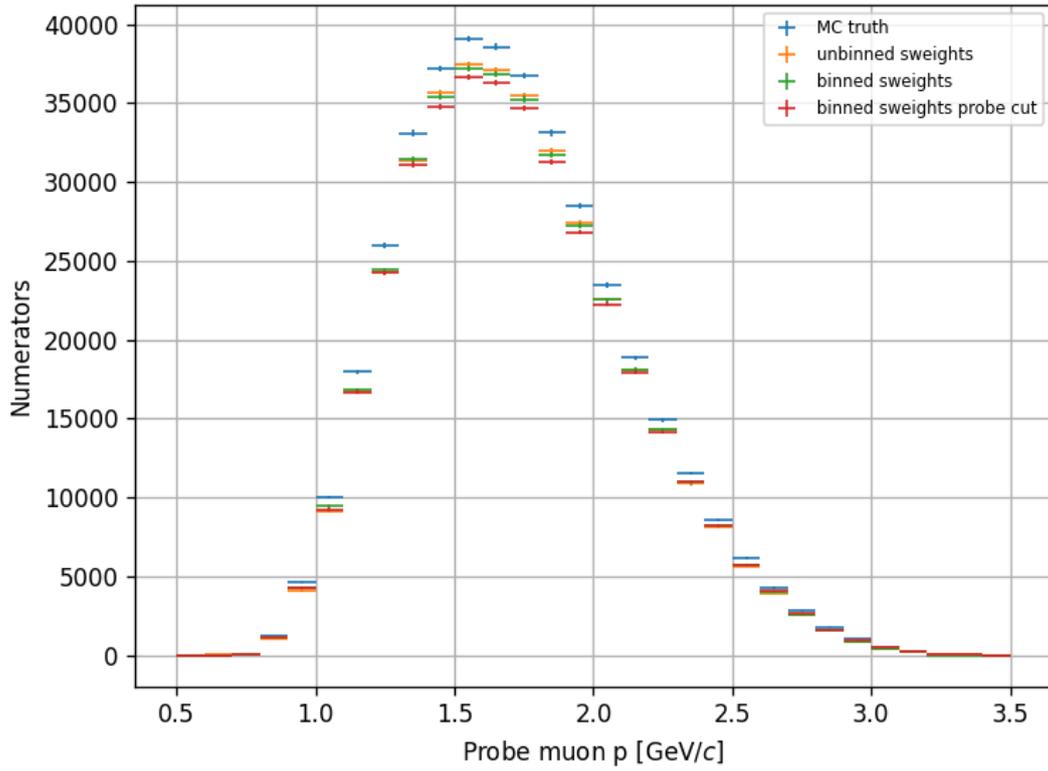
Figure 21: Numerators as in figure 11 including numerators calculated using sWeights calculated in bins of the probe particle momentum.
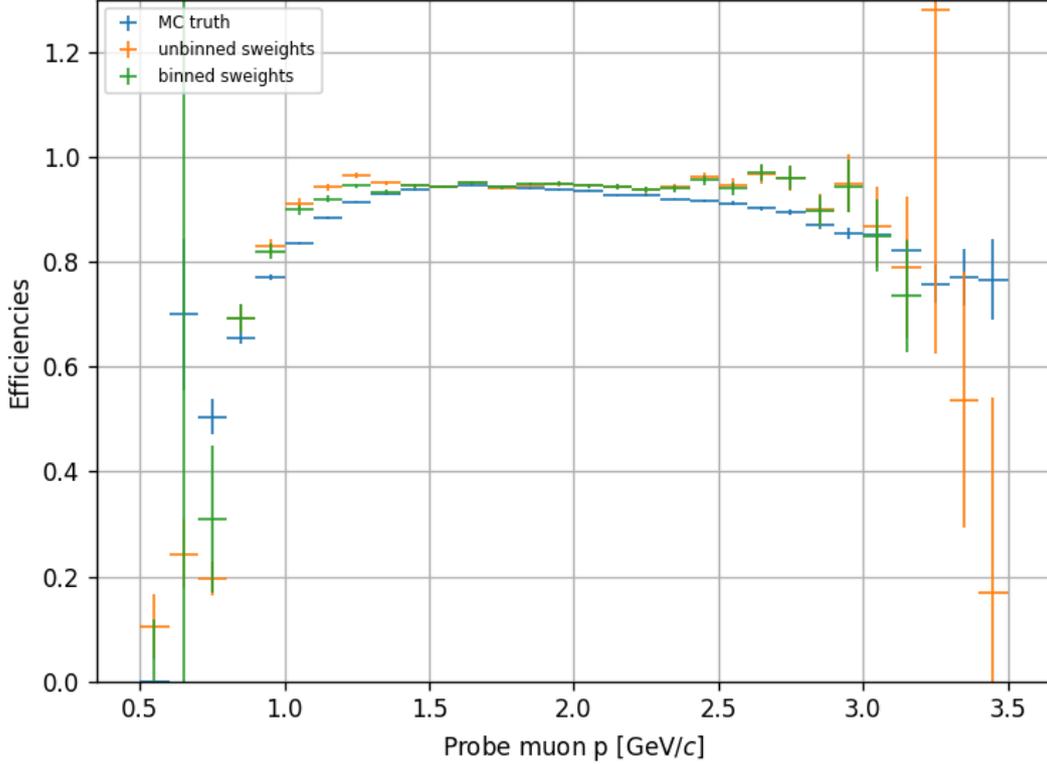
Figure 22: Efficiencies in bins of the probe particle momentum including MC truth efficiencies, efficiencies computed using numerators and denominators from individual sWeights calculated without binning in the momentum and efficiencies computed using numerators and denominators from individual sWeights calculated in bins of the probe particle momentum.

6.5   Mismatch of Signal Counts

In all of the above denominator and numerator figures we observed, that the total number of events calculated from sWeights is systematically lower than that of the MC data. As already stated in section 6.2, this is a result of the yield parameter in the fit, not a deviation that stems directly from sWeights. In this section the effect of the yield parameters will be analyzed in more detail.

From the PDF fit to the data binned in probe muon momentum p, it is possible to compare the MC truth signal counts, signal yield parameters and sWeight sum in momentum bins. In figure 23 those signal counts are plotted. We observe, that signal yield and sWeight sum mostly agree and their relative deviation is mostly far under 1%. The MC truth signal counts however are still systematically about 7% to 15 % higher than the signal yields. Because the deviation between PDF and MC truth counts is much higher than the deviation between PDF and sWeights, this hints on an imperfection in the fit rather than in the calculation of the sWeights.
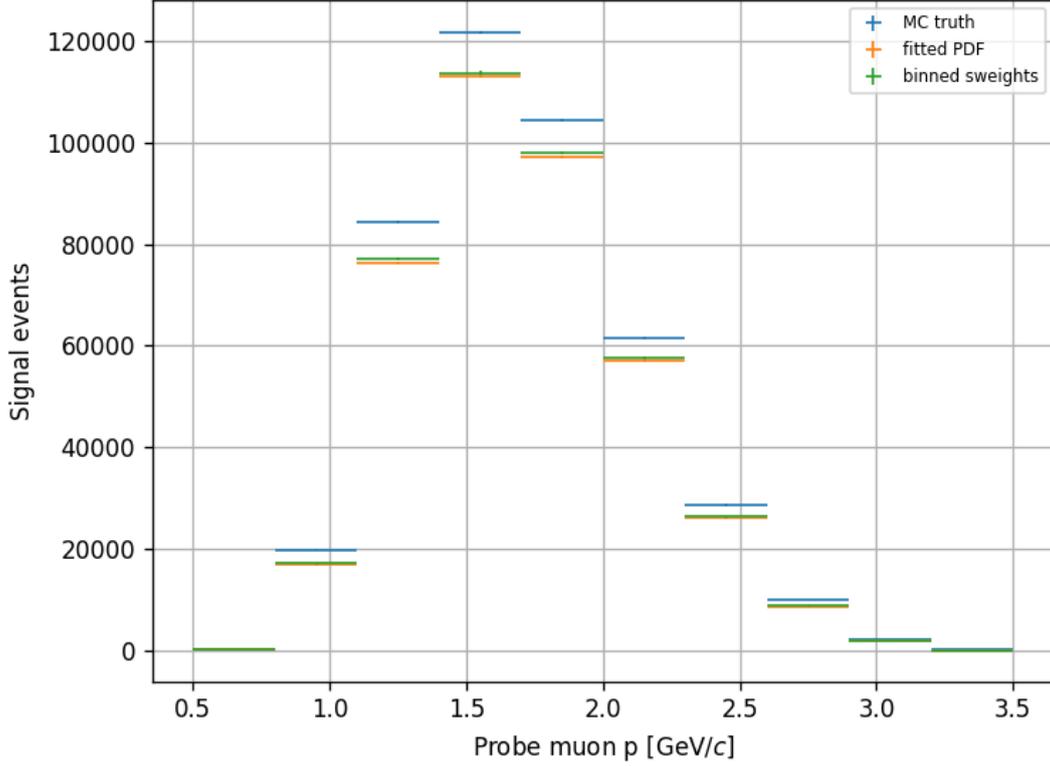
Figure 23: Number of signal events in bins of the probe particle momentum obtained from MC truth, PDF yields and from the sum of sWeights.

Since the number of MC truth events is higher than the PDF signal yield, but the total number of events is the same, there is a leakage of signal events into the background for the PDF fit. To see, if this effect is caused by a specific invariant mass region, in figure 24 the signal part of the PDF and the MC truth signal events are plotted for events from the whole momentum range between 0 GeV/c and 8 GeV/c. For this plot only the signal part of the PDF as described in section 5.4 with the fit parameter results from section 6.2 was used. We observe, that the MC truth signal has a long tail on the left side of the peak, which is not properly described by the signal part of the PDF. Also the signal peak is not fully described. This can be especially seen in the pull in the subplot. To both sides of the signal peak, there is a significant positive pull, which confirms, that the signal PDF underestimates the true distribution.
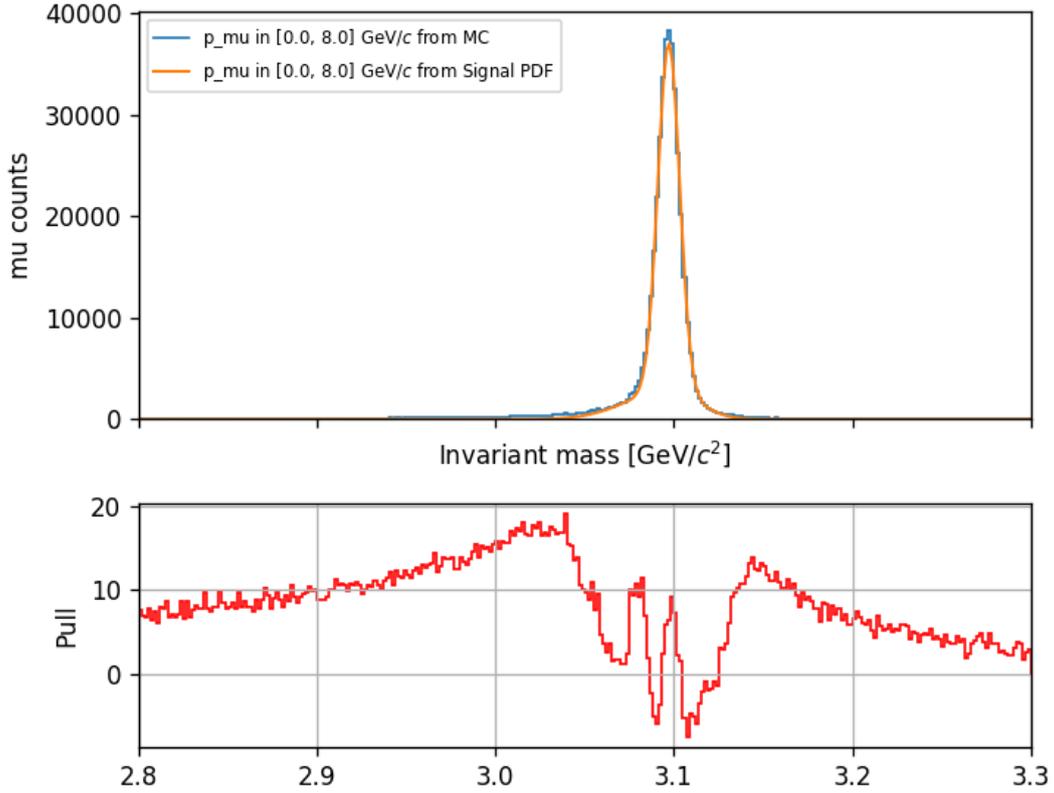
Figure 24: Invariant mass histogram of MC truth signal for probe particle momenta in [0, 8] GeV/c and signal part of the fitted PDF as well as their pull in the subplot.

A cut to tag muon ID values greater than 0.95 instead of 0.9 in the tag cut does not change the agreement between signal PDF and data. Also we tried cutting the invariant mass range of events used in the background region, to remove the signal tails. This can be done either on only one side or on both sides of the peak. However, if we are only using a reduced mass range, the fit is no longer able to determine the curvature of the background properly. This causes even more leakage between signal and background and does not lead to a valid result. Another possibility to improve the PDF fit would be, to again fit the parameters of the PDF, that were originally fixed as described in section 5.4. This would allow the PDF to better describe the signal tails by adjusting the width parameters of the bifurcated Gaussian. However, this also did not improve the agreement of signal data and signal PDF.

7   Conclusion and Outlook

We used MC simulation data to investigate the $J/\Psi \to \mu^+\mu^-$ decay mode and determine muon particle identification efficiencies directly from detector data. We use the $_s\mathcal{P}$lot method to statistically subtract the remaining background events in the selected J/Ψ sample, that could not be rejected in the event selection. For that we construct sWeights for each event, based on a PDF in the invariant mass $m_{\mu\mu}$ of the two particle system, in which signal and background exhibit different distributions. These sWeights should allow us to reconstruct the distribution of signal events in the momentum variable p, which is necessary for muon identification efficiency computation. However, we are unable to reach agreement with the real signal event distribution in the momentum p as known from Monte Carlo truth, as the signal counts after background subtraction are systematically too low. Also we observe, that the efficiencies computed using the $_s\mathcal{P}$lot are generally too high.

To improve the efficiency estimates we calculate different sWeights to obtain the number of signal events and the number of signal events after the probe cut for the efficiency calculation. This results in better agreement with the Monte Carlo truth efficiency for momenta between 1.4 GeV/c and 2.3 GeV/c. We also account for correlations between the invariant mass $m_{\mu\mu}$ and the momentum p, which could affect the validity of the sWeights. This brings only small improvements for the efficiency estimates and hardly any for the event count distribution in the momentum variable p. So accounting for these two effects does improve the resulting efficiency estimates, but is not sufficient to reach full agreement with the Monte Carlo truth efficiencies in the complete momentum range. Only adjusting correlations does not fully solve the problem we observed in the efficiencies.

We also find out, that the PDF parametrization in $m_{\mu\mu}$ used for the $_s\mathcal{P}$lot method does not describe the $m_{\mu\mu}$ distribution of signal events after the cuts explained in section 5.2. Especially the tails to both sides of the J/Ψ signal peak can not be described. This most likely leads to the systematically too low signal event number. Our hypothesis is, that this may also cause the discrepancies in the efficiency, that we still observe.

Future analysis may address this issue via two possible routes. First, one could apply additional cuts on the data set, which remove the tails of the signal. The other possibility is to improve the PDF parametrization. This PDF parametrization should be able to better describe the signal shape, especially in the tail regions. As we have seen in this analysis there exist correlations between kinematic variables in the sample, that have an effect on the efficiency determination. Therefore it will probably be also necessary to use the methods to resolve correlations presented in this thesis again on an analysis with a better PDF parametrization. Finally, the improved background subtraction could be used on real experimental data from the Belle II experiment to compare the results of muon identification efficiencies to those obtained from the Monte Carlo simulation. This enables us to better understand and improve PID at Belle II.

A Uncertainties

In this section the uncertainty propagation for the efficiency will be done at linear order. The efficiency ϵ can be calculated using the number of signal events, that are also passing the probe cut, which is referred to as N_p and the number of all signal events N_a. N_a can be again split up into passing events N_p and non passing events $N_{\bar{p}}$:

$$\epsilon = \frac{N_p}{N_a} = \frac{N_p}{N_p + N_{\bar{p}}}. \tag{18}$$

This trick of rewriting the efficiency as a function of N_p and $N_{\bar{p}}$ makes the calculation of the uncertainties easier because of N_p and $N_{\bar{p}}$ being uncorrelated. This allows us to ignore the otherwise necessary covariance terms between the uncertainties of N_p and N_a. The uncertainties of N_p and $N_{\bar{p}}$ will be denoted as σ_{N_p} and $\sigma_{N_{\bar{p}}}$. The formula for linear uncertainty propagation for the efficiency ϵ then is:

$$\begin{aligned}
\sigma_\epsilon^2 &= \left(\frac{\partial \epsilon}{\partial N_p} \sigma_{N_p} \right)^2 + \left(\frac{\partial \epsilon}{\partial N_{\bar{p}}} \sigma_{N_{\bar{p}}} \right)^2 \\
&= \left(\left(\frac{1}{N_p + N_{\bar{p}}} - \frac{N_p}{(N_p + N_{\bar{p}})^2} \right) \sigma_{N_p} \right)^2 + \left(\frac{N_p}{(N_p + N_{\bar{p}})^2} \sigma_{N_{\bar{p}}} \right)^2.
\end{aligned} \tag{19}$$

To calculate the uncertainties for MC simulated data, we only need to plug in the values for σ_{N_p} and $\sigma_{N_{\bar{p}}}$. As stated in section 5.2 the uncertainty of the count number of a bin can be calculated as the square root of its count value. Using equation 19 we get:

$$\begin{aligned}
\sigma_\epsilon^2 &= \left(\frac{1}{N_s + N_{\bar{p}}} - \frac{N_p}{(N_p + N_{\bar{p}})^2} \right)^2 N_s + \left(\frac{N_p}{(N_p + N_{\bar{p}})^2} \right)^2 N_{\bar{p}} \\
&= \frac{\epsilon}{N_p + N_{\bar{p}}} \left(1 - \frac{N_p}{N_p + N_{\bar{p}}} \right) \\
&= \epsilon^2 \left(\frac{1}{N_p} - \frac{1}{N_a} \right).
\end{aligned} \tag{20}$$

For the sWeights the event numbers are obtained as sum of the sWeight, which will for simplicity also noted using N_p and $N_{\bar{p}}$. As stated in section 16 the standard deviation for the sWeight event number is the square root of the sum of the squares of the sWeights. For simplicity the sWeights will be denoted as $\mathcal{P}$ and the sum will just index, which types of events will be summed. So p means summing the events of all events passing the probe cut and $\bar{p}$ the weights of all non passing events. Plugging all that into 19, we obtain a simplified formula:

$$\begin{aligned}
\sigma_\epsilon^2 &= \left(\frac{1}{N_p + N_{\bar{p}}} - \frac{N_p}{(N_p + N_{\bar{p}})^2} \right)^2 \sum_p \mathcal{P}^2 + \left(\frac{N_p}{(N_p + N_{\bar{p}})^2} \right)^2 \sum_{\bar{p}} \mathcal{P}^2 \\
&= \frac{1}{(N_p + N_{\bar{p}})^4} \left(N_{\bar{p}} \sum_p \mathcal{P}^2 + N_p \sum_{\bar{p}} \mathcal{P}^2 \right).
\end{aligned} \tag{21}$$

The last case we need to determine is that of different sWeights for numerator and denominator of the efficiency. For this case we can no longer use the trick to split into passed and non passed event numbers, since the signal event number in the denominator and in the numerator are the sum of different sWeights. However we do not know their correlation. Since this analysis only aims to estimate the effects, that the corrections done, have on the efficiencies and the uncertainties of the efficiencies are small, it is justified to make an estimate of the uncertainties. We assume, that the signal event number for those efficiencies is the same for numerator and denominator. This allows us to again use equation 21 to calculate the uncertainties for the efficiencies in case of individual sWeights for numerator and denominator.

B Fit Results

In this section the fit results for the parameters of the PDF fitted in bins of the probe muon momentum are listed. There is a total of twenty parameter sets, two for each of the ten momentum bins. The results for the fit before the probe cut are listed in table 8 and those for the fit after the probe cut are listen in table 9.

Table 8: Resulting PDF parameters in bins of the probe particle momentum before probe cut.

Parameter	Result for [0.5, 0.8]	Result for [0.8, 1.1]	Result for [1.1, 1.4]
N_s	155	16861	76172
N_b	52836	394841	957577
μ [GeV/c^2]	3.098989	3.09716	3.09713
σ_{gaus} [GeV/c^2]	0.00765234	0.00675343	0.00612061
C_0 [c^2/GeV]	-0.122792	-0.517511	-0.590550
C_1 [c^2/GeV]	-0.031883	0.0526368	0.0734798

Parameter	Result for [1.4, 1.7]	Result for [1.7, 2.0]	Result for [2.0, 2.3]
N_s	113010	97299	57020
N_b	1096785	881795	623686
μ [GeV/c^2]	3.09709	3.09709	3.09712
σ_{gaus} [GeV/c^2]	0.00597886	0.00598443	0.00610819
C_0 [c^2/GeV]	-0.540896	-0.528848	-0.532402
C_1 [c^2/GeV]	0.0557320	0.0561118	0.0572429

Parameter	Result for [2.3, 2.6]	Result for [2.6, 2.9]	Result for [2.9, 3.2]
N_s	26041	8678	1750
N_b	404996	248536	146114
μ [GeV/c^2]	3.09703	3.09744	3.09765
σ_{gaus} [GeV/c^2]	0.006169819	0.00638223	0.00622218
C_0 [c^2/GeV]	-0.554545	-0.559809	-0.535587
C_1 [c^2/GeV]	0.0704204	0.0605022	0.0585554

Parameter	Result for [3.2, 3.5]
N_s	1
N_b	87349
μ [GeV/c^2]	3.03408
σ_{gaus} [GeV/c^2]	0.0100545
C_0 [c^2/GeV]	-0.479895
C_1 [c^2/GeV]	0.0439832

Table 9: Resulting PDF parameters in bins of the probe particle momentum for fit after the probe cut.

Parameter	Result for [0.5, 0.8]	Result for [0.8, 1.1]	Result for [1.1, 1.4]
N_s	56	14615	71784
N_b	4536	37129	104688
μ [GeV/c^2]	3.09637	3.09710	3.09707
σ_{gaus} [GeV/c^2]	0.00431533	0.00648870	0.00608026
C_0 [c^2/GeV]	-0.070264	-0.624335	-0.611539
C_1 [c^2/GeV]	-0.060084	0.0679278	0.0310872

Parameter	Result for [1.4, 1.7]	Result for [1.7, 2.0]	Result for [2.0, 2.3]
N_s	107467	92533	54191
N_b	144702	131902	90480
μ [GeV/c^2]	3.09705	3.09707	3.09707
σ_{gaus} [GeV/c^2]	0.00587752	0.00587150	0.00601073
C_0 [c^2/GeV]	-0.519476	-0.537071	-0.582552
C_1 [c^2/GeV]	0.0015867	0.0104130	0.0410569

Parameter	Result for [2.3, 2.6]	Result for [2.6, 2.9]	Result for [2.9, 3.2]
N_s	24854	8320	1645
N_b	48265	18194	6616
μ [GeV/c^2]	3.09707	3.09713	3.09760
σ_{gaus} [GeV/c^2]	0.00626672	0.00647046	0.00690280
C_0 [c^2/GeV]	-0.664348	-0.721070	-0.536558
C_1 [c^2/GeV]	0.0671912	0.0928582	0.0784475

Parameter	Result for [3.2, 3.5]
N_s	131
N_b	3513
μ [GeV/c^2]	3.10029
σ_{gaus} [GeV/c^2]	0.00707937
C_0 [c^2/GeV]	-0.436334
C_1 [c^2/GeV]	0.0053636

References

[1] E. Kou et al., "The Belle II Physics Book", Progress of Theoretical and Experimental Physics **2019.12** (2019) 123C01, DOI: 10.1093/ptep/ptz106.

[2] K. Hanagaki et al., "Experimental Techniques in Modern High Energy Physics, A Beginner's Guide", Springer (2022), DOI: 10.1007/978-4-431-56931-2.

[3] M. Milesi et al., "Muon and electron identification performance with 189 fb^{-1} of Belle II data", BELLE2-NOTE-TE-2021-11 (2024), URL: https://docs.belle2.org/record/2340/files/BELLE2-NOTE-TE-2021-011_v6_1.pdf.

[4] M. Pivk and F.R. Le Diberder, "$_s\mathcal{P}lot$: a statistical tool to unfold data distributions", Nucl.Instrum.Meth. **A555** (2005) 356-369, arXiv: physics/0402083.

[5] S. Bilokin, "Systematic Corrections Framework", Software Documentation (2024), URL: https://syscorrfw.readthedocs.io/en/latest/index.html.

[6] D. Griffiths, "Introduction to Elementary Particles", Wiley-VCH Verlag (2008).

[7] J. Sakurai and J. Napolitano, "Modern Quantum Mechanics, Third Edition", Cambridge University Press (2021), DOI: 10.1017/9781108587280.

[8] B. Povh et al., "Particles and Nuclei, An Inroduction to the Physical Concepts", Springer (2008).

[9] S. Navas et al. (Particle Data Group), Phys. Rev. **D 110**, 030001 (2024), DOI: 10.1103/PhysRevD.110.030001.

[10] Ed. A.J. Bevan, B. Golob, Th. Mannel, S. Prell, and B.D. Yabsley, "The Physics of the B-factories", Eur. Phys. J. **C74** (2014) 3026, SLAC-PUB-15968, KEK Preprint 2014-3, arXiv: 1406.6311.

[11] K. Akai et al., "SuperKEKB collider", Nucl. Inst. Meth. in Phys. Res. **A 907** (2018) 188-199, DOI: 10.1016/j.nima.2018.08.017.

[12] T. Hara, T. Kuhr, and Y. Ushiroda, "Belle II Coordinate System and Guideline of Belle II Numbering Scheme", Guideline (2011), URL: https://indico.mpp.mpg.de/event/2308/contributions/4092/attachments/3414/3799/Belle2NumberingScheme.pdf.

[13] Belle II Collaboration, "Belle II Experiment", Accesed from webpage of the KIT (2024), URL: https://www.etp.kit.edu/belle2.php.

[14] H. Svidras, "The Central Drift Chamber of Belle 2", 5th Belle II Starterkit Workshop (2020), URL: https://indico.belle2.org/event/1307/sessions/378/attachments/3070/5652/CDC_B2SK.pdf.

[15] S. Uno, "Belle-II CDC", Tracking meeting Bayrischzell (2012), URL: https://indico.mpp.mpg.de/event/1747/contributions/2050/attachments/1929/2180/BelleII-CDC.pdf.

[16] C. Pulvermacher, "dE /dx Particle Identification and Pixel Detector Data Reduction for the Belle II Experiment", Diplomarbeit am Karlsruher Institut für Technologie (2012), URL: https://docs.belle2.org/record/560/files/BELLE2-MTHESIS-2017-006.pdf.

[17] L. Zani et al., "The Silicon Vertex Detector of the Belle II Experiment", Nucl.Instrum.Meth. A1038 (2022), arXiv:2206.11648.

[18] J. Schmitz, "The New and Complete Belle II DEPFET Pixel Detector: Comissioning and Previous Operational Experience", 13th International 'Hiroshima' Symposium (2023), URL: https://indico.cern.ch/event/1184921/contributions/5576921/attachments/2766920/4819786/72-schmitz-belle_II-pxd.pdf.

[19] O. Hartbrich, "Particle ID in Belle II ARICH and TOP", Belle II Summer Workshop (2019), URL: https://indico.belle2.org/event/1307/sessions/378/attachments/3070/5668/2020_02_01_B2SKW_PID.pdf.

[20] E. Torassa, "TOP detector for particle identification at Belle II", International Workshop on Future Tau Charm Facilities (2024), arXiv: 2406.02935.

[21] Y. Guan, "The TOP Detector for the Belle II Experiment", International workshop on the high energy Circular Electron-Positron Collider (2023), URL: https://indico.ihep.ac.cn/event/19316/contributions/143479/attachments/72850/88879/TOP_BelleII_Guan_v4.pdf.

[22] S. Longo, "Crystal Calorimeters and the Belle II ECL", Belle II Physics Week (2021), URL: https://indico.belle2.org/event/5492/contributions/28667/attachments/14257/21521/Longo_PhysicsWeekECL2021.pdf.

[23] L. Piilonen, "Belle II KLM K_L and Muon Detection", Meeting of the Division of Particles and Fields 2015 (2015), URL: https://indico.cern.ch/event/361123/contributions/856620/attachments/1135126/1625995/KLM-Piilonen.pdf.

[24] Belle II collaboration, "basf2", Software Documentation (2024), URL: https://software.belle2.org/light-2406-ragdoll/sphinx/analysis/doc/Variables.htmlvariable-muonID.

[25] S. Bilokin, "Systematic Corrections Framework", Softwarestack on Gitlab, URL: https://gitlab.desy.de/belle2/performance/systematic_corrections_framework/-/tree/main?ref_type=heads.

[26] R. Brun and F. Rademakers, "ROOT Reference Guide", Software Documentation (2024), URL: https://root.cern.ch/doc/master/group_Roofitmain.html.