

Technical University of Munich  
Department of Physics



**Master Thesis**

# **Measuring the Weinberg Angle at the Belle II Experiment**

Lukas Grubisch  
November 2, 2022

Erstgutachter (Themensteller): Prof. Dr. S. Paul  
Zweitgutachter: Prof. Dr. L. Oberauer

# Abstract

The Weinberg angle is known precisely only at high energies around the  $Z^0$  mass. At Belle II, we have the opportunity to measure it at much lower energies via  $e^+e^- \rightarrow \mu^+\mu^-$  as a process. A measurement could solve the issue of the deviation observed by the NuTeV experiment. I calculate the differential cross section for this process at tree-level to relate the Weinberg angle to the polar angle distribution of the muons. This calculation is used in a dedicated software that I developed to fit the differential cross section to the polar angle distribution and extract the Weinberg angle. I study different background sources and criteria to select signal events and reject background events. By calculating muon detection efficiencies, I study the muon identification performance of Belle II in the process  $e^+e^- \rightarrow \mu^+\mu^-\gamma$  both in recorded data and using simulation. The ratio of the efficiencies in recorded data and simulation allows us to correct differences between recorded and simulated data and provides systematic uncertainties related to these corrections.

# Contents

<b>Abstract</b>	<b>i</b>
<b>1 Introduction</b>	<b>1</b>
1.1 The Weinberg angle $\theta_W$ . . . . .	1
<b>2 The Weinberg angle in <math>e^+e^- \rightarrow \mu^+\mu^-</math></b>	<b>4</b>
2.1 The differential cross section . . . . .	4
2.2 Weinberg angle fit . . . . .	10
<b>3 Study of background sources</b>	<b>16</b>
3.1 SuperKEKB and the Belle II detector . . . . .	17
3.2 Study of $\mu^+\mu^-X$ kinematics . . . . .	18
3.3 Calculating the selection efficiency . . . . .	25
3.4 Optimizing the radial energy deviation selection . . . . .	27
3.5 Data composition . . . . .	30
3.5.1 Ratio of two efficiencies . . . . .	30
3.5.2 Expected purity with $R_{\mu\mu}$ selection . . . . .	32
<b>4 Study of the muon identification performance</b>	<b>41</b>
4.1 The Belle II charged particle identification . . . . .	41
4.2 The track isolation . . . . .	42
4.3 Selection criteria . . . . .	43
4.4 The muon identification performance . . . . .	44
<b>5 Conclusion</b>	<b>58</b>

# Chapter 1

## Introduction

### 1.1 The Weinberg angle $\theta_W$

The Standard Model of particle physics is in large parts based on the theory of the electroweak unification proposed by Glashow [1], Weinberg [2] and Salam [3]. This relates the photon and the gauge bosons of the weak interaction and introduces the Weinberg angle, which is of interest to us [4].

The leptons and quarks are charged under the weak isospin and the weak hypercharge  $Y$ . The  $SU(2)$  symmetry of the weak isospin has three massless gauge bosons  $W^1$ ,  $W^2$  and  $W^3$ , which couple with the coupling strength  $g$ . The  $U(1)_Y$  symmetry has a single massless gauge boson  $B$ , which couples with the coupling strength  $g'$ .

The Higgs-mechanism [5, 6] breaks the initial symmetries and gives mass to three linear combinations of the initial gauge bosons.  $W^1$  and  $W^2$  combine to form the  $W$ -bosons  $W^+$  and  $W^-$  with mass  $M_W$ . The  $W^3$  and  $B$  mix together to form the  $Z^0$ -boson with mass  $M_Z$  and the massless photon  $\gamma$ . This mixing can be described as the rotation of the initial gauge bosons into the set of  $\gamma$  and  $Z^0$

$$\begin{pmatrix} Z^0 \\ \gamma \end{pmatrix} = \begin{pmatrix} \cos \theta_W & -\sin \theta_W \\ \sin \theta_W & \cos \theta_W \end{pmatrix} \begin{pmatrix} W^3 \\ B \end{pmatrix}. \quad (1.1)$$

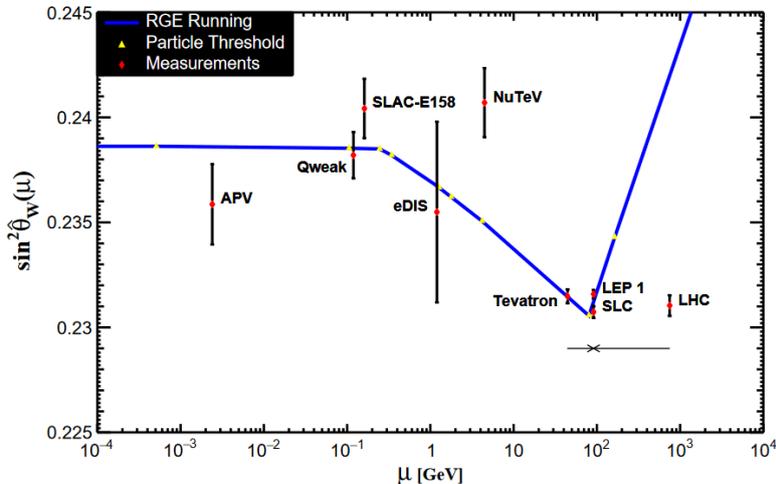
This defines the Weinberg angle  $\theta_W$  as the angle by which the initial states  $B$  and  $W^3$  are rotated to form the new basis  $\gamma$  and  $Z^0$ . As a consequence, the Weinberg angle also relates the coupling couplings  $g, g'$  through

$$\cos \theta_W = \frac{g}{\sqrt{g^2 + g'^2}}. \quad (1.2)$$

One way to measure the Weinberg angle is by measuring the ratio of the  $W$  and  $Z^0$  masses

$$\cos \theta_W = \frac{M_W}{M_Z}. \quad (1.3)$$

The value of the Weinberg angle depends on the energy scale at which it is measured [7, 8]. Figure 1.1, taken from [9], shows this behavior for a large energy



**Figure 1.1:** The running of  $\sin^2 \theta_W$  as a function of the energy scale  $\mu$  in the  $\overline{MS}$  renormalization scheme. Plot taken from [9], using [7, 22]. The measurements of LHC and Tevatron have been shifted to the sides (real position indicated by the arrows) to improve visibility.

range. It also shows different measurements, which are used to constrain the theory curve in blue. Above 10 GeV, collider experiments measure the Weinberg angle with high precision using the angular asymmetry of lepton pairs. In hadron colliders like Tevatron [10] and LHC [11, 12, 13], measuring the angular asymmetry is more challenging, since the definition of the directions is not as easy as for electron-positron colliders like LEP [14] and SLC [15].

Low energy experiments below 10 GeV have a lower precision. APV [16] uses atomic parity violation in different nuclei. SLAC-E158 [17] measures the parity violating asymmetry in fixed target polarized Møller scattering. Qweak [18] uses a similar setup measuring the parity violating asymmetry in polarized electron-positron scattering. eDis [19] uses deep inelastic scattering with polarized electrons and deuterons to measure a left-right asymmetry. NuTeV [20] uses a technique based on the measurement of the cross sections for neutral current and charged current cross sections using a neutrino beam scattering on a fixed target. The NuTeV result is displayed at roughly the average momentum transfers used.

The NuTeV result is the only one in significant disagreement with the theory prediction. However, A. Thomas and others [21] claim that the original analysis neglected several corrections to their result. Taking these corrections into account would allegedly resolve the tension.

Due to the low precision of the low energy measurements, the theory prediction is determined by the high-energy measurements. The value of the Weinberg angle at lower energies is therefore also determined by these measurements. However, if we do not want to rely solely on the predictions of the Standard Model, additional measurements at different energies are needed, to test if the theory matches reality.

Belle II and SuperKEKB operate at KEK in Tsukuba, Japan [23]. SuperKEKB is an electron-positron collider with a center-of-mass energy of 10.58 GeV. This is far below the  $Z^0$  mass (about 90 GeV) and in the region of the NuTeV measurement. At Belle II, the Weinberg angle can be measured using the angular distribution of leptons in  $e^+e^- \rightarrow \ell^+\ell^-$  events. This requires understanding of how well leptons are measured at Belle II. For this reason, I helped the Charged Particle Identification group at Belle II to understand the muon identification. Measuring the Weinberg angle with different final state leptons allows to test for lepton universality, a key ingredient of the Standard Model.

## Chapter 2

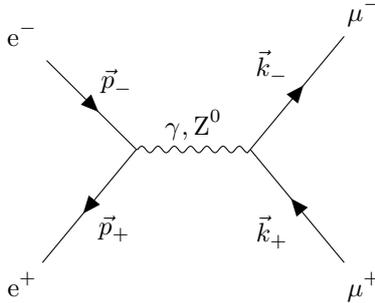
# The Weinberg angle in $e^+e^- \rightarrow \mu^+\mu^-$

### 2.1 The differential cross section

In general, the final state of two particles is described by eight real numbers, namely the two four-momenta of the particles. At a fixed center-of-mass energy  $\sqrt{s}$ , four of these numbers are fixed by energy and momentum conservation, so there are four degrees of freedom. We study processes with leptons in the final state, so the four-momentum relation  $p^2 = m_\ell^2$  fixes two further degrees of freedom. There are only two degrees of freedom left. The two degrees of freedom are chosen to be the polar angle and the azimuth of the  $\mu^-$ . The cross section for  $e^+e^- \rightarrow \ell^+\ell^-$  is independent of the azimuth and the polar angle distribution is used to extract the Weinberg angle from it. Therefore, I calculate the differential cross section with respect to the cosine of the polar angle  $\alpha$ . The polar angle is defined as the angle between the direction of the final state lepton with respect to the same-signed beam electron in the center-of-mass frame. In the center-of-mass frame, this definition of the polar angle has the same value for both charges since the electrons and final state leptons are each back-to-back.

Out of the three leptons, the electron, muon and tau, the muon is best suited for such a measurement at Belle II. An electron final state has two important issues. First, electrons lose more energy through photon radiation while traveling through the detector. To measure the polar angle, we need to know the direction of the final state electrons, so one would need to correct for this. Second, the final and initial state particles are identical so in addition to the  $s$ -channel diagram depicted in fig. 2.1, also a  $t$ -channel diagram contributes [14]. For the muon- and tau-pair production, this diagram does not exist in the Standard Model. While technically more difficult, measuring the Weinberg angle with an electron final state might be interesting, depending on the structure of the  $t$ -channel contribution.

Measuring the Weinberg angle using taus has the issue that the tau decays quickly. Therefore, only its decay products can be measured in the detector. Lepton number conservation implies that at least one of the tau decay products has to be a neutrino, which is not measured in the detector. Therefore, the



**Figure 2.1:** Feynman diagram of  $e^+e^- \rightarrow \mu^+\mu^-$  via  $\gamma$  or  $Z^0$  exchange.

direction of the tau is only approximately known, which makes the measurement of the Weinberg angle more difficult.

The muon is stable as far as the Belle II detector is concerned. Also, compared to the electron, its higher mass means that photon radiation is not as much of an issue as for the electron. Therefore, it is the best candidate and I focus on the process  $e^+e^- \rightarrow \mu^+\mu^-$  as shown in fig. 2.1. This process is most likely mediated by a photon  $\gamma$ , but it can also proceed through exchange of a  $Z^0$ -boson

$$e^+e^- \rightarrow \gamma/Z^0 \rightarrow \ell^+\ell^-. \quad (2.1)$$

Quantum mechanically, the amplitudes of photon or  $Z^0$  exchange interfere with each other.

Therefore, the amplitude of the  $Z^0$  alters the polar angle distribution of the lepton pair compared to it being produced only through  $\gamma$  exchange. If only photon exchange is considered, the polar angle distribution is symmetric at tree-level. However, through the inclusion of the  $Z^0$  exchange, the polar angle distribution acquires an asymmetry. We can relate this asymmetry to the Weinberg angle, which appears in the coupling of the  $Z^0$  to leptons.

For now, I work in the idealized scenario, where the momentum of the electrons and muons is known exactly. Their direction is defined by their normalized momentum. With this setup, I calculate the differential cross section for muon pair production

$$\frac{d\sigma}{d\cos\alpha} = \Pi_{\text{PS}}|\mathcal{M}|^2, \quad (2.2)$$

with the phase-space factor  $\Pi_{\text{PS}}$  and the total squared amplitude  $|\mathcal{M}|^2$ . The total squared amplitude can be separated into different parts

$$\begin{aligned} |\mathcal{M}|^2 &= |\mathcal{M}_\gamma + \mathcal{M}_Z|^2 \\ &= |\mathcal{M}_\gamma|^2 + 2|\mathcal{M}_{\text{int}}|^2 + |\mathcal{M}_Z|^2. \end{aligned} \quad (2.3)$$

The amplitude for photon exchange is  $\mathcal{M}_\gamma$ , the amplitude for  $Z^0$ -boson exchange is  $\mathcal{M}_Z$  and the interference amplitude between both is  $\mathcal{M}_{\text{int}}$ .

A simplified way to extract the Weinberg angle from the differential cross section is the forward-backward asymmetry  $A_{\text{FB}}$ , which is often also reported

for such measurements. It is defined as

$$A_{\text{FB}} = \frac{\int_0^1 \frac{d\sigma}{d\cos\alpha} d\cos\alpha - \int_{-1}^0 \frac{d\sigma}{d\cos\alpha} d\cos\alpha}{\int_0^1 \frac{d\sigma}{d\cos\alpha} d\cos\alpha + \int_{-1}^0 \frac{d\sigma}{d\cos\alpha} d\cos\alpha}. \quad (2.4)$$

The phase-space factor drops out since it is independent of  $\cos\theta$ . Hence the forward-backward asymmetry in terms of the total amplitude is

$$A_{\text{FB}} = \frac{\int_0^1 |\mathcal{M}|^2 d\cos\alpha - \int_{-1}^0 |\mathcal{M}|^2 d\cos\alpha}{\int_0^1 |\mathcal{M}|^2 d\cos\alpha + \int_{-1}^0 |\mathcal{M}|^2 d\cos\alpha}. \quad (2.5)$$

This is equivalent to counting the number of muons in different  $\cos\alpha$  regions

$$A_{\text{FB}} = \frac{N_{\text{F}} - N_{\text{B}}}{N_{\text{F}} + N_{\text{B}}}. \quad (2.6)$$

$N_{\text{F}}$  is the number of muons with a positive value of  $\cos\alpha$  and  $N_{\text{B}}$  the number of muons with a negative value of  $\cos\alpha$ . We could thus be content with only classifying and counting the muons and calculating the forward-backward asymmetry. This is then related to the Weinberg angle.

Instead, we opt for a different approach, where we calculate the squared total amplitude and use it in a maximum likelihood fit to the polar angle distribution. We should then be able to directly extract the Weinberg angle from the polar angle distribution. One reason for the approach with the maximum likelihood fit was that we initially wanted to continue this study by later including detector effects. This work is left for future studies.

I use the momenta as defined in fig. 2.1 to calculate the total squared amplitude at tree-level. The square of the center-of-mass energy is  $s \equiv |p_- + p_+|^2 = |k_- + k_+|^2$ . In the center-of-mass frame, energy and momentum conservation require the four leptons to have the same energy  $\sqrt{s}/2$  and the lepton pairs have to be back-to-back. The normalized momentum of the electron defines the  $z$ -direction. The magnitude of the momentum of the leptons is then given by  $|p_{\pm}| = |k_{\pm}| = \frac{\sqrt{s}}{2} z_{\ell}$  with

$$z_{\ell} \equiv \sqrt{1 - \frac{4m_{\ell}^2}{s}}. \quad (2.7)$$

In the limit  $m_{\ell}^2 \ll s$ ,  $z_{\ell} \approx 1$ . This is a good approximation at Belle II, since at  $\sqrt{s} = 10.58 \text{ GeV}$ ,  $z_e$  deviates from 1 only at the ninth digit and  $z_{\mu}$  at the fourth digit. The product of four momenta are

$$p_- p_+ = \frac{s}{4} (1 + z_e^2), \quad (2.8)$$

$$k_- k_+ = \frac{s}{4} (1 + z_{\mu}^2), \quad (2.9)$$

$$p_- k_{\mp} = p_+ k_{\pm} = \frac{s}{4} (1 \mp z_e z_{\mu} \cos\alpha). \quad (2.10)$$

To calculate the amplitude corresponding to fig. 2.1, I introduce the spinors  $u_{e,\mu}$  and  $\bar{u}_{e,\mu}$  for the  $e^-$  and  $\mu^-$  and the spinors  $v_{e,\mu}$  and  $\bar{v}_{e,\mu}$  for the  $e^+$  and  $\mu^+$ . The spinors obey the Dirac equation [4]. In the case of the electron

$$(\not{p}_- - m_e)u_e = \bar{u}_e(\not{p}_- - m_e) = 0, \quad (2.11)$$

$$(\not{p}_+ + m_e)v_e = \bar{v}_e(\not{p}_+ - m_e) = 0. \quad (2.12)$$

The muon spinors obey the same equations with the respective momentum and mass replacements. Here the Dirac slash notation  $\not{q} = \gamma^\nu q_\nu$  is used.

The photon propagator is  $-ig_{\mu\nu}/(q^2 + i\epsilon)$  with the square of the four momentum transfer  $q^2 = s$ . The  $Z^0$  propagator is

$$\frac{-i}{q^2 - M_Z^2 + iM_Z\Gamma_Z} \left( g_{\mu\nu} - \frac{q_\mu q_\nu}{M_Z^2} \right) \quad (2.13)$$

and takes into account the mass  $M_Z$  and width  $\Gamma_Z$  of the  $Z^0$ -boson [9]

$$M_Z = (91.1876 \pm 0.0021) \text{ GeV} \text{ and } \Gamma_Z = (2.4952 \pm 0.0023) \text{ GeV}. \quad (2.14)$$

The photon-lepton vertex is  $ie\gamma^\nu$ , with the electron charge  $e$ . Assuming lepton flavor universality, the  $Z^0$ -lepton vertex is

$$\frac{-ie}{\sin 2\theta_W} \gamma^\nu (c_V - c_A \gamma^5) \quad (2.15)$$

with  $c_V = -\frac{1}{2} + 2\sin^2\theta_W$  and  $c_A = -\frac{1}{2}$ .

This defines the set of Feynman rules we use. The squared amplitude for the photon exchange is then

$$|\mathcal{M}_\gamma|^2 = \frac{e^4}{s^2} [\bar{v}_e \gamma^\nu u_e \bar{u}_\mu \gamma_\nu v_\mu] \cdot [\bar{v}_\mu \gamma_\rho u_\mu \bar{u}_e \gamma^\rho v_e]. \quad (2.16)$$

At Belle II, the beams are unpolarized and the spins of the final state muons are also not measured. Therefore, the total squared amplitude averages over the initial state spins and sums the final state spins using the spin sums

$$\sum u_e \bar{u}_e = \not{\not{p}}_- + m_e \text{ and } \sum v_e \bar{v}_e = \not{\not{p}}_+ - m_e. \quad (2.17)$$

The muon spin sums take the same form with the corresponding momentum and mass replacements. Spin averaged and summed squared amplitudes are marked with an overline. The squared amplitude for the photon exchange is then

$$\overline{|\mathcal{M}_\gamma|^2} = \frac{e^4}{4s^2} \text{Tr} \left[ (\not{\not{p}}_+ - m_e) \gamma^\nu (\not{\not{p}}_- + m_e) \gamma^\rho \right] \text{Tr} \left[ (\not{\not{k}}_- + m_\tau) \gamma_\nu (\not{\not{k}}_+ - m_\tau) \gamma_\rho \right]. \quad (2.18)$$

Using the trace relations in [4], the squared photon amplitude is

$$\overline{|\mathcal{M}_\gamma|^2} = e^4 \left[ 2 + z_e^2 z_\mu^2 (-1 + \cos^2 \alpha) + (1 - z_e^2)(1 - z_\mu^2) \right]. \quad (2.19)$$

Similarly, the squared amplitude for  $Z^0$  exchange is

$$|\mathcal{M}_Z|^2 = A \left[ \bar{u}_\mu \gamma^\omega C^- v_\mu \left( g_{\omega\nu} - \frac{q_\omega q_\nu}{M_Z^2} \right) \bar{v}_e \gamma^\nu C^- u_e \right. \\ \left. \cdot \bar{u}_e \gamma^\rho C^- v_e \left( g_{\rho\sigma} - \frac{q_\rho q_\sigma}{M_Z^2} \right) \bar{v}_\mu \gamma^\sigma C^- u_\mu \right], \quad (2.20)$$

with

$$A = \frac{e^4}{\sin^4 2\theta_W} \frac{1}{(s - M_Z^2)^2 + M_Z^2 \Gamma_Z^2}. \quad (2.21)$$

I introduce the notation

$$C^\pm \equiv c_V \pm c_A \gamma^5 \quad (2.22)$$

$$\text{with } \gamma^\nu C^\pm = C^\mp \gamma^\nu. \quad (2.23)$$

Using  $\not{q} = \not{k}_- + \not{k}_+ = \not{p}_- + \not{p}_+$  and the Dirac equation on the terms containing  $\not{q}$  yields

$$\begin{aligned} \bar{v}_e \not{q} C^- u_e &= -m_e \bar{v}_e C^- u_e + m_e \bar{v}_e C^+ u_e \\ &= 2m_e c_A \bar{v}_e \gamma^5 u_e \end{aligned} \quad (2.24)$$

and similarly

$$\bar{u}_e \not{q} C^- v_e = -2m_e c_A \bar{u}_e \gamma^5 v_e. \quad (2.25)$$

The terms containing the muon spinors and  $\not{q}$  are identical, replacing all references to electrons by muons. The squared amplitude for  $Z^0$  exchange decomposes into four contributions

$$\begin{aligned} |\mathcal{M}_Z|^2 &= 4A \left[ \underbrace{\left( u_\mu \bar{u}_\mu \gamma^\omega C^- v_\mu \bar{v}_\mu \gamma_\rho C^- \right) \left( u_e \bar{u}_e \gamma^\rho C^- v_e \bar{v}_e \gamma_\omega C^- \right)}_{W_1} \right. \\ &\quad + a \underbrace{\left( u_\mu \bar{u}_\mu \gamma^\omega C^- v_\mu \bar{v}_\mu \gamma^5 \right) \left( u_e \bar{u}_e \gamma^5 v_e \bar{v}_e \gamma_\omega C^- \right)}_{W_2} \\ &\quad + a \underbrace{\left( u_\mu \bar{u}_\mu \gamma^5 v_\mu \bar{v}_\mu \gamma_\rho C^- \right) \left( u_e \bar{u}_e \gamma^\rho C^- v_e \bar{v}_e \gamma^5 \right)}_{W_3} \\ &\quad \left. + b \underbrace{\left( u_\mu \bar{u}_\mu \gamma^5 v_\mu \bar{v}_\mu \gamma^5 \right) \left( u_e \bar{u}_e \gamma^5 v_e \bar{v}_e \gamma^5 \right)}_{W_4} \right]. \end{aligned} \quad (2.26)$$

Again averaging over the initial state and summing over final state spins, the different terms are

$$W_1 = \text{Tr} \left[ (\not{k}_- + m_\mu) \gamma^\omega C^- (\not{k}_+ - m_\mu) \gamma_\rho C^- \right] \text{Tr} \left[ (\not{p}_- + m_e) \gamma^\rho C^- (\not{p}_+ - m_e) \gamma_\omega C^- \right], \quad (2.27)$$

$$W_2 = \text{Tr} \left[ (\not{k}_- + m_\mu) \gamma^\omega C^- (\not{k}_+ - m_\mu) \gamma^5 \right] \text{Tr} \left[ (\not{p}_- + m_e) \gamma^5 (\not{p}_+ - m_e) \gamma_\omega C^- \right] \quad (2.28)$$

$$= W_3, \quad (2.29)$$

$$W_4 = \text{Tr} \left[ (\not{k}_- + m_\mu) \gamma^5 (\not{k}_+ - m_\mu) \gamma^5 \right] \text{Tr} \left[ (\not{p}_- + m_e) \gamma^5 (\not{p}_+ - m_e) \gamma^5 \right]. \quad (2.30)$$

Using the trace relations,  $W_1$  is

$$W_1 = 4s^2 \left[ \left( c_V^2 + c_A^2 \right)^2 \left( 1 + z_e^2 z_\mu^2 \cos^2 \alpha \right) + \left( c_V^4 - c_A^4 \right) \left( 1 - z_e^2 z_\mu^2 \right) \right. \\ \left. + \left( c_V^2 - c_A^2 \right)^2 \left( 1 - z_e^2 \right) \left( 1 - z_\mu^2 \right) + 8c_V^2 c_A^2 z_e z_\mu \cos \alpha \right]. \quad (2.31)$$

$W_2$  and  $W_3$  are  $-16m_e m_\mu c_A^2 s$  and  $W_4$  is  $4s^2$ . Putting all together, the squared amplitude for  $Z^0$ -exchange is

$$\overline{|\mathcal{M}_Z|^2} = A s^2 \left[ \left( c_V^2 + c_A^2 \right)^2 \left( 1 + z_e^2 z_\mu^2 \cos^2 \alpha \right) \right. \\ \left. + \left( c_V^2 - c_A^2 \right)^2 \left( 1 - z_e^2 \right) \left( 1 - z_\mu^2 \right) + \left( c_V^4 - c_A^4 \right) \left( 1 - z_e^2 z_\mu^2 \right) \right. \\ \left. - \frac{s}{M_Z^2} c_A^4 \left( 1 - z_e^2 \right) \left( 1 - z_\mu^2 \right) \left( 2 - \frac{s}{M_Z^2} \right) + 8c_V^2 c_A^2 z_e z_\mu \cos \alpha \right]. \quad (2.32)$$

The squared  $Z^0$  amplitude introduces a contribution linear in the cosine of the polar angle, although heavily suppressed by  $s^2/M_Z^4$ .

The squared amplitude of the interference between the  $\gamma$  and  $Z^0$  mediated interaction is

$$|\mathcal{M}_{\text{int}}|^2 = B \left[ \bar{u}_\mu \gamma^\nu v_\mu \bar{v}_e \gamma_\nu u_e \bar{v}_\mu \gamma^\rho C^- u_\mu \bar{u}_e \gamma_\rho C^- v_e \right] \\ - \frac{B}{M_Z^2} \left[ \bar{u}_\mu \gamma^\nu v_\mu \bar{v}_e \gamma_\nu u_e \bar{v}_\mu \not{q} C^- u_\mu \bar{u}_e \not{q} C^- v_e \right], \quad (2.33)$$

with

$$B = \frac{e^4}{\sin^2(2\theta_W)} \frac{1}{s} \frac{(s - M_Z^2)}{(s - M_Z^2)^2 + M_Z^2 \Gamma_Z^2}. \quad (2.34)$$

Averaging over the initial spins and summing over the final state spins, the squared amplitude of the interference is

$$\overline{|\mathcal{M}_{\text{int}}|^2} = \frac{B}{4} \text{Tr} \left[ (\not{k}_- + m_\mu) \gamma^\nu (\not{k}_+ - m_\mu) \gamma^\rho C^- \right] \text{Tr} \left[ (\not{p} + m_e) \gamma_\rho C^- (\not{p}_+ - m_e) \gamma_\nu \right] \\ = B s^2 \left[ c_V^2 \left( \left( 1 + z_e^2 z_\mu^2 \cos^2 \alpha \right) + \left( 1 - z_e^2 z_\mu^2 \right) + \left( 1 - z_e^2 \right) \left( 1 - z_\mu^2 \right) \right) \right. \\ \left. + 2c_A^2 z_e z_\mu \cos \alpha \right]. \quad (2.35)$$

Here I immediately dropped the second term of eq. (2.33) since it is equal to 0.

The combination of eqs. (2.19), (2.32) and (2.35) as shown in eq. (2.3) is the final result for the total squared amplitude. The two final terms in eqs. (2.32) and (2.35) are linear in  $\cos \alpha$ . Therefore, the polar angle distribution is slightly asymmetric. This relates the polar angle distribution and its asymmetry to the Weinberg angle and allows to measure it. Both terms linear in  $\cos \alpha$  are suppressed since BelleII operates at energies far below the  $Z^0$  mass. Comparing

both linear contributions, the term in eq. (2.35) is amplified compared to the one in eq. (2.32) by  $(s - M_Z^2)/s$ . Therefore, the squared amplitude eq. (2.35) is the leading contribution to the forward-backward asymmetry and determines the polar angle distribution. I compare my expressions to the expression for the total tree-level matrix element in [24], by setting  $z_e = 1$ . The expressions match, confirming my results.

## 2.2 Weinberg angle fit

To cross check the extraction of the Weinberg angle from the polar angle distribution, I use the Markov Chain Monte Carlo sampler emcee [25] to first generate different polar angle distributions and then extract  $\sin^2 \theta_W$  back from them. Emcee uses a probability density function, which is explored, by moving around in the available phase-space. If the sampling process works, then the entries of the chains reflect the true underlying probability density function. Each entry in the chain corresponds to one so-called step. The probability density function I provide is the squared total amplitude normalized to its value integrated over the whole  $\cos \alpha$  range

$$P(\cos \alpha | \sin^2 \theta_W) = \frac{|\mathcal{M}(\cos \alpha, \sin^2 \theta_W)|^2}{\int_{-1}^1 |\mathcal{M}(\cos \alpha, \sin^2 \theta_W)|^2 d \cos \alpha}. \quad (2.36)$$

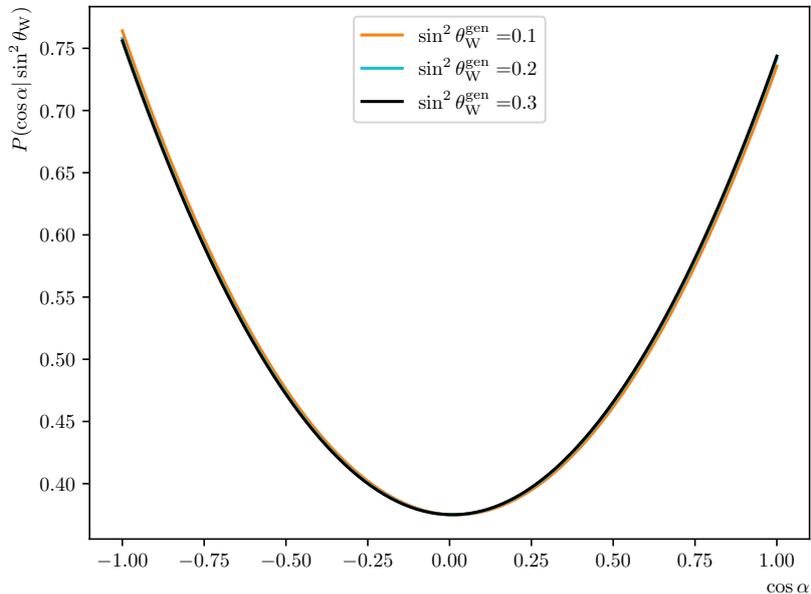
For the generation, the likelihood in eq. (2.36) takes a fixed value for  $\sin^2 \theta_W$  and  $\cos \alpha$  as a free parameter.

For the generation of the polar angle distribution, I initialize 25 chains and discard the first 5000 steps of the sampling process as so-called burn-in steps. This explores the available phase space and should get rid of dependencies on the initialization. The emcee autocorrelation time of the chains, which is given in terms of steps, is used to test if further burn-in steps are needed. This is done by calculating the ratio of the burn-in steps to the autocorrelation time and requiring the ratio to exceed 50. If this is not the case, the procedure is repeated, but with an additional 5000 burn-in steps. I do this each time, before generating polar angle distributions for the values  $\sin^2 \theta_W^{\text{gen}} = 0.1, 0.2$  and  $0.3$ . For the sample generation, the mass and width of the  $Z^0$  take the values  $M_Z = 91.1876 \text{ GeV}$  and  $\Gamma_Z = 2.50 \text{ GeV}$ . The masses of the electron  $m_e = 0.51099895 \text{ MeV}$  and muon  $m_\mu = 0.106 \text{ GeV}$  are also fixed, in agreement with their Particle Data Group values [9]. The center-of-mass energy is fixed at  $\sqrt{s} = 10.58 \text{ GeV}$ .

After the burn-in phase, the sampling process is started, until a total of  $50 \times 10^6$   $\cos \alpha$  values have been generated. These are then binned into 100 uniform, disjoint bins covering the range  $-1$  to  $1$ .

Figure 2.2 shows the polar angle probability distribution  $P(\cos \alpha | \sin^2 \theta_W)$  calculated from eq. (2.36) for three  $\sin^2 \theta_W^{\text{gen}}$  used. The only discernible difference between the distributions is between the distribution sampled from  $\sin^2 \theta_W^{\text{gen}} = 0.1$  and  $\sin^2 \theta_W^{\text{gen}} = 0.3$ . For  $\sin^2 \theta_W^{\text{gen}} = 0.1$ , the distribution is consistently higher for  $\cos \alpha < 0$  than for the others. Similarly, when  $\cos \alpha > 0$ , the situation is reversed. Therefore, the polar angle distribution for  $\sin^2 \theta_W^{\text{gen}} = 0.1$  is more asymmetric than for  $\sin^2 \theta_W^{\text{gen}} = 0.2$  or  $0.3$ .

Figure 2.3 shows the region  $\cos \alpha \in [-0.25, 0.25]$ . Here the differences are better visible. In particular, the asymmetry decreases with increasing  $\sin^2 \theta_W^{\text{gen}}$ .



**Figure 2.2:** The polar angle distribution calculated from the normalized squared total amplitude in eq. (2.36) for three values of  $\sin^2 \theta_W^{\text{gen}}$ .

The difference between the distributions for  $\sin^2 \theta_W^{\text{gen}} = 0.2$  and  $0.3$  is not as pronounced as it is comparing them to  $\sin^2 \theta_W^{\text{gen}} = 0.1$ .

Since we expect a good efficiency and Belle II plans to record about  $50 \text{ ab}^{-1}$ , we opt for a binned maximum likelihood fit to extract the value of  $\sin^2 \theta_W$  from the polar angle distribution. For the binned maximum likelihood fit, I again use emcee. Bayes' Theorem relates the posterior distribution of the Weinberg angle  $P(\sin^2 \theta_W | \text{Data})$  given the polar angle distribution to the posterior of the polar angle distribution  $P(\text{Data} | \sin^2 \theta_W)$

$$P(\sin^2 \theta_W | \text{Data}) \propto P(\text{Data} | \sin^2 \theta_W) P_0(\sin^2 \theta_W), \quad (2.37)$$

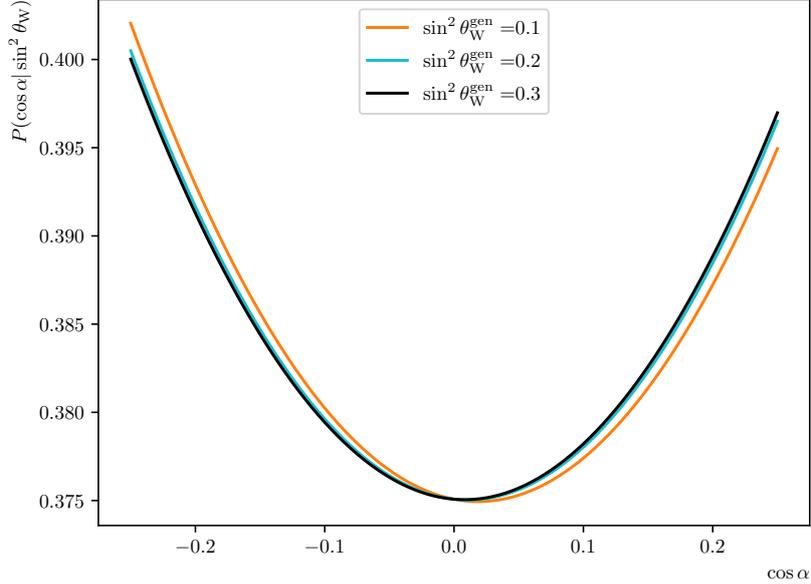
with  $P_0(\sin^2 \theta_W)$  as a uniform prior on the sine squared of the Weinberg angle.

Since the  $\cos \alpha$ -distribution is binned into a histogram, each bin has the likelihood

$\mathcal{L}_b(\text{Data} | \sin^2 \theta_W)$ . The total likelihood  $P(\text{Data} | \sin^2 \theta_W)$  is then the product of all the individual likelihoods. In each bin  $b$ , a Poisson distribution with an expected rate  $\nu_b$  determines the probability to observe  $n_b$  events in that bin

$$P(n_b | \nu_b) = \frac{\nu_b^{n_b}}{n_b!} \exp^{-\nu_b}. \quad (2.38)$$

The yield  $Y$  is the number of events that really occurred during data taking.  $N$  is then the number of events that were observed during data taking. It can be calculated by summing all the bin counts in the polar angle distribution.



**Figure 2.3:** The polar angle distribution calculated from the normalized squared total amplitude in eq. (2.36) for three values of  $\sin^2 \theta_W^{\text{gen}}$  in the region  $[-0.25, 0.25]$ .

The log likelihood of the observed data is then

$$\begin{aligned} \log P(\text{Data} | \sin^2 \theta_W, Y) &= \sum_b \log P(n_b | \nu_b), \\ &= \sum_b n_b \log \nu_b - \nu_b, \end{aligned} \quad (2.39)$$

where  $\nu_b$  depends on the Weinberg angle and the yield. I neglect the term  $\log n_b!$  since it does not depend on the Weinberg angle.

In addition, I define the normalized differential cross section

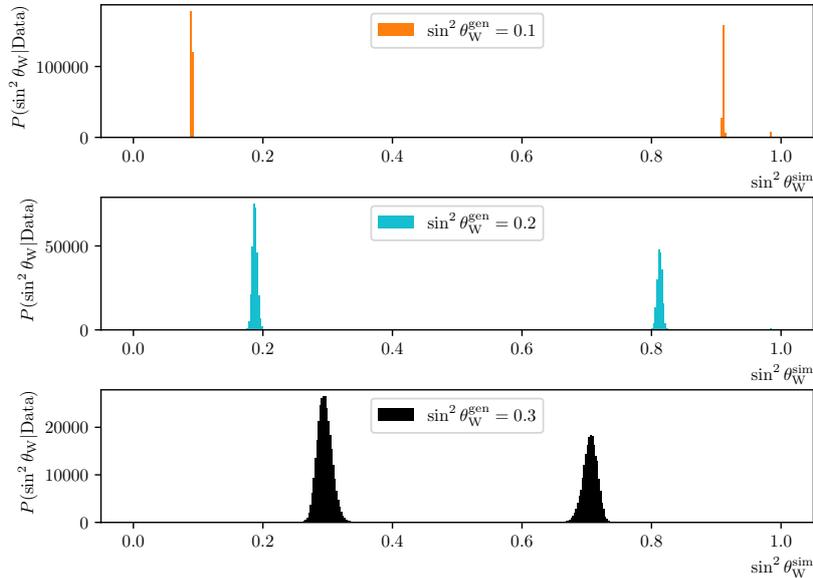
$$f(\sin^2 \theta_W, \cos \alpha) \equiv \frac{|\mathcal{M}(\cos \alpha, \sin^2 \theta_W)|^2}{\int_{-1}^1 |\mathcal{M}(\cos \alpha, \sin^2 \theta_W)|^2 d \cos \alpha}. \quad (2.40)$$

The number of events  $n_b$  in each bin  $b$  is given from the  $\cos \alpha$  histogram. The rate in each bin is

$$\nu_b = Y \int_b f(\sin^2 \theta_W, \cos \alpha) d \cos \alpha, \quad (2.41)$$

where the integral goes from the lower edge of the bin  $b$  to its upper edge. Using this, eq. (2.39) is then

$$\begin{aligned} \log P(\text{Data} | \sin^2 \theta_W, Y) &= N \log Y - Y \\ &+ \sum_b n_b \log \int_b f(\sin^2 \theta_W, \cos \alpha) d \cos \alpha. \end{aligned} \quad (2.42)$$



**Figure 2.4:** The sampled  $\sin^2 \theta_W$  probability density distributions simulated with the polar angle distributions for  $\sin^2 \theta_W^{\text{gen}} = 0.1$  (orange),  $\sin^2 \theta_W^{\text{gen}} = 0.2$  (cyan) and  $\sin^2 \theta_W^{\text{gen}} = 0.3$  (black).

The binned maximum likelihood splits into two different, independent parts. The first two terms depend on the yield but not the polar angle, while the third term contains the integration over  $\cos \alpha$  and no yield dependence. The posterior distribution for the yield is analytically solvable. Since  $N$  is large, the posterior distribution of the yield is approximately a normal distribution with mean  $N$  and standard deviation  $\sqrt{N}$ . Using a uniform prior between 0 and 1, this is the posterior distribution  $P(\sin^2 \theta_W | \text{Data})$ . The yield is required to be non-zero and positive.

To make it a consistency check, I use the same squared total amplitude, masses, widths and energies for toy sample generation and fitting. To sample from the posterior distribution of  $\sin^2 \theta_W$ , I again use emcee with the log likelihood as described in eq. (2.42). Five chains are initialized and the first 10 000 steps are discarded as burn-in steps. The ratio of burn-in steps to the autocorrelation time given in steps has to be above 50, otherwise adding 20 000 burn-in steps and starting the burn-in process again. After burn-in, the sampling process continues until a total set of 500 000 steps has been generated.

Figure 2.4 shows the sampled posterior distribution for  $\sin^2 \theta_W$  for the three values of  $\sin^2 \theta_W^{\text{gen}}$  used. The sampling process finds two values for  $\sin^2 \theta_W$ , one above and one below 0.5. This is due to the terms linear in  $\cos \alpha$  in the matrix element being proportional to  $\sin^{-2}(2\theta_W)$ . This is ambiguous in  $\sin^2 \theta_W$ . The peaks below 0.5 are larger, meaning they have a higher probability density. The reason for this is that the squared total amplitude also has a weak dependence on  $c_V$ .  $c_V$  is unambiguous in  $\sin^2 \theta_W$ . Therefore, the sampling procedure indeed finds the true value used for generating the polar angle distribution.

Table 2.1, shows the sample mean and sample standard deviation of the

**Table 2.1:** The first column shows the three values of  $\sin^2 \theta_W$  used for generating the polar angle distributions. The second and third column contain the mean and standard deviations of the posterior distributions. To calculate the mean and standard deviation of the posterior distribution of  $\sin^2 \theta_W$  only the values below 0.5 are used.

$\sin^2 \theta_W^{\text{gen}}$	$\sin^2 \theta_W^{\text{sim}}$	$Y$
0.1	$0.090 \pm 0.001$	$49\,999\,811 \pm 7058$
0.2	$0.187 \pm 0.007$	$50\,000\,091 \pm 7085$
0.3	$0.295 \pm 0.012$	$50\,000\,015 \pm 6999$

posterior of  $\sin^2 \theta_W$ , as well as for the yield. The sample mean  $\sin^2 \theta_W^{\text{sim}}$  and sample standard deviation of  $\sin^2 \theta_W$  take only the values below 0.5 into account. A different approach is to limit the available range for  $\sin^2 \theta_W$  to below 0.5 since  $\sin^2 \theta_W$  is expected to be around 0.23 [8]. Sampling the yield is not necessary since the likelihood in eq. (2.42) is analytically solvable. However, I used it as a cross check for the sampling process and the sampled results agree nicely with the expectation of a Gaussian distribution with mean  $N$  and standard deviation  $\sqrt{N}$ .

The expected value of  $\sin^2 \theta_W$  is around 0.23, therefore I estimate the relative uncertainty from the fit with  $\sin^2 \theta_W^{\text{gen}} = 0.2$ . The fit result has a relative uncertainty of 3.7%. Taking into account the efficiency calculated in section 3.4 and the angular coverage of Belle II, I assume a reconstruction efficiency of about 40% and a cross section of 1.148 nb. The  $50 \times 10^6$  events in each data set, then correspond to about  $100 \text{ fb}^{-1}$  integrated luminosity. Before its first long shutdown in 2022, Belle II recorded  $424 \text{ fb}^{-1}$  of integrated luminosity [26]. Scaled to this luminosity, the relative uncertainty would be 1.8%. This is comparable to the NuTeV result [20], which has a relative uncertainty of 1.0%.

The  $\sin^2 \theta_W^{\text{sim}}$  values in table 2.1 are consistently below the values used for generation. I calculate the pull using the value used for generating the polar angle distribution  $\sin^2 \theta_W^{\text{gen}}$ , the value extracted from the sampling  $\sin^2 \theta_W^{\text{sim}}$  and the uncertainty extracted from the sampling  $\sigma^{\text{sim}}$

$$\frac{\sin^2 \theta_W^{\text{sim}} - \sin^2 \theta_W^{\text{gen}}}{\sigma^{\text{sim}}}. \quad (2.43)$$

Table 2.2 shows the pulls for the three values of  $\sin^2 \theta_W^{\text{gen}}$ . It confirms the agreement between the  $\sin^2 \theta_W^{\text{gen}}$  and  $\sin^2 \theta_W^{\text{sim}}$  improves for the two higher values. A possible explanation of this behavior is that the total squared amplitude diverges as  $\sin^2 \theta_W$  approaches 0 or 1 since it depends on  $\sin^{-2}(2\theta_W)$ . Estimating the bias and pulls in a next step would require to produce multiple data sets for the same value  $\sin^2 \theta_W^{\text{gen}}$  and studying them.

In conclusion, I have introduced a method to extract the Weinberg angle from a polar angle distribution for  $\mu^+ \mu^-$  events. I also compared the numerical values of the different contributions to the total squared amplitude to code written by F. Krinner independently. The different contributions agreed within the numerical uncertainties.

The next step would introduce detector simulation, but we started without it since the setup is easier and the Belle II software was not public, when I started

**Table 2.2:** The first column shows the three values of  $\sin^2 \theta_W$  used for generating the polar angle distributions. The second column shows the pull of  $\sin^2 \theta_W^{\text{sim}}$  given in table 2.1 from the value  $\sin^2 \theta_W^{\text{gen}}$  used.

$\sin^2 \theta_W^{\text{gen}}$	Pull
0.1	-12.207
0.2	-1.858
0.3	-0.389

working on the sampling procedure. The results presented retain their validity since the sampling procedure is still necessary to extract the Weinberg angle.

## Chapter 3

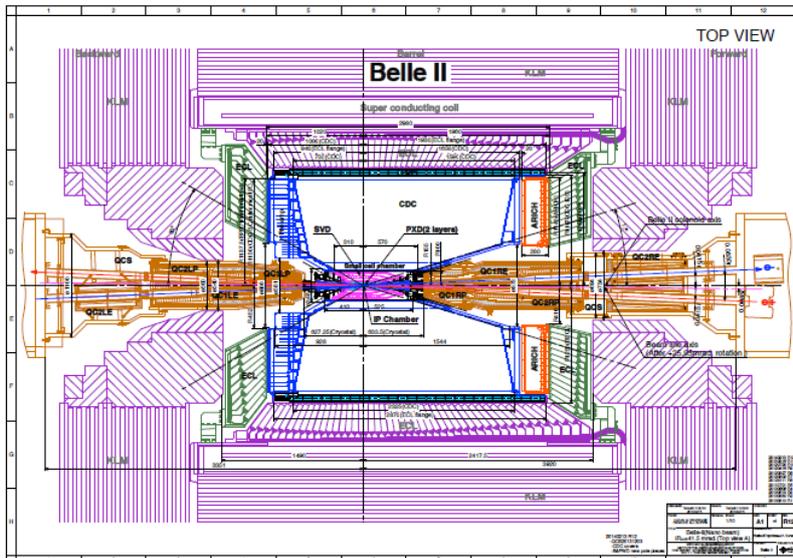
# Study of background sources

To measure the Weinberg angle with muon pairs, the signal events should contain only two muons. No other particles should be in the detector. However, there are other processes, which can look similar to the signal events in the detector. We call these background processes. They can be categorized as particle identification related background processes and detection related background processes.

The particle identification related background processes are processes, where either electrons or hadrons are mis-identified as muons. In the process  $e^+e^- \rightarrow h^+h^-$ , both hadrons  $h^\pm$  can be mis-identified as muons. Other possibilities are that the hadrons decay into electrons or positrons with two corresponding neutrinos and the electron-positron pair is mis-identified as a muon pair. The other possibility is the process  $e^+e^- \rightarrow \tau^+\tau^-$ , where the tau pair decays to a pair of electrons, a pair of hadrons or an electron and a hadron. The decay products of the tau are then mis-identified so that the total event looks like an  $e^+e^- \rightarrow \mu^+\mu^-$  event. The pair of neutrinos is not measured in the detector.

The detection related background processes are processes in which the detector detects one muon pair, but misses at least one additional particle. The general structure of these processes is  $e^+e^- \rightarrow \mu^+\mu^-X$ , where  $X$  is a system of one or more undetected particles. I treat  $X$  as a single particle, keeping in mind that it could also be a system of multiple particles. A possible reason for not detecting  $X$  is that it traveled outside the geometric acceptance of the detector, for example the beam pipe. Another reason could be that the momentum  $\vec{p}_X$  of  $X$  is simply too low so that it does not reach the necessary detector parts to be measured. Most of the time, the detection related background processes are going to be  $e^+e^- \rightarrow \mu^+\mu^-\gamma$  events. Depending on the photon energy, the angular distribution of the muon pair is changed. Therefore, we want to exclude such events above a certain threshold for the photon energy.

The particle identification at Belle II is still under development, so we focus on separating detection related background processes from our signal processes using kinematic variables. Therefore, differences in the distribution of kinematic variables between signal and background processes are studied using the official Belle II simulation software [27, 28].



**Figure 3.1:** The top view of the Belle II detector taken from [29].

### 3.1 SuperKEKB and the Belle II detector

The SuperKEKB accelerator accelerates electrons and positrons, which collide with a center-of-mass energy of 10.58 GeV. In the laboratory frame, the electron beam has an energy of 7 GeV, while the positron beam has an energy of 4 GeV. Therefore, SuperKEKB is an asymmetric collider with most of the particles boosted in the direction of the electron beam. This is the forward direction at Belle II. The targeted instantaneous luminosity of SuperKEKB is  $8 \times 10^{35} \text{ cm}^{-2} \text{ s}^{-1}$  [23]. This would be the world record for the instantaneous luminosity, but has not been achieved by SuperKEKB so far.

Belle II [23] is the new detector at the interaction point of SuperKEKB and the upgrade of the former Belle detector. The goal of Belle II is to accumulate  $50 \text{ ab}^{-1}$  [23] of integrated luminosity for precision studies of particle physics. For  $1 \text{ ab}^{-1}$  of integrated luminosity, SuperKEKB will produce about  $1.148 \times 10^9$  muon pairs. Even for a bad reconstruction efficiency of 0.1%, this still corresponds to  $1.148 \times 10^6$  muon pairs per  $\text{ab}^{-1}$ . The Belle II detector is an almost complete  $4\pi$ -detector. Here I only provide an overview of its most important parts and their usage. A complete description is given in [23, 29, 30].

Figure 3.1 shows the schematics of the detector from the top side. The sub-detectors are arranged in a cylindrical fashion around the interaction region at the center of the image. From the inner to the outer layer, the detector is composed of the silicon pixel detector (PXD), the silicon vertex detector (SVD), the central drift chamber (CDC), the time-of-propagation counter (TOP), the aerogel ring-imaging Cherenkov detector (ARICH), the electromagnetic calorimeter (ECL) and the  $K_L^0$ -and muon detector (KLM). Part of the KLM is the iron yoke, which together with the superconducting solenoid provides the magnetic field of 1.5 T for the Belle II detector.

The silicon pixel detector currently consists of one layer of silicon pixels. The silicon vertex detector consists of four layers of silicon strips. Both register hits from charged particles, which are then used in the track reconstruction. Since they are the closest sub-detectors to the interaction point, they provide the most detailed information about the origin of the track.

The central drift chamber consists of a large volume of He-C<sub>2</sub>H<sub>6</sub> gas. 56 layers of wires, grouped into 9 superlayers, pass through the volume of the central drift chamber. Charged tracks traveling through the gas ionize it and the ionization is then measured at the wires. This is used to reconstruct tracks of charged particles. In addition, the curvature of the tracks provides information about the momentum of the track. This is important for particle identification because the particle identification often relies on relating the momentum of a particle to other measurable quantities of the same particle. Another piece of information relevant to charged particle identification is the specific energy loss  $dE/dx$  of a particle in the central drift chamber. This depends on the momentum of the track and on the particle species responsible for the track. Measurements of the central drift chamber are particularly important for tracks with low momentum, which might not reach the outer sub-detectors.

The time-of-propagation counter and the aerogel ring-imaging Cherenkov detector complement each other. The time-of-propagation counter covers the barrel region and an aerogel ring-imaging Cherenkov detector in the forward endcap. Both use Cherenkov radiation to measure the velocity of charged particles passing through the detector material.

The electromagnetic calorimeter is made out of CsI(Tl) crystals arranged as a barrel with a forward endcap and a backward endcap. Photons and charged particles create clusters of energy deposition in these crystals. The clusters provide information on the energy and location of particles. This is in particular important for the  $\pi^0$  reconstruction since the  $\pi^0$  decays mainly to two photons. Other tasks of the electromagnetic calorimeter include the reconstruction of  $K_L^0$  together with the  $K_L^0$  and muon detector, the luminosity measurement and providing trigger information.

The  $K_L^0$  and muon detector is made out of a stack of alternating iron plates and detector material. Similar to the electromagnetic calorimeter, it consists of a barrel and two endcaps, one in the forward and one in the backwards region. As the name suggests, its main purpose is to detect  $\mu^\pm$  and  $K_L^0$ . These usually travel almost undisturbed through the other detector parts but can be detected in the  $K_L^0$  and muon detector due to the iron plates. Since the  $\mu^\pm$  is a lepton, whereas the  $K_L^0$  is a hadron, they interact differently with the material of the  $K_L^0$  and muon detector. In addition, the muon is a charged particle, so it should have a reconstructed track associated to it in the other sub-detectors. The neutral  $K_L^0$  does not interact with the tracking sub-detectors and should not have an associated track. Therefore, the separation between  $K_L^0$  and  $\mu^\pm$  should work well.

### 3.2 Study of $\mu^+\mu^-X$ kinematics

I focus on rejecting  $e^+e^- \rightarrow \mu^+\mu^-X$  processes. Since there is at least one missing particle in the system  $X$ , the  $\mu^+\mu^-X$  final state is described by three four-momenta, so twelve degrees of freedom. Four-momentum conservation removes four of them, so there are eight degrees of freedom left. Two particles are

identified as muons, so they need to fulfill the four-momentum relation  $p^2 = m_\mu^2$ , which removes two degrees of freedom. There remain six degrees of freedom. Two of these degrees of freedom are the azimuth and the polar angle of the  $\mu^-$ . Due to momentum conservation, the momenta of the  $\mu^+$ ,  $\mu^-$  and  $X$  all lie within one so-called decay plane. The orientation of the decay plane around the  $\mu^-$  direction is described by the angle between the decay plane and the  $e^-$ - $\mu^-$  plane. The azimuth and the orientation of the decay plane are not relevant for measuring the Weinberg angle. The polar angle is measured and therefore is not a suitable variable to discriminate signal from background events. Therefore, there are three remaining degrees of freedom, that are used to differentiate the three-body final state from the signal two-body final state.

Due to the presence of the particle  $X$ , the muon energies deviate from  $\sqrt{s}/2$ , as they would be in the signal  $e^+e^- \rightarrow \mu^+\mu^-$  process. I introduce the energies  $E_\pm$  as kinematic variables of interest. The energy measurements of the muons should not influence each other, therefore they can be combined into the radial energy deviation of the event

$$R_{\mu\mu} = \sqrt{(\langle E \rangle - E_+)^2 + (\langle E \rangle - E_-)^2}. \quad (3.1)$$

The value  $\langle E \rangle$  is not fixed for a measurement on real data. One could use the individual averaged energies of  $\mu^+$  and  $\mu^-$  in true  $e^+e^- \rightarrow \mu^+\mu^-$  events in both terms under the square root in simulated or recorded events. For now, I define  $\langle E \rangle$  as the combined average energy of  $\mu^+$  and  $\mu^-$  in true, simulated signal events  $e^+e^- \rightarrow \mu^+\mu^-$ . For the signal events with perfect reconstruction and no resolution effects  $R_{\mu\mu} = 0$  GeV.

Since the muons in  $\mu^+\mu^-X$  are no longer back-to-back, I introduce the acollinearity  $\cos \eta \equiv -\hat{p}_+\hat{p}_-$ . It describes the deviation from the muons being back-to-back in the center-of-mass frame and is defined using the normalized momenta of the muons. In the center-of-mass frame, the muons in signal events are back-to-back with  $\cos \eta = 1$ .

Finally, I study the missing invariant mass square  $M_X^2 = p_X^2$ , where  $p_X$  is the four-momentum of the missing particle  $X$ . In the two-body final state, there are no missing particles, therefore  $M_X^2 = 0$  GeV<sup>2</sup>.

As we have seen, these variables have precise values for the two-body signal final state. Therefore, they are interesting to study separating signal and background processes. In principle, next to the energies, only one additional variable is needed for a complete description of the  $\mu^+\mu^-X$  topology. However, the finite resolution of the detector means that four-momentum conservation can be broken for one kinematic variable, while it is approximately conserved for another. Therefore, we consider all the variables, that have been introduced so far. By assumption  $X$  is not measured in the detector, so all the quantities related to  $X$  are expressed in terms of measured quantities. The missing mass square is then

$$M_X^2 = s + s' - 2\sqrt{s}(E_+ + E_-), \quad (3.2)$$

where  $s' \equiv (p^+ + p^-)^2$  is the squared mass of the muon pair.

In a presentation in 2014 [31], T. Ferber performed an early study for a possible measurement of the Weinberg angle at Belle and Belle II. He introduced a

different set of variables, assuming the missing particle to be a photon, including

$$x^\pm \equiv \frac{2p^\pm p_X}{s}. \quad (3.3)$$

I treat the three-body final state more generally by letting  $X$  be unknown and undetected. It is therefore interesting, to further study the relation between the different variables that have been introduced. The four-momentum of  $X$  is replaced in the definition of  $x^\pm$ , using  $p_X = p_e^+ + p_e^- - p^+ - p^-$

$$p^\pm p_X = p^\pm (p_e^+ + p_e^- - p^+ - p^-) = E_\pm \sqrt{s} - m_\mu^2 - p^+ p^-. \quad (3.4)$$

The last two terms are replaced using

$$s' \equiv (p^+ + p^-)^2 = 2m_\mu^2 + 2p^+ p^-, \quad (3.5)$$

and so

$$x^\pm \equiv \frac{2p^\pm p_X}{s} = \frac{2E_\pm \sqrt{s} - s'}{s}. \quad (3.6)$$

Therefore, it is not necessary to study the distribution in  $x^\pm$ . The focus is then on studying the relations between  $E_\pm$ ,  $M_X^2$  and  $\cos \eta$ . Using the definition of  $s'$ , the acollinearity can be expressed in terms of the muon energies

$$s' - 2m_\mu^2 = 2p^+ p^- = 2E_+ E_- + 2\sqrt{E_+^2 - m_\mu^2} \sqrt{E_-^2 - m_\mu^2} \cos \eta. \quad (3.7)$$

Therefore, the acollinearity is

$$\cos \eta = \frac{s' - 2m_\mu^2 - 2E_+ E_-}{2\sqrt{E_+^2 - m_\mu^2} \sqrt{E_-^2 - m_\mu^2}}. \quad (3.8)$$

The acollinearity is related to the energies of the muons and the mass of the muon pair  $s'$ . At Belle II, the approximation  $m_\mu = 0$  is a good one

$$\cos \eta \approx \frac{s'}{2E_+ E_-} - 1. \quad (3.9)$$

An earlier study proposed  $\cos \eta > 0.985$  [31]. In terms of a general lower bound  $\cos \eta_0$  on the acollinearity, this means

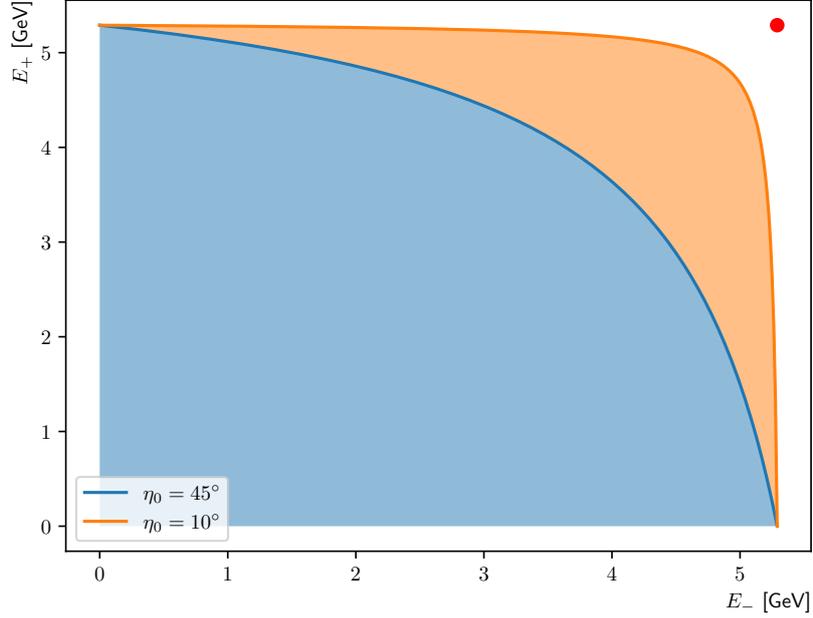
$$\frac{s'}{2E_+ E_-} - 1 > \cos \eta_0. \quad (3.10)$$

As a limit on the product of muon energies

$$E_+ E_- < \frac{s'}{2(1 + \cos \eta_0)}. \quad (3.11)$$

$s'$  depends on the muon energies. Using the definition of  $s$ , all terms depending on the muon energies are grouped together

$$s \equiv |p^+ + p^- + p_X|^2 = (p^+ + p^-)^2 + M_X^2 + 2p_X (p^+ + p^-). \quad (3.12)$$



**Figure 3.2:** The selection bounds for two values of the acollinearity angle  $\eta_0 = 10^\circ$  (orange) and  $\eta_0 = 45^\circ$  (blue) under the assumption  $M_X^2 = 0$  GeV. The red dot is the point (5.29, 5.29) GeV.

Solving this for  $s'$  yields

$$s' = 2\sqrt{s}(E_+ + E_-) + M_X^2 - s. \quad (3.13)$$

Therefore, the bound on the acollinearity can be written in terms of the muon energies and the square of the missing mass. The signal  $e^+e^- \rightarrow \mu^+\mu^-$  events have no missing mass. Therefore, I approximate  $M_X^2 = 0$

$$s' = 2\sqrt{s}(E_+ + E_-) - s. \quad (3.14)$$

Using this, the acollinearity selection bound is

$$E_{\pm} > \frac{\frac{1}{2}s - \sqrt{s}E_{\mp}}{\sqrt{s} - (1 + \cos \eta_0)E_{\mp}}. \quad (3.15)$$

Figure 3.2 shows two acollinearity selection bounds, where the areas below the curve are excluded. The red point at (5.29, 5.29) GeV marks the location of signal muon events, if the energies of the muons are perfectly reconstructed. The reconstructed signal events are distributed around this point due to imprecisions of the detector.

The selection of the signal events is optimized to achieve the greatest statistical precision on the Weinberg angle. It might be the case that some of the selection criteria I studied here introduce systematic uncertainties. They are usually only studied much later in the development of a study so it is left for future work. Should it turn out that some of the selection criteria I studied here

introduce large systematic uncertainties, it will be necessary to study which selection criteria are best used.

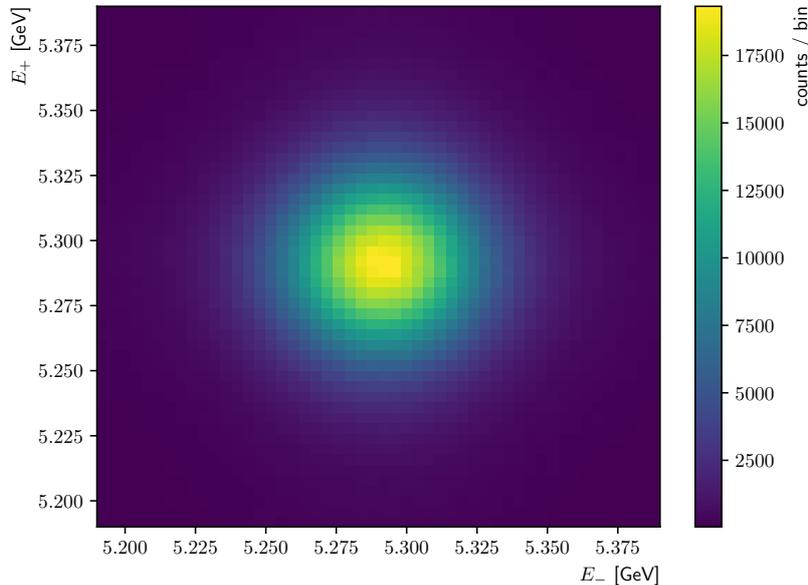
To obtain the highest precision on the Weinberg angle measurement, the distributions for the different variables introduced are studied for signal retention and background rejection. The reconstruction of the events uses some pre-selection criteria, which are aimed at reducing the total output of the simulation, but not distorting the signal distribution. There should be exactly two tracks in each event. These tracks should come from the interaction region since the muons from  $e^+e^- \rightarrow \mu^+\mu^-$  events originate at the interaction point of the beams. This means the distances of closest approach of the two tracks to the interaction point have to be less than 2 cm in the radial direction and less than 5 cm in the direction of the electron beam. If these criteria are not met, then the tracks are most likely not a result of the process of interest to us, but could be coming from cosmic rays or beam background. Since the final state particles are muons, the candidates should have a muon likelihood above 0.1. As the particle identification at Belle II is still being worked on, we are not looking to optimize the selection on the muon identification yet.

An earlier study [31] proposed a timing requirement for the tracks in the time-of-propagation counter. This is aimed at reducing the background from cosmic ray muons. A cosmic muon produces both tracks by traveling through the detector, instead of the two tracks originating at the interaction region. I drop this requirement for now since it is not relevant for this study.

To define the signal events, I use Monte Carlo truth matching. Using generator level information, the reconstructed muons have to be from  $e^+e^- \rightarrow \mu^+\mu^-$  events. A large expected background contribution are from  $\mu^+\mu^-\gamma$  events, since the cross section for the radiation of a photon from one of the beam electrons or one of the muons in the final state diverges as the photon energy approaches 0. Therefore, there are no events having only a muon pair in the final state and no photons. Theoretical corrections to the squared amplitude calculated in section 2.1 depend on a cutoff on the photon energy and will have to take into account the ability of the Belle II detector to measure photons. As the energy of the photon increases, the angular distribution of the muon pair is changed. Events with a high-energy photon are therefore background events, while the situation is more nuanced for low energy photons. For a future Weinberg angle measurement, the cutoff on the photon energy has to be discussed again taking into account the detector performance to measure photons and the theoretical model used to fit the polar angle distribution. For now, generator information is used to define signal events as events where the Monte Carlo generator KKCC [32] produced no photon coming from the beam electrons (initial state radiation) or the muon pair (final state radiation). This means that the KKCC internal cutoff on the photon energy is used. This is the energy at which KKCC deems it necessary to simulate a photon that can interact with the detector.  $\mu^+\mu^-n\gamma$  events with at least one photon are considered a background process.

Figure 3.3 shows the peak region of the reconstructed energies of the signal muons in the center-of-mass frame using a sample of 6 039 150 events. This shows approximately a one standard deviation window around the nominal energy of 5.29 GeV. The peak is tightly focused around the nominal energy.

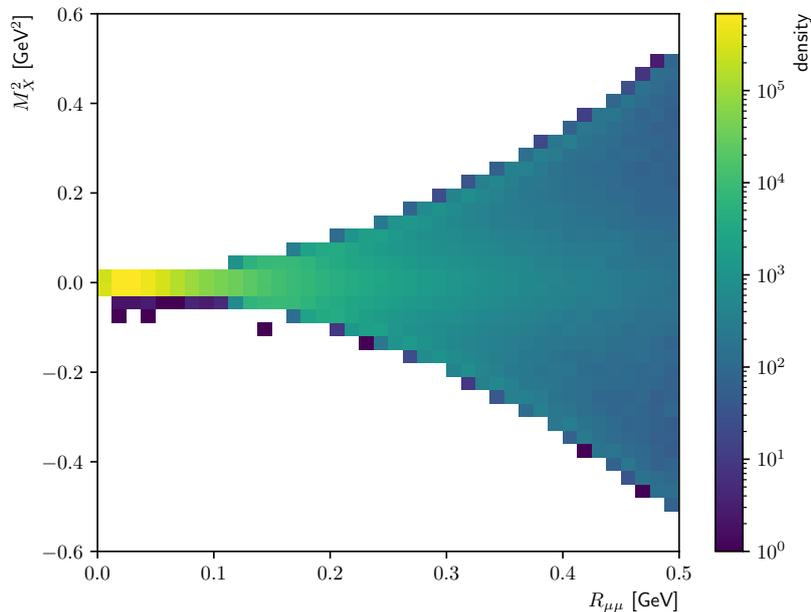
The sample mean and standard deviation for the energies are  $(5.292 \pm 0.105)$  GeV for the  $\mu^+$  and  $(5.293 \pm 0.109)$  GeV for the  $\mu^-$ . The sample means agree well with each other and with the expectation of 5.29 GeV. I use the



**Figure 3.3:** The energies of correctly reconstructed  $\mu^+$  and  $\mu^-$  in  $e^+e^- \rightarrow \mu^+\mu^-$  events.

sample standard deviation as an approximation to the resolution of the energy measurement. Therefore, I expect the selection criterion on the individual energies should be of the order of 100 MeV. The correlation coefficient between the muon energies is  $6 \times 10^{-4}$ , so there is no correlation between both energy measurements. This is consistent with expectation, since the muons of most muon pairs should be in different detector regions. Therefore, their momentum and energy measurements are independent of each other. In fact, the study of  $\mu^+\mu^-\gamma$  events I perform in section 4.4 finds that 99.87% of  $\mu^+$  are isolated in recorded events. The photon in that study needed to have a high-energy and the isolation requirement takes into account additional tracks. Since the signal events do not have any high-energetic photons, this percentage should be even higher. Therefore, the radial energy deviation seems like a good selection criterion to be optimized. The sample mean  $\langle E \rangle$  of the energies of both muons is 5.292 GeV and its sample standard deviation  $\sigma_{\langle E \rangle} = 0.107$  GeV. The sample mean fixes the origin of the circle. The radius of this region of interest is then optimized in a fit to retain signal events and reject background events. From the standard deviations, I expect that the selection criterion on the radial energy deviation will also be around 100 MeV.

For signal events, the two dimensional population distributions of the radial energy deviation and the muon likelihood, and the radial energy deviation and the acollinearity do not show any meaningful correlations. Figure 3.4 shows the two dimensional population distributions of the radial energy deviation  $R_{\mu\mu}$  with the missing mass squared  $M_X^2$  for signal events. The density distribution peaks at  $M_X^2 = 0$  GeV<sup>2</sup> and  $R_{\mu\mu}$  close to 0 GeV. The  $R_{\mu\mu}$  peak is slightly shifted from 0 GeV since the radial energy deviation is the square root of the deviation from the mean. This does not take the sign of the deviation into



**Figure 3.4:** The 2D density distribution of the missing invariant mass squared  $M_X^2$  against the radial energy deviation  $R_{\mu\mu}$  in  $\mu^+\mu^-$  events. It uses a logarithmic coloraxis.

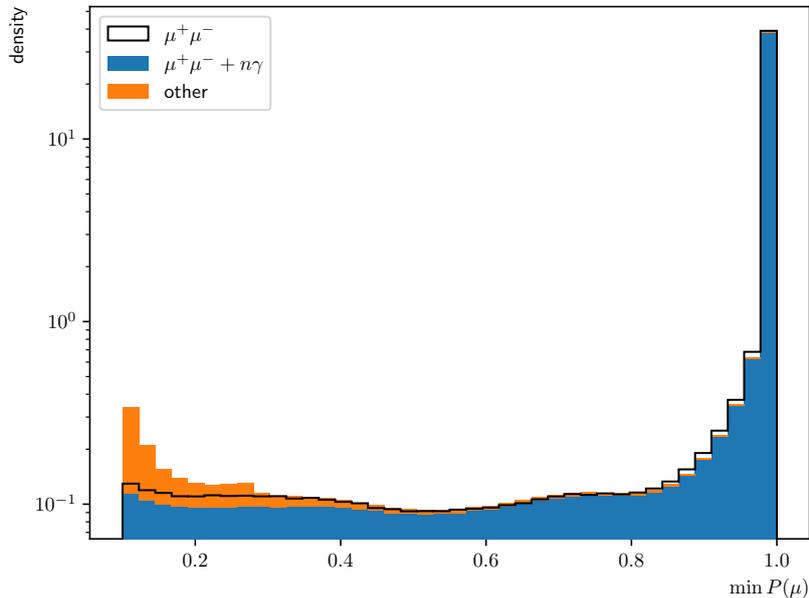
account. A selection on the radial energy deviation selects the events with a vanishing missing mass squared very well, so that no further cut on the missing mass squared should be necessary.

Figures 3.5 to 3.8 show the normalized distributions for the signal process  $\mu^+\mu^-$  (black), as well as the distributions of the  $\mu^+\mu^-\gamma$  (blue) and other background processes (orange). The histograms of both categories of background processes are stacked above each other and normalized together. The contribution of non  $\mu^+\mu^-\gamma$  background processes is very small since the pre-selection criteria already eliminate the majority of such events.

Figure 3.5 shows the minimum of both muon likelihoods  $P(\mu)$  of the two muon candidates. It cannot differentiate between  $\mu^+\mu^-$  and  $\mu^+\mu^-\gamma$ , however setting a higher cutoff would get rid of most of the other background events. In most of these additional events, the reconstruction algorithm confuses an electron or a pion for a muon.

Figures 3.6a and 3.6b show the radial energy deviation distribution for a larger range zoomed-in on the signal peak region respectively. Figure 3.6a shows that the distributions for  $\mu^+\mu^-$  and  $\mu^+\mu^-\gamma$  are very similar at large deviations, however below approximately 3 GeV, the  $\mu^+\mu^-\gamma$  distribution forms a plateau, which is not present in the  $\mu^+\mu^-$  distribution. This is an avenue to reject  $\mu^+\mu^-\gamma$  events.

Figure 3.6b shows the region around the signal peak. While the signal and  $\mu^+\mu^-\gamma$  distributions peak at the same value, the signal distribution drops quicker. The red line in fig. 3.6b marks the optimal value for a selection bound on the radial energy deviation to separate signal from background events as



**Figure 3.5:** The minimal muon likelihood  $\min P(\mu)$  in background and signal (black line) events.

calculated in eq. (3.33) in section 3.4. The second axis on top of the histogram shows the radial energy deviation in units of the sample standard deviation of the energies of both muons.

Figures 3.7a and 3.7b show the acollinearity distribution over the full range and zoomed-in on the signal peak respectively. The difference between  $\mu^+\mu^-$  and  $\mu^+\mu^-\gamma$  is even more pronounced, with the  $\mu^+\mu^-$  distribution forming a very narrow peak at 1, while the background processes basically cover the entire available range.

Finally, fig. 3.8 shows the missing invariant mass squared  $M_X^2$ . The signal and background distributions peak at  $0 \text{ GeV}^2$ . However, the distribution of the combined backgrounds (including  $\mu^+\mu^-\gamma$ ) is broader than the signal distribution. The events away from the peak contain most likely by events with at least one missing final-state particle. Examples for this can be four lepton final states, where two leptons were not seen in the detector or  $\tau$  decays involving at least one neutrino.

Equation (3.15) relates the acollinearity to the muon energies and fig. 3.4 shows a clear dependence of the missing invariant mass square on the radial energy deviation. Therefore, the radial energy deviation is the only selection, that is optimized to reduce the statistical uncertainty and investigate its influence on signal retention and background rejection.

### 3.3 Calculating the selection efficiency

To optimize the radial energy deviation, we have to define the selection efficiency. The selection efficiency  $\epsilon$  is the probability to measure an individual event. As a statistical process for  $N$  events, we want to determine how many

events are measured. The probability to measure  $k$  events out of  $N$  is described by the binomial distribution given  $\epsilon$

$$P(k|N, \epsilon) = \binom{N}{k} \epsilon^k (1 - \epsilon)^{N-k}. \quad (3.16)$$

Bayes' Theorem then relates the probability  $P(k|\epsilon, N)$  to the probability of the efficiency

$$P(\epsilon|k, N) \propto P(k|N, \epsilon) P_0(\epsilon), \quad (3.17)$$

where  $P_0(\epsilon)$  is the prior knowledge of  $\epsilon$ .  $P_0(\epsilon)$  can be chosen freely on its support  $[0, 1]$ . As priors I consider a flat prior

$$P_0(\epsilon) = 1 \quad (3.18)$$

and a double inverse prior

$$P_0(\epsilon) = \frac{1}{\epsilon} \frac{1}{1 - \epsilon}. \quad (3.19)$$

The advantage of the double inverse prior is that we do not provide it any information on the scale of the efficiency. The flat prior implicitly assumes the efficiency is of the order 0.1 or larger. The drawback of the double inverse prior is that it is improper, meaning the integral over its support  $[0, 1]$  does not converge. This drawback is not an issue, as long as the posterior is properly normalized. The priors are combined into the more general form

$$P_0(\epsilon) = \epsilon^{\delta-1} (1 - \epsilon)^{\delta-1}, \quad (3.20)$$

where  $\delta = 0$  corresponds to the double inverse prior, while  $\delta = 1$  corresponds to the flat prior. Using the general prior, the posterior distribution of the efficiency is

$$P(\epsilon|k, N) \propto \binom{N}{k} \epsilon^{k+\delta-1} (1 - \epsilon)^{N-k+\delta-1}. \quad (3.21)$$

This has the form of the beta distribution [33, 34] with support  $[0, 1]$

$$P(\epsilon|k, N) = \text{Beta}(\epsilon|k + \delta, N - k + \delta). \quad (3.22)$$

The expectation value of the efficiency is

$$\begin{aligned} \text{E}[\epsilon] &= \int_0^1 \epsilon \text{Beta}(\epsilon|k + \delta, N - k + \delta) d\epsilon \\ &= \frac{k + \delta}{N + 2\delta}. \end{aligned} \quad (3.23)$$

Similarly, the variance of the efficiency is

$$\text{Var}[\epsilon] = \text{E}[\epsilon]^2 \frac{N - k + \delta}{(k + \delta)(N + 2\delta + 1)}. \quad (3.24)$$

**Table 3.1:** Comparison of the expectation value and standard deviation (separated from expectation value by a semi-colon) of the efficiency for the two different choices of priors (flat and double inverse).

choice of prior	$k = 0$	$k = N$	precision
flat ( $\delta = 1, N \gg 1$ )	$(\frac{1}{N}); (\frac{1}{N})$	$(1 - \frac{1}{N}); (\frac{1}{N})$	(approximated)
double inverse ( $\delta = 0$ )	$(0); (0)$	$(1); (0)$	(exact)

For the flat prior ( $\delta = 1$ ), expectation value and variance are

$$\mathbb{E}[\epsilon] = \frac{k+1}{N+2} \text{ and } \text{Var}[\epsilon] = \mathbb{E}[\epsilon]^2 \frac{N-k+1}{(k+1)(N+3)}. \quad (3.25)$$

For the double inverse prior ( $\delta = 0$ ) they are

$$\mathbb{E}[\epsilon] = \frac{k}{N} \text{ and } \text{Var}[\epsilon] = \mathbb{E}[\epsilon]^2 \frac{N-k}{k(N+1)}. \quad (3.26)$$

Table 3.1 compares both choices of prior, when  $k = 0$  or  $k = N$ , where the efficiency approaches 0 or 1. The results for the flat prior are approximated for large  $N$ . The results for the double inverse prior are exact. In particular the variance of the double inverse prior is exactly zero independent of  $N$ . This does not seem like a suitable definition for the efficiency. If we want to determine the selection efficiency of a simulated process, no events passing would mean the selection efficiency is known to be exactly 0, independent of the number of simulated events. The issue of the efficiency at  $k = 0$  and  $k = N$  is that the parameters of the beta distribution are no longer on their support. The efficiency with a flat prior asymptotically approaches 0 or 1 (if  $k = 0$  or  $N$ ) and always has a finite variance, as we would expect. For this reason, I use the results for the efficiency using the flat prior in eq. (3.25).

Information about the distribution of the efficiency is given by the expectation value and variance of the efficiency. This is often interpreted as a Gaussian distribution. However, this interpretation is only valid, if the efficiency is sufficiently far away from the boundaries of its support.

Figure 3.9 shows the probability density functions for  $N = 100$  and three different values of  $k$ . The probability densities with  $k = 0$  or  $k = 100$  are asymmetric and cannot be described by a normal distribution. For  $k = 50$  however, the probability density function is well described by a Gaussian distribution.

### 3.4 Optimizing the radial energy deviation selection

I use the official Belle II Monte Carlo campaign MC14ri\_a and MCri.d, where "ri" stands for run-independent. Run-independent Monte Carlo does not take into account small differences, which occur during different runs of measurements. As an example, the beam energies are fixed to the same value for each event. I include all the processes, which are generically generated at Belle II and list them in table 3.2. The names are determined by the first set of particles generated by the Monte Carlo generators after the virtual particle created

in the electron-positron collision. Depending on which particles are generated, they are then further decayed into different final state particles.

For the reconstruction and detector simulation, the release release-06-00-03 of the basf2 software [27, 28] is used. The same set of pre-selection criteria and the same signal definition as in section 3.2 are applied. This means, the muons are truth-matched and generator-level information is used to assure that no photons from the beam electrons or muons have been generated by KKMC.

With this setup, the radial energy deviation selection is optimized. The goal of the optimization process is to achieve the smallest statistical uncertainty on the measurement of the Weinberg angle. The statistical uncertainty is proportional to the relative statistical variance, which is approximated by the estimated relative statistical variance  $\Phi$ . The estimated relative statistical variance is the inverse square of the so-called figure of merit

$$\Phi = \frac{S + B}{S^2}. \quad (3.27)$$

$S$  is the number of signal events and  $B$  the number of background events.

The expectation value of the estimated relative statistical variance is then minimized given certain efficiencies of the selection criterion on the different processes

$$\mathbb{E}[\Phi] = \mathbb{E}\left[\frac{S+B}{S^2}\right] = \mathbb{E}\left[\frac{1}{S}\right] + \mathbb{E}\left[\frac{B}{S^2}\right]. \quad (3.28)$$

For a certain process  $i$ , the expected number of events is

$$\mathbb{E}[N_i] = \mathbb{E}[L\sigma_i\epsilon_i], \quad (3.29)$$

using its cross section  $\sigma_i$  and efficiency  $\epsilon_i$  and a given luminosity  $L$ . I assume these to be independent distributions. Therefore, the expectation value is the product of the individual expectation values

$$\mathbb{E}[N_i] = \mathbb{E}[L]\mathbb{E}[\sigma_i]\mathbb{E}[\epsilon_i]. \quad (3.30)$$

The estimated relative statistical variance is

$$\mathbb{E}[\Phi] = \mathbb{E}\left[\frac{1}{L}\right]\mathbb{E}\left[\frac{1}{\sigma_s}\right]\mathbb{E}\left[\frac{1}{\epsilon_s}\right] + \sum_b \mathbb{E}\left[\frac{1}{L}\right]\mathbb{E}\left[\frac{\sigma_b}{\sigma_s^2}\right]\mathbb{E}\left[\frac{\epsilon_b}{\epsilon_s^2}\right]. \quad (3.31)$$

The subscript  $s$  indicates values for the signal process and the sum over  $b$  takes into account all the background processes. The cross sections and efficiencies for the background processes are independent of those of the signal process. Using the algebra of random variables, the expectation value of a ratio factorizes into a product of expectation values. Since the expectation value of the luminosity can be pulled out of the sum on the right, it is sufficient to minimize

$$\mathbb{E}[\Phi L] = \mathbb{E}\left[\frac{1}{\epsilon_s}\right]\mathbb{E}\left[\frac{1}{\sigma_s}\right] + \sum_b \mathbb{E}[\epsilon_b]\mathbb{E}\left[\frac{1}{\epsilon_s^2}\right]\mathbb{E}[\sigma_b]\mathbb{E}\left[\frac{1}{\sigma_s^2}\right]. \quad (3.32)$$

In principle, the cross sections are distributed according to Gaussian distributions with a certain mean and standard deviation. However, for the cross

**Table 3.2:** Comparison of the signal and background processes and how well they pass the selection criteria. The last column gives the efficiencies on the  $R_{\mu\mu}$  selection.

process	events generated	events passing	$\epsilon$ [%]
$\mu^+\mu^-$	7 651 236	5 793 621	$75.72 \pm 0.02$
$\mu^+\mu^- + n\gamma$	31 548 764	10 746 740	$34.06 \pm 0.01$
$e^+e^-\mu^+\mu^-$	570 625 835	38 887	$(6.81 \pm 0.03) \times 10^{-3}$
$e^+e^-$	740 699 373	3209	$(4.33 \pm 0.08) \times 10^{-4}$
$\mu^+\mu^-\mu^+\mu^-$	519 833	58	$(1.13 \pm 0.15) \times 10^{-2}$
$u\bar{u}$	158 626 781	5	$<1.45 \times 10^{-6}$
$\tau^+\tau^-$	91 900 000	1	$<2.51 \times 10^{-6}$
$d\bar{d}$	48 237 077	0	$<4.77 \times 10^{-6}$
$s\bar{s}$	50 492 936	0	$<4.56 \times 10^{-6}$
$c\bar{c}$	129 999 202	0	$<1.77 \times 10^{-6}$
$B^0\bar{B}^0$	51 000 000	0	$<4.51 \times 10^{-6}$
$B^+B^-$	54 000 000	0	$<4.26 \times 10^{-6}$
$\gamma\gamma$	354 845 927	0	$<6.49 \times 10^{-7}$
$e^+e^-e^+e^-$	582 792 993	0	$<3.95 \times 10^{-7}$
$e^+e^-\tau^+\tau^-$	18 103 215	0	$<1.27 \times 10^{-5}$
$\mu^+\mu^-\tau^+\tau^-$	287 270	0	$<8.02 \times 10^{-4}$
$\tau^+\tau^-\tau^+\tau^-$	2114	0	$<0.11$
$e^+e^-\pi^+\pi^-$	189 500 000	0	$<1.22 \times 10^{-6}$
$e^+e^-K^+K^-$	79 800 000	0	$<2.89 \times 10^{-6}$
$e^+e^-p\bar{p}$	11 700 000	0	$<1.97 \times 10^{-5}$
$\pi^+\pi^-ISR$	166 700 000	0	$<1.38 \times 10^{-6}$
$\pi^+\pi^-\pi^0ISR$	23 780 000	0	$<9.68 \times 10^{-6}$
$K^+K^-ISR$	16 300 000	0	$<1.41 \times 10^{-5}$
$K^0\bar{K}^0ISR$	8 864 000	0	$<2.60 \times 10^{-5}$

sections, I was unable to find values with uncertainties at the Belle II center-of-mass energy. Therefore, I take them as fixed values from the official Monte Carlo 14 campaign [35].

I minimize the expectation value of  $\Phi L$  using the Powell solver from `scipy` [36]. Figure 3.10 shows the region around the minimum. While the region between 0.22 and 0.23 flattens, it is a minimum of the fit function. The optimal value for the radial energy deviation selection is found for

$$R_{\mu\mu} < 229 \text{ MeV}. \quad (3.33)$$

Table 3.2 lists the number of simulated events  $N$ , the number of events  $k$  passing the  $R_{\mu\mu}$  selection and the expectation value and standard deviation of the efficiency.

For highly suppressed processes, where the efficiency is close to zero, I calculate the 90% interval which covers the smallest range from the boundary using the cumulative distribution and thus define the upper and lower limits. The cumulative distribution of  $\text{Beta}(\epsilon|k+1, N-k+1)$  is the regularized incomplete

beta function  $I_\epsilon(k+1, N-k+1)$  [34]. It has the special properties

$$I_\epsilon(1, N+1) = 1 - (1-\epsilon)^{N+1}, \quad (3.34)$$

$$I_\epsilon(N+1, 1) = \epsilon^{N+1}. \quad (3.35)$$

Equation (3.34) corresponds to the case where  $k=0$ , while eq. (3.35) corresponds to the case  $k=N$ . To calculate the limits on the efficiency, one must solve the equality  $I_\epsilon(1, k+1) = 0.9$ . I also use eq. (3.34) as an approximation, if almost no events pass the selection.

### 3.5 Data composition

The  $\tau^+\tau^-\tau^+\tau^-$  and  $\mu^+\mu^-\tau^+\tau^-$  processes have high upper limits due to the low amount of simulated events that had been produced during the Monte Carlo campaigns. To get a better understanding of the efficiencies, more events would need to be simulated. To estimate, if these processes contribute with the efficiencies calculated from the available data, we need to calculate the number of events at a given luminosity according to eq. (3.30). In order to be independent of the luminosity, I calculate the ratio  $R$  of expected events after the  $R_{\mu\mu}$  selection for each background process compared to the signal process using eq. (3.30)

$$\mathbb{E}[R] = \mathbb{E}\left[\frac{N_b}{N_s}\right] = \mathbb{E}\left[\frac{L\sigma_b\epsilon_b}{L\sigma_s\epsilon_s}\right] = \mathbb{E}\left[\frac{\sigma_b}{\sigma_s}\right]\mathbb{E}\left[\frac{\epsilon_b}{\epsilon_s}\right]. \quad (3.36)$$

I do not do this for the  $\mu^+\mu^- + n\gamma$  process since there the distinction between background and signal events is not obvious. I also calculate the variance to estimate the uncertainty. Since the cross sections are fixed values, I need to calculate the expectation value and variance of the ratio of efficiencies.

#### 3.5.1 Ratio of two efficiencies

I define the ratio  $\rho$  of two efficiencies  $\epsilon_i$  and  $\epsilon_j$  as

$$\rho \equiv \frac{\epsilon_i}{\epsilon_j}. \quad (3.37)$$

The probability distribution of each efficiency is

$$P(\epsilon) = \text{Beta}(\epsilon|k+\delta, N-k+\delta), \quad (3.38)$$

where I used the general prior from eq. (3.20). Using the ratio distribution of two random variables, the probability distribution of  $\rho$  is

$$\begin{aligned} P(\rho) &= \int_0^1 \epsilon_j \text{Beta}(\rho\epsilon_j|k_i+\delta, N_i-k_i+\delta) \text{Beta}(\epsilon_j|k_j+\delta, N_j-k_j+\delta) d\epsilon_j \\ &= \text{B}^{-1}(k_i+\delta, N_i-k_i+\delta) \text{B}^{-1}(k_j+\delta, N_j-k_j+\delta) \rho^{k_i+\delta-1} \\ &\quad \times \int_0^1 \epsilon^{k_i+k_j+2\delta-1} (1-\rho\epsilon_j)^{N_i-k_i+\delta-1} (1-\epsilon_j)^{N_j-k_j+\delta-1} d\epsilon_j. \end{aligned} \quad (3.39)$$

This has the form of the hypergeometric function  ${}_2F_1$  [37, Eq. 15.6.1]

$${}_2F_1(a, b; c; z) = B^{-1}(b, c - b) \int_0^1 t^{b-1} (1-t)^{c-b-1} (1-zt)^{-a} dt, \quad (3.40)$$

with parameters:

$$a = k_i - N_i + 1 - \delta, \quad (3.41)$$

$$b = k_i + k_j + 2\delta, \quad (3.42)$$

$$c = N_j + k_i + 3\delta, \quad (3.43)$$

$$z = \rho. \quad (3.44)$$

Therefore

$$\begin{aligned} P(\rho) = & B^{-1}(k_i + \delta, N_i - k_i + \delta) B^{-1}(k_j + \delta, N_j - k_j + \delta) \rho^{k_i + \delta - 1} \\ & \times B(k_i + k_j + 2\delta, N_j - k_j + \delta) \\ & \times {}_2F_1(k_i - N_i - 1 + \delta, k_i + k_j + 2\delta; N_j + k_i + 3\delta, \rho). \end{aligned} \quad (3.45)$$

The expectation value and variance involve integrating eq. (3.45). The results of these integrals are not intuitively helpful. It is therefore better to calculate the expectation value and variance directly from the ratio of beta distributions using the algebra of random variables. Since the efficiencies are independent, the expectation value of the ratio is

$$E[\rho] = E[\epsilon_i] E[1/\epsilon_j] \quad (3.46)$$

and the variance of the ratio is

$$\text{Var}[\rho] = \text{Var}[\epsilon_i] \text{Var}[1/\epsilon_j] + (E[\epsilon_i])^2 \text{Var}[1/\epsilon_j] + \text{Var}[\epsilon_i] (E[1/\epsilon_j])^2. \quad (3.47)$$

The expectation value  $E[\epsilon_i]$  and its variance  $\text{Var}[\epsilon_i]$  are given in eqs. (3.23) and (3.24). The expectation value of  $1/\epsilon_j$  is

$$\begin{aligned} E[1/\epsilon_j] &= \frac{\Gamma(N_j + 2\delta)}{\Gamma(k_j + \delta) \Gamma(N_j - k_j + \delta)} \int_0^1 \epsilon^{k_j + \delta - 2} (1 - \epsilon)^{N_j - k_j + \delta - 1} d\epsilon \\ &= \frac{N_j + 2\delta - 1}{k_j + \delta - 1}. \end{aligned} \quad (3.48)$$

For the variance of  $1/\epsilon_j$ , I also need

$$\begin{aligned} E[1/\epsilon_j^2] &= \frac{\Gamma(N_j + 2\delta)}{\Gamma(k_j + \delta) \Gamma(N_j - k_j + \delta)} \int_0^1 \epsilon^{k_j + \delta - 3} (1 - \epsilon)^{N_j - k_j + \delta - 1} d\epsilon \\ &= \frac{(N_j + 2\delta - 1)(N_j + 2\delta - 2)}{(k_j + \delta - 1)(k_j + \delta - 2)}. \end{aligned} \quad (3.49)$$

The variance of  $1/\epsilon_j$  is then

$$\begin{aligned} \text{Var}[1/\epsilon_j] &= E[1/\epsilon_j^2] - (E[1/\epsilon_j])^2 \\ &= (E[1/\epsilon_j])^2 \frac{N_j - k_j + \delta}{(N_j + 2\delta - 1)(k_j + \delta - 2)}. \end{aligned} \quad (3.50)$$

For the flat prior ( $\delta = 1$ ), the expectation value and variance of  $1/\epsilon_j$  are

$$\text{E}[1/\epsilon_j] = \frac{N_j + 1}{k_j} \quad (3.51)$$

$$\text{and Var}[1/\epsilon_j] = (\text{E}[1/\epsilon_j])^2 \frac{N_j - k_j + 1}{(k_j - 1)(N_j + 1)}. \quad (3.52)$$

Together with eq. (3.25), the expectation value of the ratio of two efficiencies is

$$\text{E}[\rho] = \frac{k_i + 1}{N_i + 2} \frac{N_j + 1}{k_j}. \quad (3.53)$$

The variance of the ratio of two efficiencies is

$$\text{Var}[\rho] = (\text{E}[\rho])^2 [\lambda_i \lambda_j + \lambda_i + \lambda_j], \quad (3.54)$$

where  $\lambda_i$  and  $\lambda_j$  are the relative variances

$$\lambda_i = \frac{N_i - k_i + 1}{(k_i + 1)(N_i + 3)} \quad (3.55)$$

$$\text{and } \lambda_j = \frac{N_j - k_j + 1}{(k_j - 1)(N_j + 1)}. \quad (3.56)$$

For  $k_i$ ,  $k_j$ ,  $N_i$  and  $N_j$  all much larger than one and neglecting  $\lambda_i \lambda_j$  since it is small, the relative variance of the ratio is approximately

$$\frac{\text{Var}[\rho]}{(\text{E}[\rho])^2} \approx \frac{1}{k_i} - \frac{1}{N_i} + \frac{1}{k_j} - \frac{1}{N_j}. \quad (3.57)$$

I compare this to a simplified estimation of the relative variance assuming  $k_i$ ,  $k_j$ ,  $N_i$  and  $N_j$  to be independent of each other and each one following a Poisson distribution. With the same assumptions of  $k_i$ ,  $k_j$ ,  $N_i$  and  $N_j$  much larger than one and doing Gaussian error propagation, the relative variance is approximately

$$\frac{\text{Var}[\rho]}{(\text{E}[\rho])^2} \approx \frac{1}{k_i} + \frac{1}{N_i} + \frac{1}{k_j} + \frac{1}{N_j}. \quad (3.58)$$

Comparing eq. (3.57) and eq. (3.58), the signs of two terms flipped. Therefore, the simplified approach used for eq. (3.58) overestimates the uncertainties. From the start this simplified approach had the issue that it assumes the number of the events passing the selection criterion to be independent of the number of total events. One should therefore be careful, when assuming distributions to be approximately Gaussian or doing Gaussian error propagation.

### 3.5.2 Expected purity with $R_{\mu\mu}$ selection

I calculate the ratio of expected background to signal events using eqs. (3.53) and (3.54). If the efficiency is only given as an upper limit, the value of the upper limit is used as calculated in eq. (3.34). Table 3.3 shows the cross sections used in the BelleII MC14 campaign [35] for the different processes in the second column. The efficiencies  $\epsilon$  are given in the third column and the expected ratio of background to signal ( $\mu^+ \mu^-$ ) events in the fourth.

**Table 3.3:** The cross sections [35], selection efficiencies and ratio  $R$  of expected number of events for background processes.  $<$  indicates that the 90% upper limit on the efficiency was calculated.

process	cross section [nb]	$\epsilon$ [%]	$R$
$\mu^+\mu^-$	1.148	$75.72 \pm 0.02$	1.00
$e^+e^-\mu^+\mu^-$	18.83	$(6.81 \pm 0.03) \times 10^{-3}$	$(1.5 \pm 0.1) \times 10^{-3}$
$e^+e^-$	295.8	$(4.33 \pm 0.08) \times 10^{-4}$	$(1.47 \pm 0.03) \times 10^{-3}$
$\mu^+\mu^-\mu^+\mu^-$	$0.3512 \times 10^{-3}$	$(1.13 \pm 0.15) \times 10^{-2}$	$(4.59 \pm 0.60) \times 10^{-8}$
$u\bar{u}$	1.605	$<1.45 \times 10^{-6}$	$<2.68 \times 10^{-8}$
$\tau^+\tau^-$	0.919	$<2.51 \times 10^{-6}$	$<2.65 \times 10^{-8}$
$d\bar{d}$	0.401	$<4.77 \times 10^{-6}$	$<2.20 \times 10^{-8}$
$s\bar{s}$	0.383	$<4.56 \times 10^{-6}$	$<2.01 \times 10^{-8}$
$c\bar{c}$	1.329	$<1.77 \times 10^{-6}$	$<2.71 \times 10^{-8}$
$B^0\bar{B}^0$	0.510	$<4.51 \times 10^{-6}$	$<2.65 \times 10^{-8}$
$B^+B^-$	0.540	$<4.26 \times 10^{-6}$	$<2.65 \times 10^{-8}$
$\gamma\gamma$	5.10	$<6.49 \times 10^{-7}$	$<3.81 \times 10^{-8}$
$e^+e^-e^+e^-$	39.55	$<3.95 \times 10^{-7}$	$<1.80 \times 10^{-7}$
$e^+e^-\tau^+\tau^-$	0.01836	$<1.27 \times 10^{-5}$	$<2.69 \times 10^{-9}$
$\mu^+\mu^-\tau^+\tau^-$	$0.1441 \times 10^{-3}$	$<8.02 \times 10^{-4}$	$<1.33 \times 10^{-9}$
$\tau^+\tau^-\tau^+\tau^-$	$0.2114 \times 10^{-6}$	$<0.11$	$<2.65 \times 10^{-10}$
$e^+e^-\pi^+\pi^-$	1.895	$<1.22 \times 10^{-6}$	$<2.65 \times 10^{-8}$
$e^+e^-K^+K^-$	0.0798	$<2.89 \times 10^{-6}$	$<2.65 \times 10^{-9}$
$e^+e^-p\bar{p}$	0.0117	$<1.97 \times 10^{-5}$	$<2.65 \times 10^{-9}$
$\pi^+\pi^-ISR$	0.1667	$<1.38 \times 10^{-6}$	$<2.65 \times 10^{-9}$
$\pi^+\pi^-\pi^0ISR$	0.02378	$<9.68 \times 10^{-6}$	$<2.65 \times 10^{-9}$
$K^+K^-ISR$	$16.30 \times 10^{-3}$	$<1.41 \times 10^{-5}$	$<2.65 \times 10^{-9}$
$K^0\bar{K}^0ISR$	$8.864 \times 10^{-3}$	$<2.60 \times 10^{-5}$	$<2.65 \times 10^{-9}$

At the beginning of this section, the contribution of the  $\tau^+\tau^-\tau^+\tau^-$  and  $\mu^+\mu^-\mu^+\mu^-$  processes to the total data composition was to be studied. Their relative contributions to the total number of events are small even with a high upper limit on the efficiencies since their cross sections are small. Overall, the ratio of expected events is of the order  $1 \times 10^{-8}$  or smaller. Notable exceptions are the  $e^+e^- \rightarrow e^+e^-$  and  $e^+e^- \rightarrow e^+e^-\mu^+\mu^-$  processes.

Figure 3.11 shows the minimal muon likelihood  $P(\mu)$  of both muons, after applying the  $R_{\mu\mu} < 0.229$  GeV selection for signal (black) and  $e^+e^-$  events (orange). The  $e^+e^-$  events peak at 0.1 due to the pre-selection criterion, whereas the signal events peak at 1. The  $e^+e^-$  events are therefore easily removed by setting a tighter selection on the muon likelihood of both candidates. At BelleII, studies involving muons at BelleII often require at least one of the muon candidates to have a muon likelihood above 0.9. Figure 3.11 shows that such a selection would eliminate all the remaining  $e^+e^-$  events.

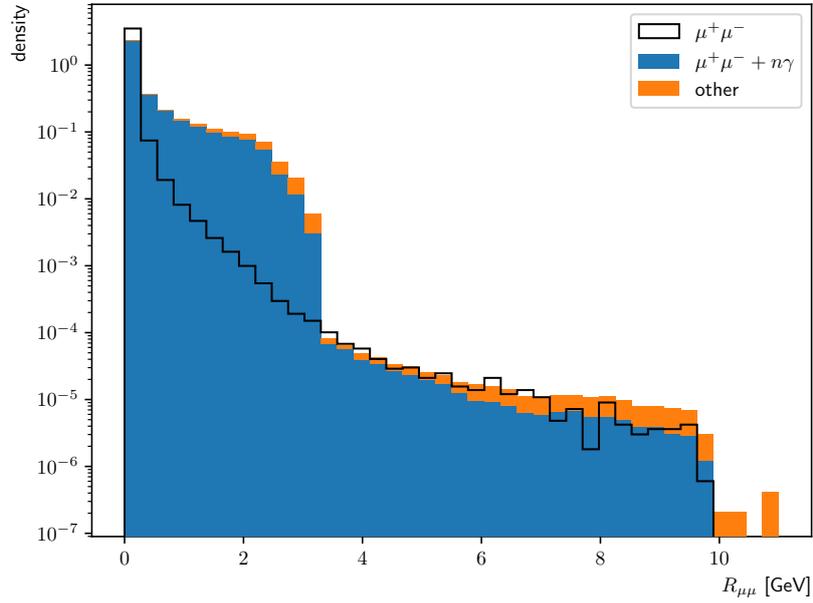
The  $e^+e^-\mu^+\mu^-$  events are more difficult and it is most likely not possible to eliminate these events completely. For a very low energetic electron-positron pair, these events are almost indistinguishable from  $\mu^+\mu^-$  events. In this case,

$E_{\pm} \approx 5.29 \text{ GeV}$ ,  $\cos \eta \approx 1$  and  $M_X^2 \approx 0 \text{ GeV}^2$ , which matches the expectation for the signal events. The muon likelihood also does not reject these events because the two reconstructed particles are actually muons.

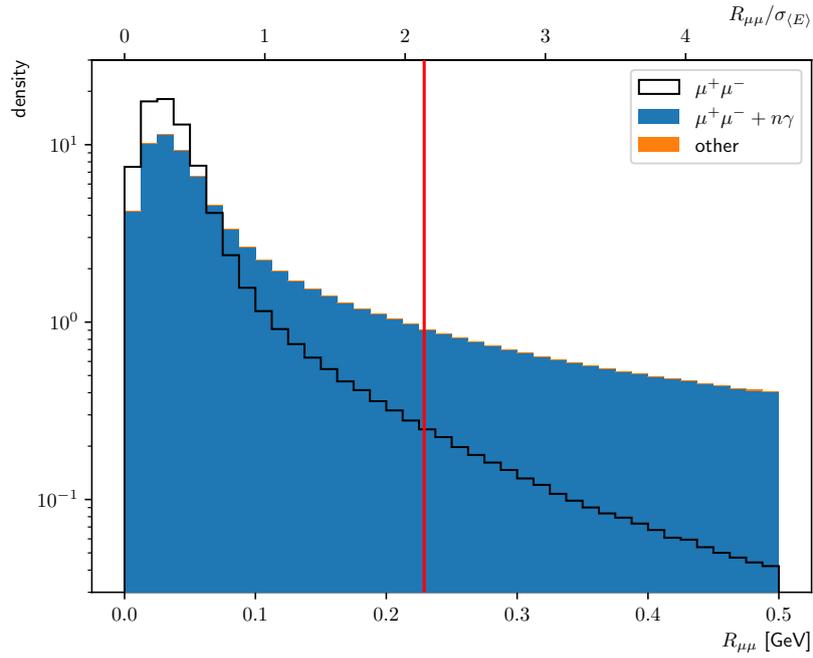
Figure 3.12 show the radial energy deviation for  $\mu^+\mu^-$  (black) and  $e^+e^-\mu^+\mu^-$  (orange) events. Similarly, figs. 3.13 and 3.14 show the acollinearity and the missing mass squared respectively.

While overall the distributions of  $e^+e^-\mu^+\mu^-$  events are very similar to the  $\mu^+\mu^-$  distributions, the peaks are a bit broader. Therefore, tighter cuts on the aforementioned variables ( $R_{\mu\mu}$ ,  $\eta$  and  $M_X^2$ ) still show some room for improvement.

The next steps to measure the Weinberg angle are to refine the selection on the radial energy deviation, the acollinearity and the muon identification, once it is finalized at Belle II. Also the systematic uncertainties have to be studied.

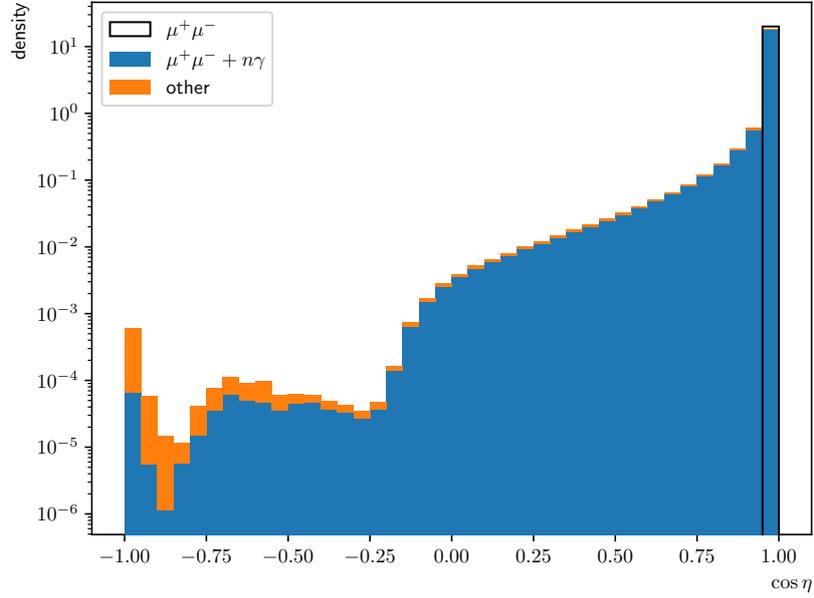


(a) The radial energy deviation shown in the complete available range.

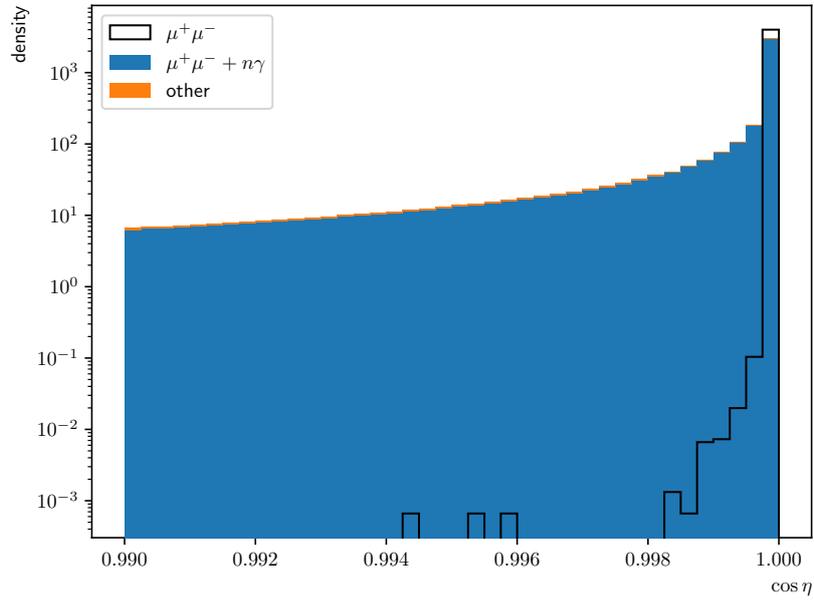


(b) The radial energy deviation in the region around the signal peak. The horizontal axis on top of the histogram shows the radial energy deviation in terms of the sample standard deviation of the energies of both muons. The red line marks the optimal selection bound as found in the fit in section 3.4.

**Figure 3.6:** The radial energy deviation of background and signal (black line) events.

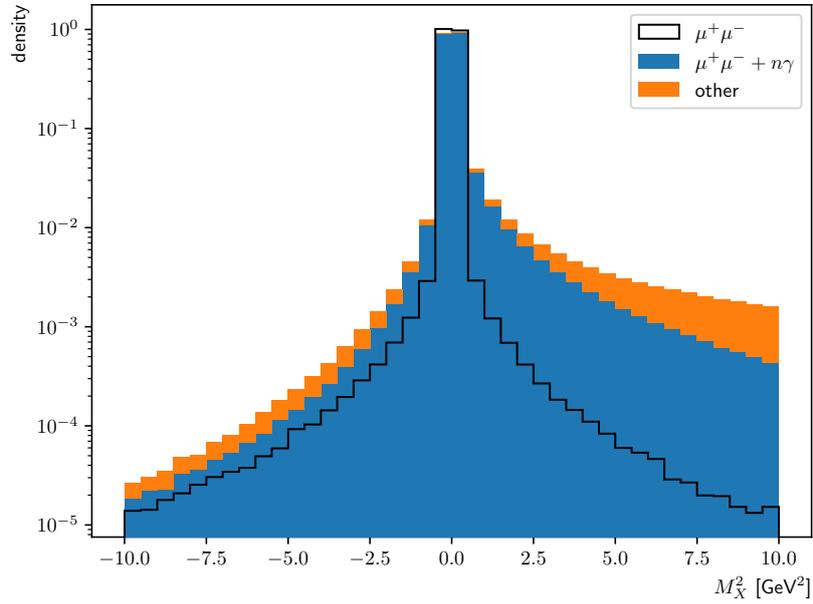


(a) The acollinearity shown in the complete available range.

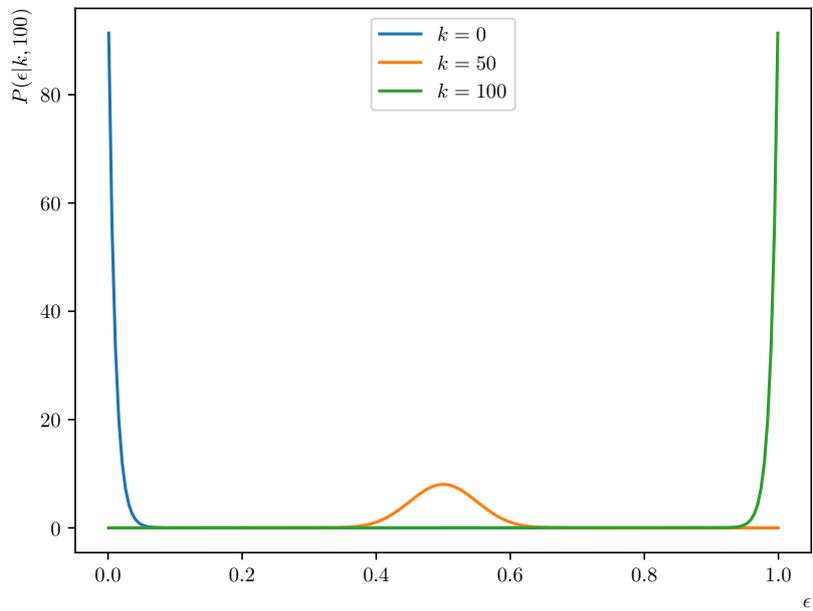


(b) The acollinearity in the region around the signal peak.

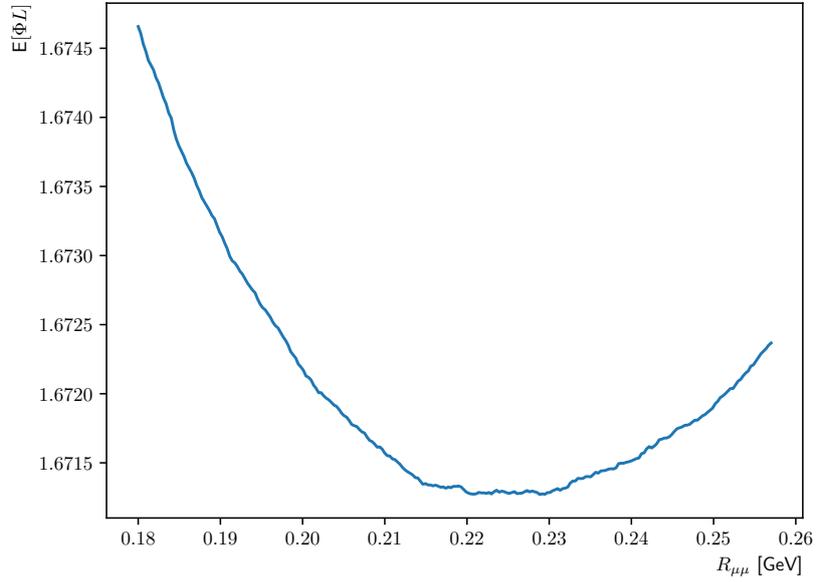
**Figure 3.7:** The acollinearity of the muon candidates of background and signal (black line) events.



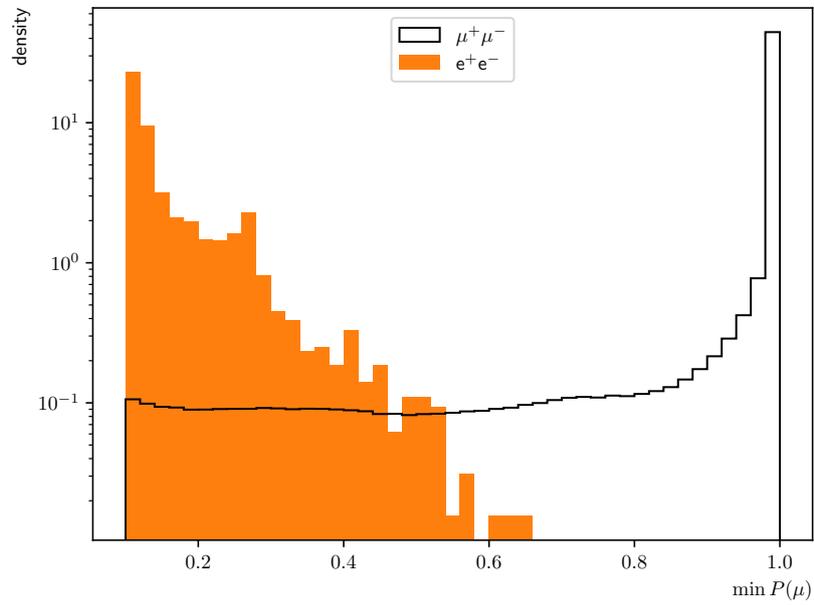
**Figure 3.8:** The missing invariant mass squared of background and signal (black line) events.



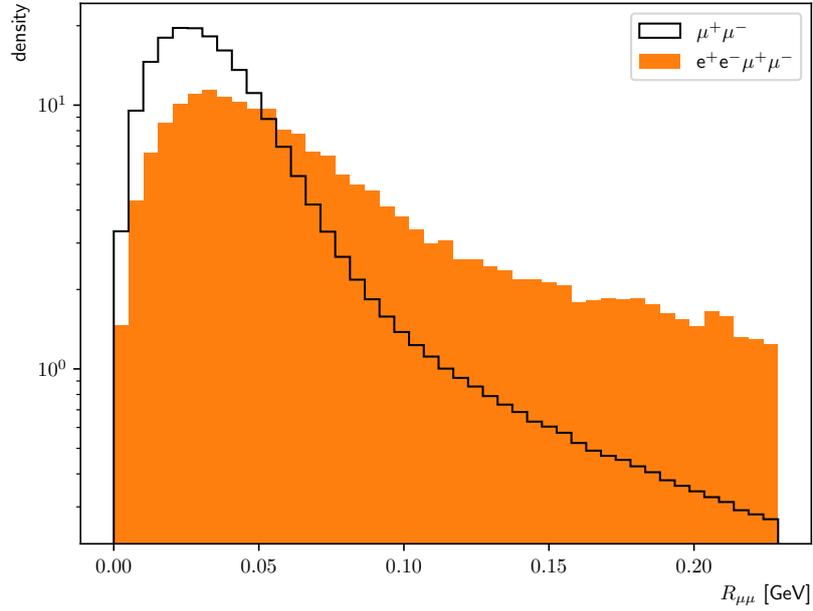
**Figure 3.9:** The posterior probability density function of the efficiency  $\epsilon$  with a flat prior for  $k = 0, 50, 100$  and  $N = 100$ .



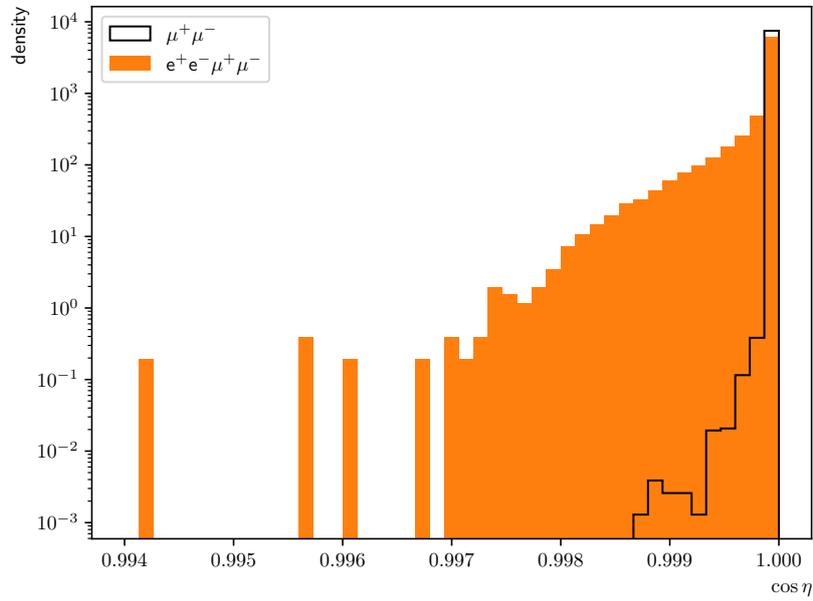
**Figure 3.10:** The expected relative statistical variance in the region around the minimum of the radial energy deviation fit at 0.229 GeV.



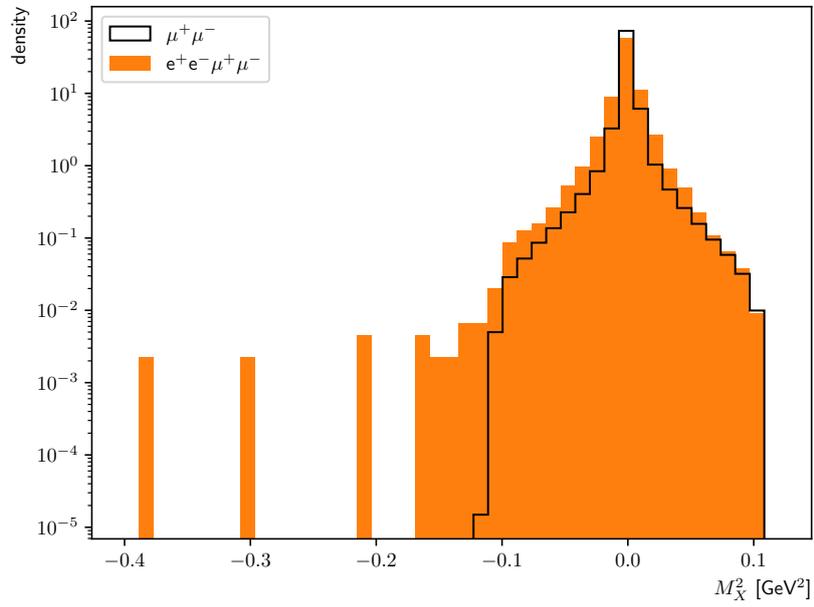
**Figure 3.11:** The minimal muon likelihood  $\min P(\mu)$  for  $e^+e^-$  (orange) and  $\mu^+\mu^-$  (black line) events after applying  $R_{\mu\mu} < 0.229$  GeV.



**Figure 3.12:** The radial energy deviation for  $e^+e^-\mu^+\mu^-$  (orange) and  $\mu^+\mu^-$  (black line) events after applying  $R_{\mu\mu} < 0.229$  GeV.



**Figure 3.13:** The acollinearity for  $e^+e^-\mu^+\mu^-$  (orange) and  $\mu^+\mu^-$  (black line) events after applying  $R_{\mu\mu} < 0.229$  GeV.



**Figure 3.14:** The missing invariant mass squared  $M_X^2$  for  $e^+e^-\mu^+\mu^-$  (orange) and  $\mu^+\mu^-$  (black line) events after applying  $R_{\mu\mu} < 0.229$  GeV.

## Chapter 4

# Study of the muon identification performance

After the optimization of the event selection, we obtain a data set of polar angles to extract the Weinberg angle. However, the measured angular distribution is not the true underlying distribution. It is distorted by detection efficiencies and resolution effects. The efficiency is calculated from simulated events. Since the final analysis is done on recorded data, it is necessary to understand how the efficiency needs to be scaled from simulated to recorded events. I do this for the muon identification efficiencies, which I study in the independent channel  $\mu^+\mu^-\gamma$ , where the photon has an energy of at least 1 GeV.

Previous studies [30] have shown that the particle identification performance depends on the momentum and polar angle of the track of interest in the laboratory frame. A study by P. Feichtinger in the  $e^+e^- \rightarrow \tau^\pm(1P)\tau^\mp(3P)$  channel [38], where one tau decays to one charged particle and the other to three, showed that the performance also depends on the activity around the track. This means that the performance changes, if there is a lot of activity (other tracks) in the detector around the track of interest. We call tracks with almost no other tracks around them isolated and consecutively will talk about the track isolation, instead of the activity around a track.

To study if this dependence is also present in other channels, I perform a similar study in the  $e^+e^- \rightarrow \mu^+\mu^-\gamma$  channel, where most tracks should be well isolated. If a similar dependence as in the  $e^+e^- \rightarrow \tau^\pm(1P)\tau^\mp(3P)$  channel is seen, this would be a strong motivation to investigate the influence of the track isolation in other channels with a higher number of charged final state particles. In addition, we then calculate corrections to the simulation of muons at Belle II based on the ratio of the muon identification efficiencies in recorded and simulated events.

### 4.1 The Belle II charged particle identification

The Belle II-framework currently has two methods to determine the particle species of a given candidate, namely the global particle identification and the binary particle identification. These define the probability that a candidate belongs to the particle species  $i$ . The final state species considered at Belle II are electrons  $e$ , muons  $\mu$ , pions  $\pi$ , kaons  $K$ , protons  $p$  and deuterons  $d$ . The detec-

tors providing information for the particle identification have been described in section 3.1. For muons, the  $K_L^0$  and muon detector is of particular importance.

For each final state species  $i$ , we use the set of information  $x$  from the sub-detector  $d$  to calculate the likelihood  $\mathcal{L}_d(x|i)$  for that detector. Using Bayes' Theorem, the probability  $P(i|x)$  for the candidate to belong to a certain particle species is

$$P(i|x) = \frac{\prod_d \mathcal{L}_d(x|i)}{\sum_j \prod_d \mathcal{L}_d(x|j)} = \frac{\mathcal{L}_i}{\sum_j \mathcal{L}_j}. \quad (4.1)$$

The sum in the denominator takes into account all the different final state particle species. Equation (4.1) defines the global likelihood ratio at Belle II. This assumes the measurements of the sub-detectors to be independent.

If the goal is only to differentiate between two separate particle hypothesis  $i$  and  $j$ , often the so-called binary likelihood does a better job

$$P(i/j|x) = \frac{\mathcal{L}_i}{\mathcal{L}_i + \mathcal{L}_j}. \quad (4.2)$$

There are also currently other approaches in development, where the detector likelihoods are reweighted. These weights are either implemented based on studies or using neural networks [39].

Regardless of which method will be the finalized version of the particle identification at Belle II, the agreement between recorded and simulated events needs to be studied. In the following, I study the agreement using the global muon likelihood.

## 4.2 The track isolation

At Belle II, track finding for charged particles takes into account detector hits at the different sub-detectors and interpolates them into a helix, due to the magnetic field. This helix is only an approximation of the real trajectory since the charged particles interact with the detector material and the magnetic field is not homogeneous. The helix is then extrapolated throughout the detector. For more information see the Belle II physics book [29].

The initial definition of the track isolation variable looked at the entry surfaces of the different sub-detectors except for the silicon pixel detector and the silicon vertex detector. It then calculated the distance of other surrounding tracks at one surface to the track of interest. We advocated to change the definition of the track isolation variable since the initial definition would consider two tracks as non-isolated, if they simply crossed at one detector surface and then diverged again. They are therefore only locally non-isolated, while they are well isolated throughout most of the detector. This old definition was therefore abandoned in favor of an isolation score presented by M. Milesi in an internal presentation [40].

The track isolation score defines an isolated track by taking into account information from the central drift chamber, the aerogel ring-imaging Cherenkov detector, the time-of-propagation counter, the electromagnetic calorimeter and the  $K_L^0$  and muon detector. For the central drift chamber, it defines nine layers, which correspond to the nine superlayers of the central drift chamber. For the aerogel ring-imaging Cherenkov detector and the time-of-propagation counter,

the only layer is the inner surface of the sub-detectors. For the electromagnetic calorimeter, the two layers are the inner surface and one about 15 cm further out. Finally, for the  $K_L^0$  and muon detector, the single layer is again the entry surface. Since the  $K_L^0$  and muon detector extends in the radial direction, it would make sense to define further layers. However, the magnetic field configuration does not allow for that in the current simulation. For each sub-detector, the different layers are approximated as cylinders with the dimensions approximating reality.

At each sub-detector different thresholds are defined based on the granularity of the sub-detectors. The thresholds are: 5 cm for the central drift chamber, 22 cm for the time-of-propagation counter, 10 cm for the aerogel ring-imaging Cherenkov detector, 36 cm for the electromagnetic calorimeter and 20 cm for the  $K_L^0$  and muon detector. Using this, the number of layers  $n_d$  is counted, at which at least one extrapolated track is closer to the track of interest than the respective threshold value at that layer. For the central drift chamber and the time-of-propagation counter, a 2D distance in the azimuth and radial plane is calculated at each layer since they are only segmented in the radial and azimuth plane. For the other sub-detectors, a 3D distance connecting the crossing points of all the tracks to the track of interest is calculated at each layer.

This currently does not take into account activity from neutral particles or detector hits, which could not be reconstructed into a track, for example photon clusters in the electromagnetic calorimeter. However, there are considerations in the Lepton ID group at Belle II to change this in an updated version of the isolation score.

The ablation metric  $\Delta s_d = s_d - s$  [30] uses the separation metric  $s$ , which differentiates between muons and pions based on their likelihoods. The separation metric is bounded between  $-1$  and  $1$  and  $s_d$  is the same separation metric, but excluding the sub-detector  $d$ . A positive ablation metric means the information from this sub-detector worsens the overall particle identification performance, while a negative ablation metric means that this sub-detector improves the overall particle identification performance. For sub-detectors, that worsen the muon particle identification performance, the ablation metric is set to 0. Therefore, the track isolation score weighs the detector input to the track isolation based on the importance of the detector for the identification of the particle species of interest. This effectively excludes them from the track isolation calculation. Using the ablation metric, the semi-continuous isolation score  $I$  is

$$I = 1 - \left[ \sum_d -\Delta s_d \left( \frac{n_d}{N_d} \right) \right], \quad (4.3)$$

where the sum takes all the discussed sub-detectors into account.  $N_d$  is the total number of layers of the sub-detector  $d$ . The isolation score is then normalized to be in the range  $[0, 1]$ , where an isolation score of 1 signifies that the track is completely isolated.

### 4.3 Selection criteria

I perform the track isolation study with the available data set of Belle II with an integrated luminosity of  $(189.88 \pm 0.01) \text{ fb}^{-1}$ . The uncertainty is statistical only. This data set covers proc 12 chunks 1 and 2, as well as buckets 16 to 25. The processing of the recorded and simulated data was done by A. Narimani and the simulated data (MC) contains  $1.148 \times 10^9 \mu^+ \mu^- \gamma$  events.

Events need to fulfill the high level trigger requirements of the `hlt_radmumu skim` [30]. These require at most three track candidates per event with a minimal transverse momentum of 0.2 GeV. In addition, the point of closest approach to the interaction point of each track candidate has to be closer than 2 cm in the radial direction and closer than 4 cm in the  $z$ -direction. The tracks need to have at least one hit in the CDC or the KLM and a cluster with an energy of less than 0.4 GeV associated with them. The events are required to have a recoil momentum of at least 0.1 GeV, which is consistent with the emission of a single photon. In these events, there needs to be one pair of tracks, which have a cluster energy in the ECL of at most 0.25 GeV for both tracks and for which the norm of the azimuth between the two clusters has to be above  $\pi/2$ . Finally, the momentum of the higher momentum track has to be above 1 GeV and the momentum of the lower momentum track has to be below 3 GeV. Since these requirements are imposed on all events, to good approximation, high-level trigger effects should cancel out when calculating efficiencies and ratios of efficiencies. Therefore, the influence of the high-level trigger on this study was not investigated.

For the reconstruction, only events with exactly two tracks with a point of closest approach to the interaction point closer than 2 cm in the radial direction and 5 cm in the  $z$ -direction are allowed. The event also needs to have at least one photon with an energy of at least 1 GeV. In addition, the polar angle  $\theta_\gamma$  of the photon has to be between  $-0.8660$  and  $0.9563$ . This is the angular acceptance of the ECL. Finally, the weighted sum of the ECL crystals of the photon cluster has to be above 1.5.

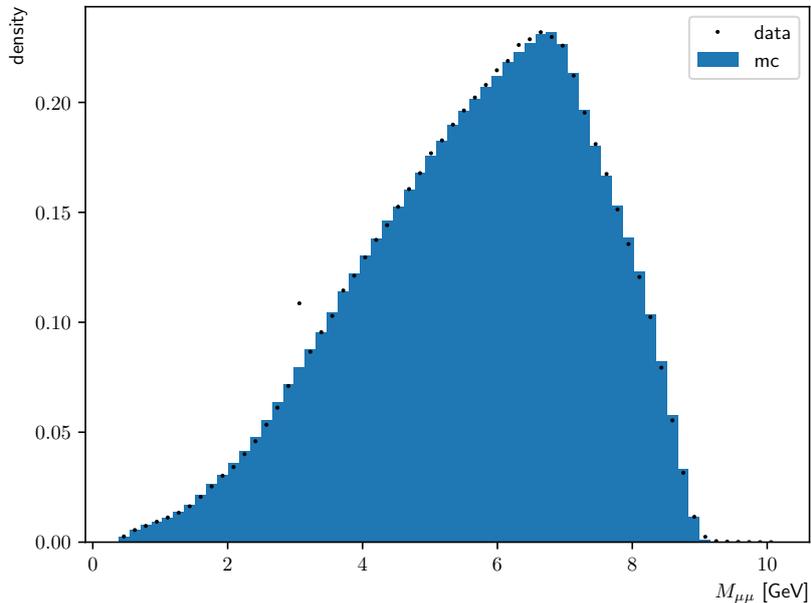
In addition, I require the mass of the  $\mu^+\mu^-\gamma$  system to be between 10.2 GeV and 10.8 GeV and the mass of the muon pair  $M_{\mu\mu}$  to be below 3.0964 GeV or above 3.0974 GeV. This excludes the region of around five times the width of the  $J/\psi(1S)$ -resonance around its peak position. Figure 4.1 shows the mass of the muon pair in recorded (data, black points) and simulated events (mc, blue histogram). The outlier at 3 GeV only shows up in the recorded events and matches well with the expectation for the  $J/\psi(1S)$ , which has a mass of  $(3096.900 \pm 0.006)$  MeV and a width of only  $(92.6 \pm 1.7)$  keV. The exclusion of this region therefore guarantees a more consistent comparison between recorded and simulated data. The  $J/\psi(1S)$  is the only visible resonance in the mass spectra of the  $\mu^+\mu^-$  and  $\mu^+\mu^-\gamma$  systems.

Finally, I use a tag-and-probe method, meaning I study the efficiencies for the  $\mu^+$  (probe-muon), by tagging the events with the requirement that the global muon likelihood of the  $\mu^-$  has to be larger than 0.9 and vice-versa.

Overall these selection criteria produce a very clean sample [30], so we are certain that only  $\mu^+\mu^-\gamma$  events are selected.

## 4.4 The muon identification performance

The muon likelihood efficiencies are calculated in bins of the laboratory frame momentum  $p$  and the laboratory frame polar angle  $\theta$  using eq. (3.25) and the global muon likelihood. For each bin,  $N$  in the denominator of the efficiency is given by the number of events in the bin and  $k$  is the number of events, where the probe-muon has a muon likelihood above 0.9. As the probe-muon I choose the  $\mu^+$ . Since the distributions for the  $\mu^+$  and the  $\mu^-$  are very similar, I only show those of the  $\mu^+$ . The momentum bins are six disjoint contiguous intervals



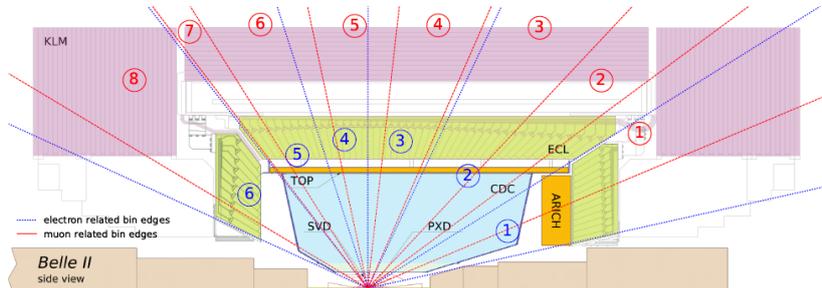
**Figure 4.1:** The density histograms of the  $\mu^+\mu^-$  invariant mass in generated events (blue histogram) and recorded events (black points).

[0.2, 0.7, 1.0, 2.0, 3.0, 5.0, 7.0] GeV. They are a combination of the official bins from the Lepton identification Group (LID) [30].

The angular binning consists of 10 disjoint contiguous intervals [0.22, 0.40, 0.64, 0.82, 1.16, 1.46, 1.78, 2.13, 2.22, 2.60, 2.71] rad. It reflects the official LID binning [30] and is shown in fig. 4.2 in red. In addition, I extend the angular binning at the edges down to 0.22 rad and up to 2.71 rad. The angle of 0.22 rad marks the end of the ECL forward endcap, while 2.71 rad marks the end of the ECL backward endcap. I define the region of 0.64 rad to 2.22 rad as the KLM barrel region.

The total set of events is divided into two parts based on the track isolation score of the probe-muon. If the track isolation score is below 0.999, the track is called non-isolated. If the track isolation score is above 0.999, it is called an isolated track. The value 0.999 is based on a study by M. Milesi in  $J/\psi \rightarrow \mu^+\mu^-$  events [40]. Using this value, 99.87% of  $\mu^+$  in recorded events are isolated. In simulated events the percentage of isolated  $\mu^+$  is even a bit higher at 99.90%.

Figure 4.3 shows the muon likelihood efficiency in bins of laboratory frame momentum of the  $\mu^+$ . The upper plot shows the muon likelihood efficiencies in recorded events, while the bottom plot shows the efficiencies in simulated events. The combination of non-isolated and isolated tracks is plotted in gray. The horizontal bars mark the extend of the bin and should not be confused with error bars. Similarly, the points marked with an "x" are calculated for the whole bin, but displaced a bit to the right from the bin center for better visibility. Uncertainties on the efficiency are also plotted. However, for the isolated tracks they are too small to be visible. Overall, the distributions behave similarly between simulated and recorded events. The muon likelihood efficiency



**Figure 4.2:** A view from the side of the Belle II detector taken from [38]. The angles determining the binning in [30] are shown in red. I extend the binning to the outermost blue lines.

is consistently lower for non-isolated tracks, than for isolated tracks. In most bins, the efficiency of the combined set is almost identical with the efficiency of the isolated data set. This is expected since almost all the tracks are well isolated. Using a higher value does not increase the amount of non-isolated tracks significantly. The uncertainties in the non-isolated bins are still under control, so that we can make statements about the behavior of the distributions.

We expect the muon likelihood efficiency to rise from very low momenta, since low momentum tracks might not reach all detector parts, in particular the KLM, which is very important for the muon identification. At larger momenta, it should then plateau.

As expected, the muon likelihood efficiency in fig. 4.3 increases with increasing momentum until it reaches a plateau at about 2 GeV. The plateau drops in the highest momentum bin. This drop had been seen previously by the Lepton ID group, but was not studied. For the tracks in the angular acceptance of the KLM (from 0.4 rad to 2.60 rad), the drop disappears. We conclude that the drop is due to tracks, which do not reach the KLM. Furthermore, the efficiencies of tracks in the KLM are almost identical for isolated and non-isolated tracks, once the tracks have a momentum of more than 1 GeV. For low momenta, the influence of the KLM is not as visible since a minimum momentum is needed for a particle to reach the KLM. As the KLM is the outermost detector, low momentum tracks will not reach it. This shows clearly that the KLM is the most important detector for the muon identification, as expected.

Figure 4.4 shows the muon likelihood efficiency as a function of the laboratory frame polar angle of the  $\mu^+$  in recorded (top) and simulated (bottom) events. The difference in the efficiencies of isolated and non-isolated tracks is even more striking. For the isolated tracks, the efficiency is relatively constant over a large angular region, only showing significant drops towards the end of the endcaps (lowest and largest angles) and at the transition from endcaps to barrel at around 0.6 rad and 2.2 rad. For the non-isolated tracks, this behavior is mirrored and amplified. The efficiency is even lower, where it drops for the isolated tracks. In particular, the plateau in the barrel region (0.64 rad – 2.2 rad) is less stable and does not cover the same range. Instead, the muon likelihood efficiency is best in the region 1.2 rad – 1.8 rad. The forward endcap (< 0.8 rad) performs better than the backward endcap (> 2.2 rad). This matches the expect-

tation from the asymmetric setup at SuperKEKB. Since the electron beam has a higher energy than the positron beam, the tracks are boosted in the forward direction. The Belle II detector was constructed with this in mind, so more focus was put into optimizing the forward region of the detector, instead of the backward region. Therefore, the dedicated particle identification detectors, the time-of-flight counter and the aerogel ring-imaging Cherenkov detector, do not have a backward endcap and the performance is worse in that region.

To better compare the agreement between the distributions in recorded and simulated events, I calculate the ratio of muon likelihood efficiencies according to eqs. (3.53) and (3.54) with the efficiency in recorded events  $\epsilon_d$  in the numerator and the efficiency in simulated events  $\epsilon_{mc}$  in the denominator. A ratio very close to 1 means almost perfect agreement between the efficiency in recorded and in simulated events. A higher or lower ratio means that the performance of the Belle II detector is different in simulated events and in recorded events.

Figure 4.5 shows the ratio of efficiencies in recorded over simulated events in bins of laboratory frame momentum of the  $\mu^+$ . The subplot underneath shows the relative statistical uncertainty for isolated (red) and non-isolated (blue) tracks. The ratio of efficiencies for the isolated tracks dips at around 0.6 GeV, before it increases, approaching 1. Over the complete momentum range, it indicates good, but not perfect agreement between simulated and recorded events. The relative uncertainties are on the percent-level or smaller.

For non-isolated tracks, there is no obvious trend. The ratio is consistently small at lower momenta, with exception of the very first bin. At higher momenta, the ratio increases and sometimes even overshoots 1.

While a quantitative analysis is difficult, the relative uncertainties can be as large as a few 10%, the ratio of efficiencies behaves very different for isolated and non-isolated tracks. This is good news, as it strengthens the claim of the previous analysis [38] that the muon identification performance depends on the track isolation.

Figure 4.6 shows the ratio of efficiencies in recorded over simulated events in bins of the laboratory frame polar angle of the  $\mu^+$ . Again the subplot below shows the relative statistical uncertainties for the two isolation bins. For the isolated tracks, the ratio of efficiencies is overall in very good agreement with 1. It drops slightly in the backward endcaps of the KLM and ECL. The statistical uncertainties are again well under control, as the relative uncertainties are below 1%.

There is no discernible trend for the non-isolated tracks and the statistical uncertainties can be large. The ratio of efficiencies fluctuate between the different bins. The largest deviation from a ratio of 1 is the bin from 0.64 rad to 0.82 rad. There the ratio of efficiencies is significantly above 1. This is the last bin of the KLM barrel before the gap to its forward endcap, which is a possible explanation.

Although the amount of data is limited for the non-isolated tracks, there is no clear indication of different behavior between isolated and non-isolated tracks as a function of the polar angle.

Figure 4.7 shows the ratio of efficiencies for isolated  $\mu^+$  tracks. The upper plot shows the ratio itself, while the lower plot shows the relative statistical uncertainty  $\hat{\sigma}$  calculated from eqs. (3.53) and (3.54) on the ratio. The solid red lines mark the beginning and end of the angular coverage of the KLM. If a bin is empty, then it did not contain enough events to calculate the ratio of

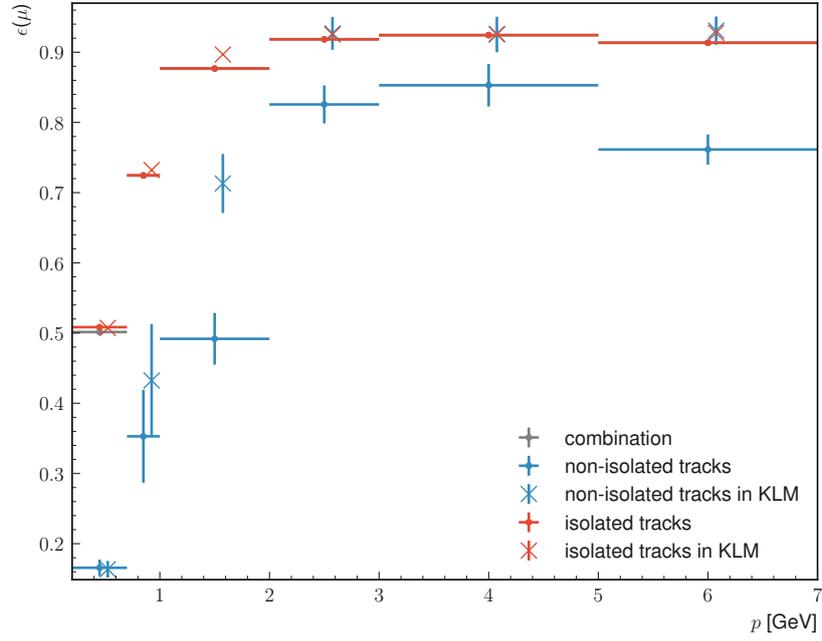
efficiencies.

Overall the ratio of efficiencies is relatively uniform for tracks above 1 GeV in the barrel region. It drops a bit in the backward endcap of the KLM. In the forward ECL endcap the ratio is lower than in the KLM barrel. In the backwards ECL endcap, there are even fewer tracks due to the asymmetry of SuperKEKB. Independent of the angle, the efficiency ratio for low momentum tracks is lower, as we expect it from the one dimensional distribution in fig. 4.5. The transition from endcaps to barrel (0.64 rad – 0.82 rad and 2.13 rad – 2.22 rad) shows a slightly worse agreement between simulation and recorded events so there could be a mis-modeling of the detector. Overall the relative statistical uncertainties are well under control as fig. 4.7b shows.

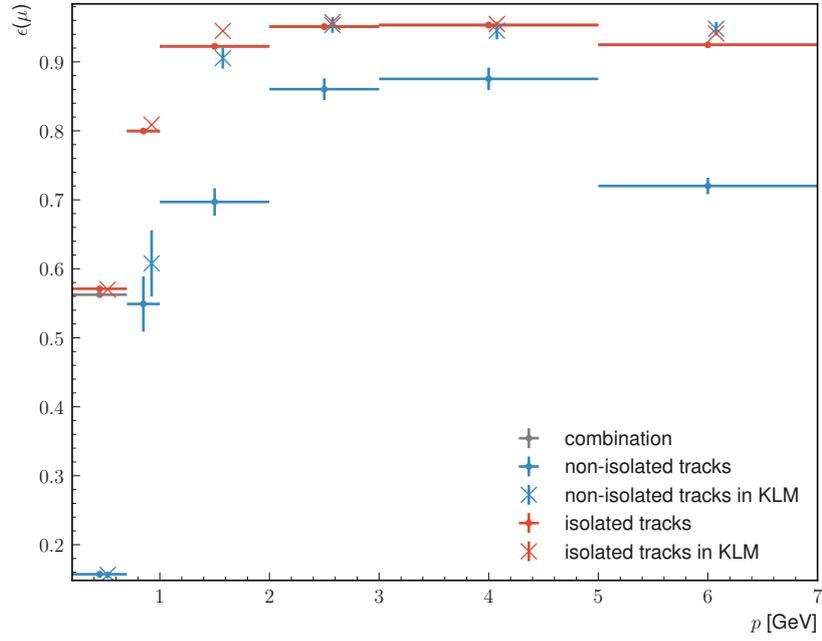
Figure 4.8 shows the same momentum-polar angle distribution for the non-isolated  $\mu^+$  tracks. The backwards endcap of the ECL (to the right of the red line at 2.6 rad) is not populated at all. For the highest momentum bin, the backward region of the KLM (part of the barrel and the backward endcap) is also not populated. Therefore, the forward region contributes more, in particular the first polar angle bin, which is outside the KLM acceptance. This explains the drop of the efficiency in the last momentum bin of fig. 4.3. There are missing entries in the region 2.13 rad – 2.22 rad, where the ECL and KLM transition from their barrel to the backward endcap. The missing entries are due to low amount of non-isolated tracks. This makes it difficult to evaluate if the efficiency ratio in the transition from barrel to endcaps shows the same behavior as in fig. 4.7a. In the barrel region, figs. 4.7a and 4.8a show the same trend, with the ratio being closer to 1 for momenta larger than 1 GeV. However, for the non-isolated tracks in fig. 4.7a, the efficiency ratio is worse and is affected by larger statistical uncertainties.

Figures 4.7 and 4.8 are the most important plots of this study since they are the visualization of the correction tables used to scale the simulated efficiencies to the ones in recorded data. I give the correction tables for the  $\mu^+$  in tables 4.1 and 4.2. The relative statistical uncertainty  $\hat{\sigma}$  indicates the size of the systematic uncertainty that an analysis picks up, when correcting its efficiencies. Tables 4.3 and 4.4 are the same correction tables for the  $\mu^-$ .

Finally, I want to stress the importance of the conclusion of fig. 4.3. Muon tracks with low momentum or that are outside the angular acceptance of the KLM suffer a drastic drop in the efficiency, in particular the non-isolated tracks. Studies that depend on these tracks are therefore challenging and affected by large systematic uncertainties.

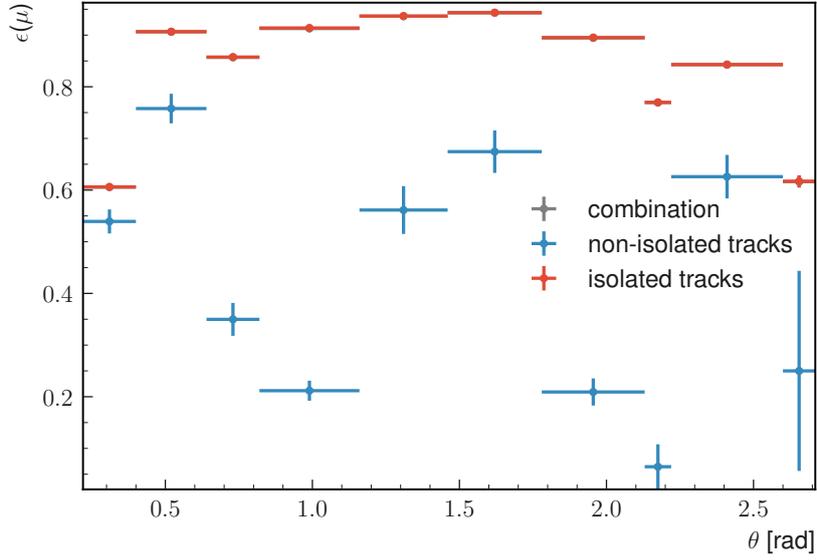


(a)  $\mu^+$  efficiency in recorded events.

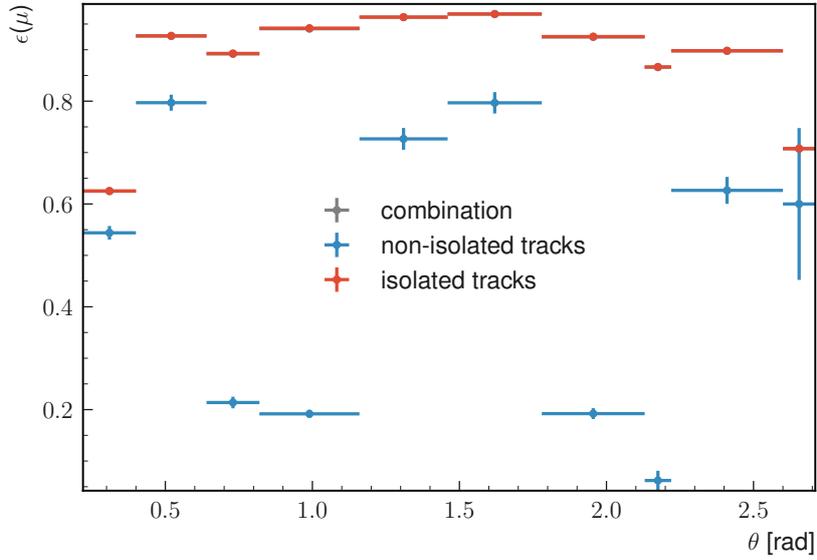


(b)  $\mu^+$  efficiency in simulated events.

**Figure 4.3:** The subfigures show the muon identification efficiency as a function of the momentum for the  $\mu^+$ . Grey are all the candidates. Blue are all candidates with an isolation score below 0.999. Red are all the candidates with an isolation score above 0.999. The bins marked with "x" are required to be within the angular acceptance of the KLM.

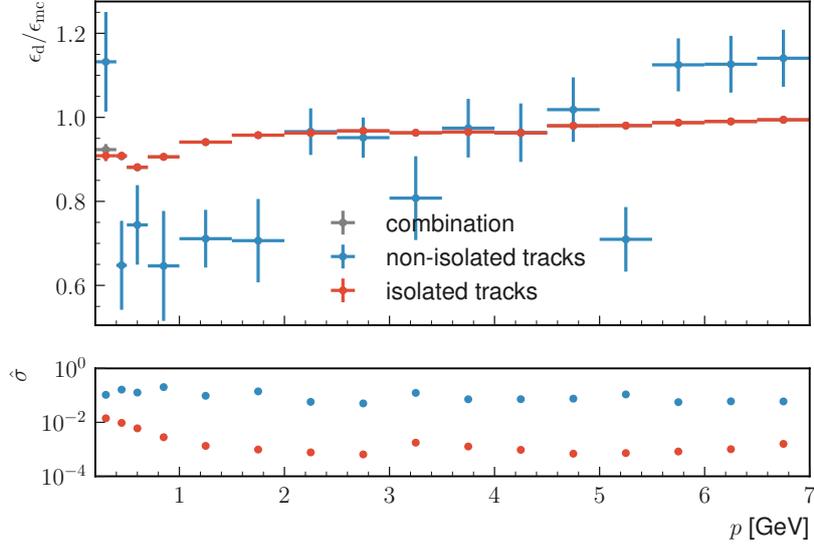


(a)  $\mu^+$  efficiency in recorded events.

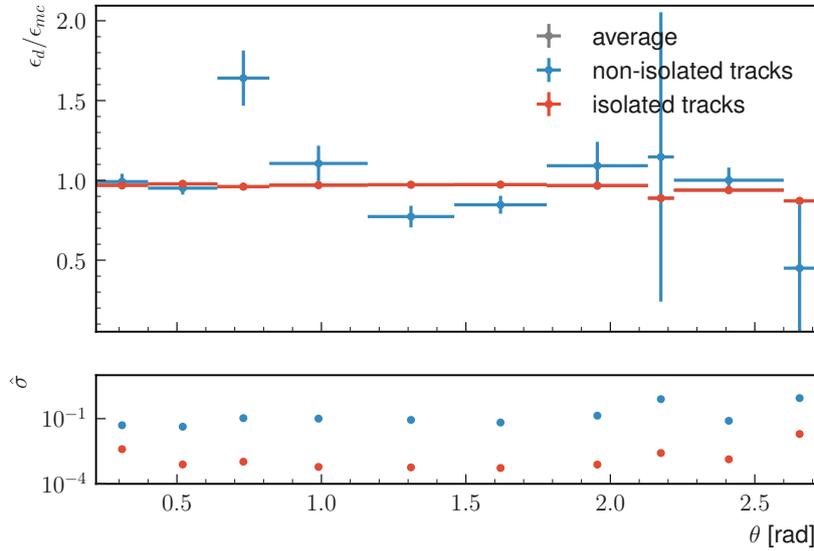


(b)  $\mu^+$  efficiency in simulated events.

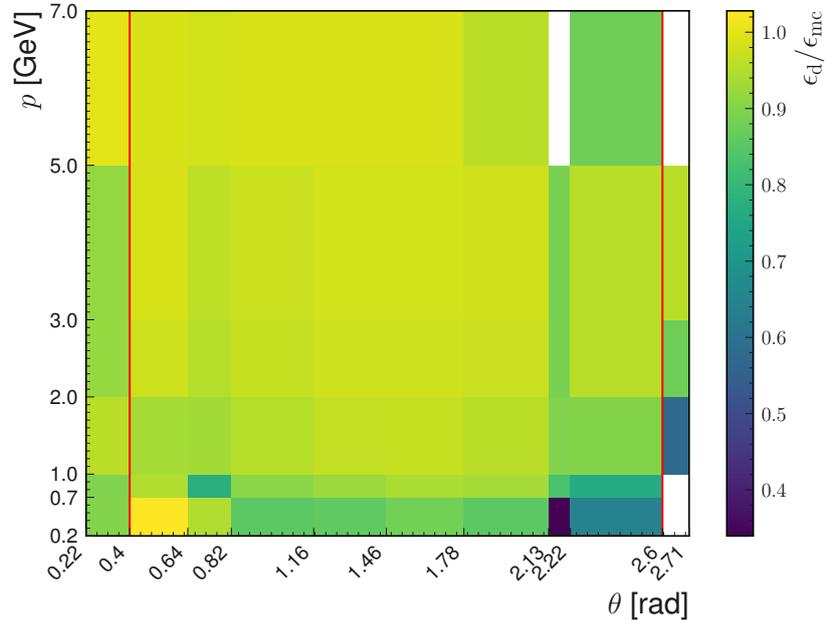
**Figure 4.4:** The subfigures show the muon identification efficiency as a function of the polar angle for the  $\mu^+$ . Grey are all the candidates. Blue are all candidates with an isolation score below 0.999. Red are all the candidates with an isolation score above 0.999.



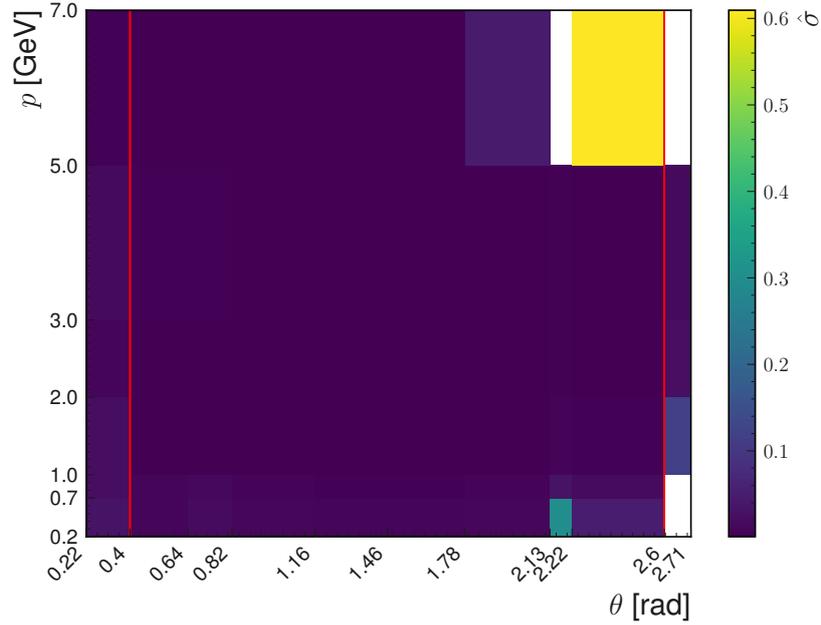
**Figure 4.5:** The subfigures show the ratio of efficiencies (top) and its uncertainty (bottom) as a function of their momentum for the  $\mu^+$ . Grey are all the candidates. Blue are all candidates with an isolation score below 0.999. Red are all the candidates with an isolation score above 0.999. The smaller subplot underneath shows the relative uncertainty for the two isolation bins.



**Figure 4.6:** The subfigures show the ratio of efficiencies (top) and its uncertainty (bottom) as functions of their polar angle for the  $\mu^+$ . Grey are all the candidates. Blue are all candidates with an isolation score below 0.999. Red are all the candidates with an isolation score above 0.999.

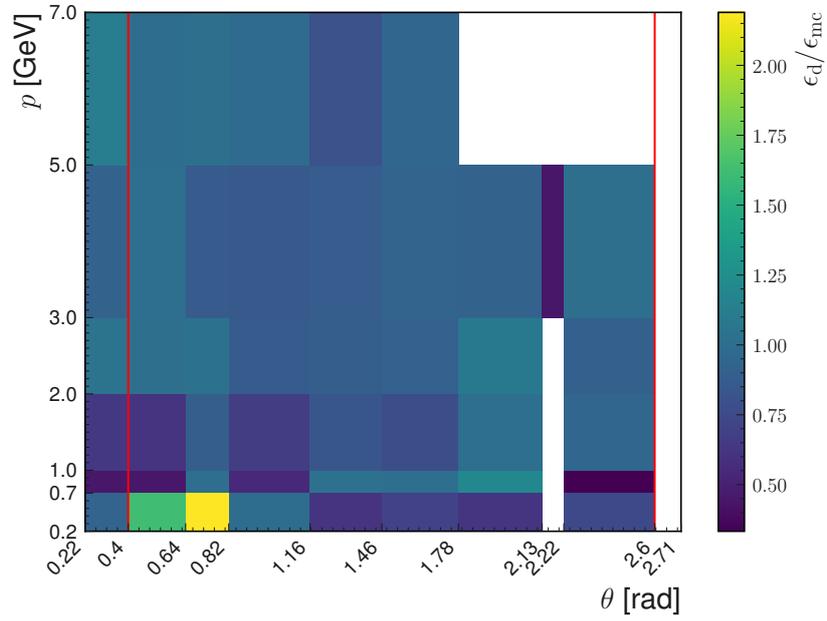


(a) The ratio of the muon identification efficiencies in bins of the polar angle and momentum of isolated  $\mu^+$ .

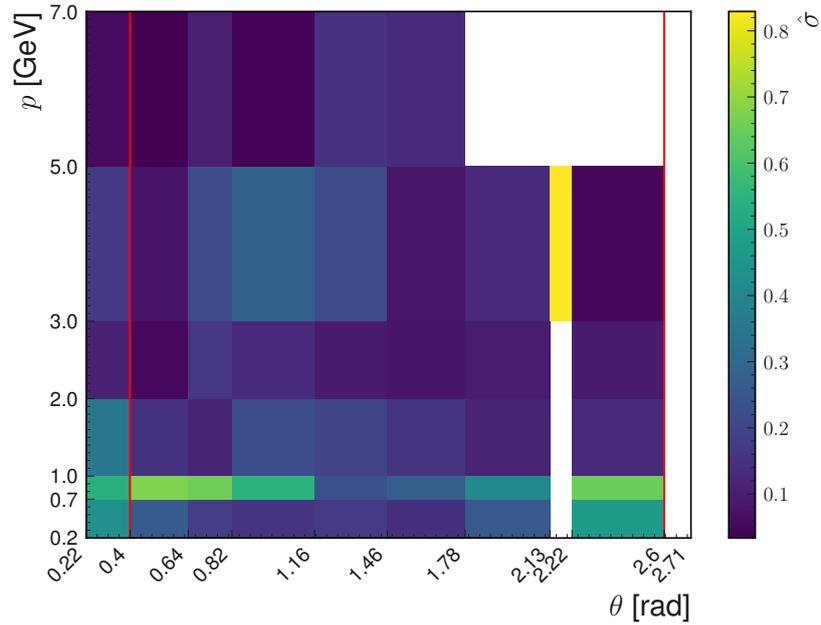


(b) The relative statistical uncertainty on the ratio of the muon identification efficiencies in bins of the polar angle and momentum of isolated  $\mu^+$ .

**Figure 4.7:** The ratio of efficiencies (top) and its relative statistical uncertainty (bottom) for the  $\mu^+$  in bins of the polar angle and momentum for isolated tracks (isolation score above 0.999). The solid red lines mark the beginning and end of the KLM.



(a) The ratio of the muon identification efficiencies in bins of the polar angle and momentum of non-isolated  $\mu^+$ .



(b) The relative statistical uncertainty on the ratio of the muon identification efficiencies in bins of the polar angle and momentum of non-isolated  $\mu^+$ .

**Figure 4.8:** The ratio of efficiencies (top) and its relative statistical uncertainty (bottom) for the  $\mu^+$  in bins of the polar angle and momentum for non-isolated tracks (isolation score below 0.999). The red lines mark the beginning and end of the KLM.

**Table 4.1:** The efficiency ratio for isolated  $\mu^+$  in bins of momentum and polar angle.

$p$ [GeV]	$\theta$ [rad]	$\epsilon_d/\epsilon_{mc}$	$\hat{\sigma}$ [%]	$p$ [GeV]	$\theta$ [rad]	$\epsilon_d/\epsilon_{mc}$	$\hat{\sigma}$ [%]
0.2–0.7	0.22–0.40	0.901	3.02	2.0–3.0	0.22–0.40	0.919	0.99
	0.40–0.64	1.028	0.88		0.40–0.64	0.976	0.16
	0.64–0.82	0.946	1.88		0.64–0.82	0.952	0.18
	0.82–1.16	0.855	1.24		0.82–1.16	0.968	0.10
	1.16–1.46	0.859	0.99		1.16–1.46	0.976	0.09
	1.46–1.78	0.881	0.84		1.46–1.78	0.976	0.09
	1.78–2.13	0.854	1.23		1.78–2.13	0.969	0.13
	2.13–2.22	0.340	29.64		2.13–2.22	0.890	0.42
	2.22–2.60	0.642	4.73		2.22–2.60	0.953	0.19
2.60–2.71	-	-	2.60–2.71	0.877	2.23		
0.7–1.0	0.22–0.40	0.895	2.18	3.0–5.0	0.22–0.40	0.916	1.72
	0.40–0.64	0.947	0.80		0.40–0.64	0.985	0.33
	0.64–0.82	0.775	1.31		0.64–0.82	0.958	0.38
	0.82–1.16	0.907	0.63		0.82–1.16	0.974	0.16
	1.16–1.46	0.923	0.52		1.16–1.46	0.981	0.10
	1.46–1.78	0.940	0.44		1.46–1.78	0.981	0.07
	1.78–2.13	0.936	0.62		1.78–2.13	0.978	0.10
	2.13–2.22	0.831	3.01		2.13–2.22	0.888	0.34
	2.22–2.60	0.764	1.86		2.22–2.60	0.954	0.15
2.60–2.71	-	-	2.60–2.71	0.957	1.56		
1.0–2.0	0.22–0.40	0.957	2.05	5.0–7.0	0.22–0.40	1.000	0.43
	0.40–0.64	0.936	0.30		0.40–0.64	0.987	0.09
	0.64–0.82	0.931	0.26		0.64–0.82	0.980	0.13
	0.82–1.16	0.951	0.14		0.82–1.16	0.987	0.07
	1.16–1.46	0.964	0.14		1.16–1.46	0.985	0.08
	1.46–1.78	0.967	0.13		1.46–1.78	0.987	0.13
	1.78–2.13	0.958	0.19		1.78–2.13	0.957	4.36
	2.13–2.22	0.898	0.61		2.13–2.22	-	-
	2.22–2.60	0.900	0.40		2.22–2.60	0.875	60.94
2.60–2.71	0.576	11.69	2.60–2.71	-	-		

**Table 4.2:** The efficiency ratio for non-isolated  $\mu^+$  in bins of momentum and polar angle.

$p$ [GeV]	$\theta$ [rad]	$\epsilon_d/\epsilon_{mc}$	$\hat{\sigma}$ [%]	$p$ [GeV]	$\theta$ [rad]	$\epsilon_d/\epsilon_{mc}$	$\hat{\sigma}$ [%]
0.2–0.7	0.22–0.40	0.926	42.88	2.0–3.0	0.22–0.40	1.041	10.81
	0.40–0.64	1.619	26.18		0.40–0.64	1.004	5.15
	0.64–0.82	2.189	17.79		0.64–0.82	1.029	15.94
	0.82–1.16	0.990	15.55		0.82–1.16	0.868	12.75
	1.16–1.46	0.616	17.16		1.16–1.46	0.880	9.12
	1.46–1.78	0.702	14.34		1.46–1.78	0.905	7.93
	1.78–2.13	0.618	25.12		1.78–2.13	1.084	9.55
	2.13–2.22	-	-		2.13–2.22	-	-
	2.22–2.60	0.744	47.01		2.22–2.60	0.894	9.22
2.60–2.71	-	-	2.60–2.71	-	-		
0.7–1.0	0.22–0.40	0.443	53.79	3.0–5.0	0.22–0.40	0.919	17.02
	0.40–0.64	0.444	67.63		0.40–0.64	0.994	7.56
	0.64–0.82	1.000	65.47		0.64–0.82	0.857	21.68
	0.82–1.16	0.550	54.65		0.82–1.16	0.844	27.31
	1.16–1.46	1.019	22.79		1.16–1.46	0.873	21.40
	1.46–1.78	1.000	26.73		1.46–1.78	0.927	8.08
	1.78–2.13	1.200	40.82		1.78–2.13	0.918	12.13
	2.13–2.22	-	-		2.13–2.22	0.444	82.92
	2.22–2.60	0.333	64.62		2.22–2.60	1.005	4.80
2.60–2.71	-	-	2.60–2.71	-	-		
1.0–2.0	0.22–0.40	0.635	35.16	5.0–7.0	0.22–0.40	1.110	5.58
	0.40–0.64	0.612	14.94		0.40–0.64	0.987	3.36
	0.64–0.82	0.879	11.33		0.64–0.82	1.005	10.42
	0.82–1.16	0.661	22.17		0.82–1.16	0.975	4.29
	1.16–1.46	0.826	19.27		1.16–1.46	0.800	14.82
	1.46–1.78	0.762	15.57		1.46–1.78	0.944	12.83
	1.78–2.13	0.996	11.28		1.78–2.13	-	-
	2.13–2.22	-	-		2.13–2.22	-	-
	2.22–2.60	0.939	12.45		2.22–2.60	-	-
2.60–2.71	-	-	2.60–2.71	-	-		

**Table 4.3:** The efficiency ratio for isolated  $\mu^-$  in bins of momentum and polar angle.

$p$ [GeV]	$\theta$ [rad]	$\epsilon_d/\epsilon_{mc}$	$\hat{\sigma}$ [%]	$p$ [GeV]	$\theta$ [rad]	$\epsilon_d/\epsilon_{mc}$	$\hat{\sigma}$ [%]
0.2–0.7	0.22–0.40	0.783	5.64	2.0–3.0	0.22–0.40	0.900	1.71
	0.40–0.64	0.999	1.29		0.40–0.64	0.980	0.23
	0.64–0.82	0.884	2.33		0.64–0.82	0.954	0.21
	0.82–1.16	0.829	1.40		0.82–1.16	0.969	0.11
	1.16–1.46	0.859	0.97		1.16–1.46	0.977	0.09
	1.46–1.78	0.871	0.75		1.46–1.78	0.977	0.08
	1.78–2.13	0.837	1.03		1.78–2.13	0.972	0.10
	2.13–2.22	0.341	19.40		2.13–2.22	0.907	0.26
	2.22–2.60	0.655	3.09		2.22–2.60	0.960	0.11
2.60–2.71	-	-	2.60–2.71	0.913	1.00		
0.7–1.0	0.22–0.40	0.849	3.67	3.0–5.0	0.22–0.40	0.909	2.14
	0.40–0.64	0.922	1.14		0.40–0.64	0.988	0.38
	0.64–0.82	0.767	1.68		0.64–0.82	0.963	0.38
	0.82–1.16	0.900	0.70		0.82–1.16	0.973	0.20
	1.16–1.46	0.917	0.52		1.16–1.46	0.981	0.13
	1.46–1.78	0.945	0.38		1.46–1.78	0.981	0.08
	1.78–2.13	0.924	0.51		1.78–2.13	0.980	0.11
	2.13–2.22	0.854	2.18		2.13–2.22	0.916	0.32
	2.22–2.60	0.805	1.21		2.22–2.60	0.968	0.14
2.60–2.71	-	-	2.60–2.71	0.953	1.26		
1.0–2.0	0.22–0.40	0.868	3.63	5.0–7.0	0.22–0.40	0.990	0.49
	0.40–0.64	0.940	0.41		0.40–0.64	0.987	0.09
	0.64–0.82	0.933	0.32		0.64–0.82	0.974	0.12
	0.82–1.16	0.951	0.16		0.82–1.16	0.979	0.08
	1.16–1.46	0.960	0.14		1.16–1.46	0.984	0.08
	1.46–1.78	0.970	0.11		1.46–1.78	0.984	0.14
	1.78–2.13	0.958	0.15		1.78–2.13	0.936	4.27
	2.13–2.22	0.911	0.41		2.13–2.22	-	-
	2.22–2.60	0.911	0.25		2.22–2.60	0.643	54.43
2.60–2.71	0.650	5.55	2.60–2.71	-	-		

**Table 4.4:** The efficiency ratio for non-isolated  $\mu^-$  in bins of momentum and polar angle.

$p$ [GeV]	$\theta$ [rad]	$\epsilon_d/\epsilon_{mc}$	$\hat{\sigma}$ [%]	$p$ [GeV]	$\theta$ [rad]	$\epsilon_d/\epsilon_{mc}$	$\hat{\sigma}$ [%]
0.2–0.7	0.22–0.40	0.500	85.63	2.0–3.0	0.22–0.40	1.094	17.96
	0.40–0.64	1.786	32.03		0.40–0.64	0.823	16.93
	0.64–0.82	1.870	18.35		0.64–0.82	0.877	17.73
	0.82–1.16	0.662	19.92		0.82–1.16	0.882	11.64
	1.16–1.46	0.770	14.33		1.16–1.46	0.988	5.45
	1.46–1.78	0.860	11.35		1.46–1.78	1.003	4.58
	1.78–2.13	0.836	16.44		1.78–2.13	0.970	5.42
	2.13–2.22	1.083	126.11		2.13–2.22	1.120	34.07
	2.22–2.60	0.693	32.43		2.22–2.60	0.923	4.57
2.60–2.71	-	-	2.60–2.71	-	-		
0.7–1.0	0.22–0.40	0.343	100.26	3.0–5.0	0.22–0.40	0.764	31.41
	0.40–0.64	0.655	54.65		0.40–0.64	0.942	9.81
	0.64–0.82	0.600	81.65		0.64–0.82	0.865	28.02
	0.82–1.16	0.444	47.67		0.82–1.16	-	-
	1.16–1.46	0.771	39.02		1.16–1.46	0.850	21.39
	1.46–1.78	1.146	18.91		1.46–1.78	0.970	6.63
	1.78–2.13	-	-		1.78–2.13	0.869	14.07
	2.13–2.22	-	-		2.13–2.22	0.844	28.91
	2.22–2.60	0.969	24.73		2.22–2.60	1.115	4.09
2.60–2.71	-	-	2.60–2.71	0.889	51.54		
1.0–2.0	0.22–0.40	0.545	61.81	5.0–7.0	0.22–0.40	1.144	6.39
	0.40–0.64	0.827	16.09		0.40–0.64	0.980	4.72
	0.64–0.82	0.772	20.10		0.64–0.82	0.889	17.45
	0.82–1.16	0.697	35.68		0.82–1.16	0.930	7.79
	1.16–1.46	0.984	12.70		1.16–1.46	0.987	5.71
	1.46–1.78	0.832	11.91		1.46–1.78	0.829	20.74
	1.78–2.13	0.732	20.28		1.78–2.13	-	-
	2.13–2.22	1.125	43.03		2.13–2.22	-	-
	2.22–2.60	0.957	9.08		2.22–2.60	-	-
2.60–2.71	-	-	2.60–2.71	-	-		

## Chapter 5

# Conclusion

I investigated our ability to measure the Weinberg angle in  $e^+e^- \rightarrow \mu^+\mu^-$  events at Belle II. The tree-level squared amplitude, calculated for  $e^+e^- \rightarrow \mu^+\mu^-$ , as a function of the scattering angle is used in a model to generate data. Fitting the model to the generated data allows us to extract the Weinberg angle and cross-check the fitting procedure. A systematic study of any biases needs to be carried out, once a higher order model has been developed.

By examining different sources of background processes and selection criteria, I obtained a clean signal sample using the official Belle II Monte Carlo and reconstruction software. I identified the energies of the muons to be the most important variables to differentiate between signal and background events. While most background processes are well rejected, the model and selection criteria need to be studied further to determine to what degree  $e^+e^- \rightarrow \mu^+\mu^- \gamma$  events need to be rejected and to what degree they need to be accounted for by a higher order model.

The efficiency as a function of the polar angle is necessary for fitting the Weinberg angle to recorded data. I used simulated and recorded  $e^+e^- \rightarrow \mu^+\mu^- \gamma$  events, where the photon has an energy above 1 GeV, to study the muon identification efficiency as functions of the polar angle, the momentum and the track isolation. From this, the ratio of efficiencies in recorded to simulated events is calculated, which allows the correction of the simulated efficiency. The relative uncertainty on the efficiency ratio will be used by the Belle II collaboration to estimate systematic uncertainties.

# Bibliography

- [1] S. L. Glashow. “The renormalizability of vector meson interactions”. In: *Nuclear Physics* 10 (1959), pp. 107–117.
- [2] S. Weinberg. “A Model of Leptons”. In: *Phys. Rev. Lett.* 19 (21 Nov. 1967), pp. 1264–1266. DOI: 10.1103/PhysRevLett.19.1264.
- [3] A. Salam and J.C. Ward. “Electromagnetic and weak interactions”. In: *Physics Letters* 13.2 (1964), pp. 168–171. ISSN: 0031-9163. DOI: [https://doi.org/10.1016/0031-9163\(64\)90711-5](https://doi.org/10.1016/0031-9163(64)90711-5).
- [4] M. E. Peskin, D. V. Schroeder, and N. Balakrishnan. *Introduction to Quantum Field Theory*. 1995.
- [5] P. W. Higgs. “Broken Symmetries and the Masses of Gauge Bosons”. In: *Phys. Rev. Lett.* 13 (16 Oct. 1964), pp. 508–509. DOI: 10.1103/PhysRevLett.13.508.
- [6] F. Englert and R. Brout. “Broken Symmetry and the Mass of Gauge Vector Mesons”. In: *Phys. Rev. Lett.* 13 (9 Aug. 1964), pp. 321–323. DOI: 10.1103/PhysRevLett.13.321.
- [7] J. Erler and R. Ferro-Hernández. “Weak mixing angle in the Thomson limit”. In: *Journal of High Energy Physics* 2018.3 (Mar. 2018). DOI: 10.1007/jhep03(2018)196.
- [8] J. Erler and M. Schott. “Electroweak precision tests of the Standard Model after the discovery of the Higgs boson”. In: *Progress in Particle and Nuclear Physics* 106 (May 2019), pp. 68–119. DOI: 10.1016/j.pnpnp.2019.02.007.
- [9] R. L. Workman et al. “Review of Particle Physics”. In: *PTEP* 2022 (2022), p. 083C01. DOI: 10.1093/ptep/ptac097.
- [10] T. Aaltonen et al. “Tevatron Run II combination of the effective leptonic electroweak mixing angle”. In: *Physical Review D* 97.11 (June 2018). DOI: 10.1103/physrevd.97.112007.
- [11] R. Aaij et al. “Measurement of the forward-backward asymmetry in  $Z/\gamma \rightarrow \mu^+ \mu^-$  decays and determination of the effective weak mixing angle”. In: *Journal of High Energy Physics* 2015.11 (Nov. 2015). DOI: 10.1007/jhep11(2015)190.
- [12] G. Aad et al. “Measurement of the forward-backward asymmetry of electron and muon pair-production in pp collisions at  $\sqrt{s} = 7$  TeV with the ATLAS detector”. In: *Journal of High Energy Physics* 2015.9 (Sept. 2015). DOI: 10.1007/jhep09(2015)049.

- [13] S. Chatrchyan et al. “Measurement of the weak mixing angle with the Drell-Yan process in proton-proton collisions at the LHC”. In: *Physical Review D* 84.11 (Dec. 2011). DOI: 10.1103/physrevd.84.112002.
- [14] “Precision electroweak measurements on the Z resonance”. In: *Physics Reports* 427.5 (2006), pp. 257–454. ISSN: 0370-1573. DOI: <https://doi.org/10.1016/j.physrep.2005.12.006>.
- [15] K. Abe et al. “Improved Direct Measurement of Leptonic Coupling Asymmetries with Polarized Z Bosons”. In: *Physical Review Letters* 86.7 (Feb. 2001), pp. 1162–1166. DOI: 10.1103/physrevlett.86.1162.
- [16] C. S. Wood et al. “Measurement of Parity Nonconservation and an Anapole Moment in Cesium”. In: *Science* 275.5307 (1997), pp. 1759–1763. DOI: 10.1126/science.275.5307.1759.
- [17] P. L. Anthony et al. “Precision Measurement of the Weak Mixing Angle in Møller Scattering”. In: *Phys. Rev. Lett.* 95 (8 Aug. 2005), p. 081601. DOI: 10.1103/PhysRevLett.95.081601.
- [18] D. Androic et al. “First Determination of the Weak Charge of the Proton”. In: *Physical Review Letters* 111.14 (Oct. 2013). DOI: 10.1103/physrevlett.111.141803.
- [19] C.Y. Prescott et al. “Further measurements of parity non-conservation in inelastic electron scattering”. In: *Physics Letters B* 84.4 (1979), pp. 524–528. ISSN: 0370-2693. DOI: [https://doi.org/10.1016/0370-2693\(79\)91253-X](https://doi.org/10.1016/0370-2693(79)91253-X).
- [20] G. P. Zeller et al. “Precise Determination of Electroweak Parameters in Neutrino-Nucleon Scattering”. In: *Phys. Rev. Lett.* 88 (9 Feb. 2002), p. 091802. DOI: 10.1103/PhysRevLett.88.091802.
- [21] A. W. Thomas. “The Determination of  $\sin^2 \theta_W$  in Neutrino Scattering: no more anomaly”. In: *AIP Conference Proceedings*. American Institute of Physics, 2011. DOI: 10.1063/1.3667317.
- [22] J. Erler and M. J. Ramsey-Musolf. “Weak mixing angle at low energies”. In: *Phys. Rev. D* 72 (7 Oct. 2005), p. 073003. DOI: 10.1103/PhysRevD.72.073003.
- [23] T. Abe et al. *Belle II Technical Design Report*. 2010. DOI: 10.48550/ARXIV.1011.0352.
- [24] M. Böhm et al. “Forward-backward asymmetries”. In: Jan. 1989.
- [25] D. Foreman-Mackey et al. “emcee: The MCMC Hammer”. In: *PASP* 125 (2013), pp. 306–312. DOI: 10.1086/670067. eprint: 1202.3665.
- [26] The Belle II collaboration. *Belle II Luminosity*. Accessed October 20, 2022. URL: <https://confluence.desy.de/display/BI/Belle+II+Luminosity#BelleIILuminosity-TotalIntegratedluminosityforgoodruns>.
- [27] T. Kuhr et al. “The Belle II Core Software”. In: *Computing and Software for Big Science* 3.1 (Nov. 2018). DOI: 10.1007/s41781-018-0017-9.
- [28] The Belle II Collaboration. *Belle II Analysis Software Framework (basf2)*. Version release-06-00-03. Aug. 2022. DOI: 10.5281/zenodo.6949414.
- [29] E. Kou et al. “The Belle II Physics Book”. In: *Progress of Theoretical and Experimental Physics* 2019.12 (Dec. 2019). DOI: 10.1093/ptep/ptz106.

- [30] Lepton ID group. “Muon and electron identification efficiencies and hadron-lepton mis-identification probabilities with BelleII Phase III data”. In: (2022). BELLE2-NOTE-TE-2021-011 (Only available to BelleII members).
- [31] T. Ferber. *Precision Measurement of the Muon Pair Asymmetry in  $e^+e^-$  Annihilation at Belle*. Presentation in the Tau/2Photon Working Group. (Only available to Belle members). Oct. 2014. URL: [https://kds.kek.jp/event/16739/contributions/32657/attachments/26630/31706/tau2photon\\_torbenferber\\_v2.pdf](https://kds.kek.jp/event/16739/contributions/32657/attachments/26630/31706/tau2photon_torbenferber_v2.pdf).
- [32] S. Jadach, B.F.L. Ward, and Z. Was. “The precision Monte Carlo event generator for two-fermion final states in collisions”. In: *Computer Physics Communications* 130.3 (Aug. 2000), pp. 260–325. DOI: 10.1016/S0010-4655(00)00048-5.
- [33] W. Feller. *An Introduction to Probability Theory and Its Applications (Second Edition)*. Vol. 2. 1971.
- [34] N. L. Johnson, S. Kotz, and N. Balakrishnan. *Continuous Univariate Distributions (Second Edition)*. Vol. 2. 1995.
- [35] P. Urquijo and T. Ferber. “Overview of the BelleII Physics Generators”. In: (2016). BELLE2-NOTE-PH-2015-006 (Only available to Belle II members).
- [36] P. Virtanen et al. “SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python”. In: *Nature Methods* 17 (2020), pp. 261–272. DOI: 10.1038/s41592-019-0686-2.
- [37] *NIST Digital Library of Mathematical Functions*. <http://dlmf.nist.gov/>, Release 1.1.6 of 2022-06-30. F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller, B. V. Saunders, H. S. Cohl, and M. A. McClain, eds.
- [38] P. Feichtinger. “Particle identification performance and track isolation effects on LID using pions from taupair events”. In: (2021). BELLE2-NOTE-PH-2021-020 (Only available to Belle II members).
- [39] I. Tsaklidis and F. Bernlochner. “Improving kaon-pion separation with Neural Networks”. In: (2021). BELLE2-NOTE-TE-2021-028 (Only available to Belle II members).
- [40] M. Milesi. *Updates to track-based distance calculation*. Presentation in the Charged PID meeting (Only available to Belle II members). July 2022.

# Acknowledgements

I want to thank Dr. Daniel Greenwald and Prof. Dr. Stephan Paul for supporting me throughout my work. I am especially grateful to Daniel for guiding me throughout all aspects of this thesis. In addition, I want to thank Katarina Dugic, Lukas Bierwirth and Andrei Rabusov for their interesting discussions and support with all the challenges that I encountered. Finally, I want to thank Abtin Narimani for helping me with the study of the influence of the track isolation.