

Measuring the Branching Fraction of $B \rightarrow \rho \nu_\ell$ Decays with the Belle II Experiment

Zur Erlangung des akademischen Grades eines

DOKTORS DER NATURWISSENSCHAFTEN

(Dr. rer. nat.)

von der KIT-Fakultät für Physik des
Karlsruher Instituts für Technologie (KIT)

genehmigte

DISSERTATION


von

M.Sc. Moritz David Bauer

Tag der mündlichen Prüfung: 27. Oktober 2023

Referent: Prof. Dr. Torben Ferber

Korreferent: Prof. Dr. Günter Quast

 This document is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-SA 4.0): <https://creativecommons.org/licenses/by-sa/4.0/deed.en>

Disclaimer

Data analyses in high-energy physics such as the measurement presented in this doctoral thesis are a collaborative effort. The SuperKEKB particle accelerator which provides the particle beams essential for all studies at Belle II was built and is operated and maintained by the SuperKEKB accelerator group. The Belle II detector was built and is maintained and operated by the Belle II collaboration. The Belle II collaboration also creates the simulated and recorded data sets and maintains the computing infrastructure necessary to process them. The software environment necessary for studies with Belle II data plays an important role and was created and is maintained by the collaboration. The author of this work has been a part of the Belle II collaboration since 2019 and performed all studies detailed in this thesis except for the following:

- The branching fractions and uncertainty estimates of charmed semileptonic decays in the simulated samples as described in Chapter 5
- The determination of calibration factors and the associated systematic uncertainties for:
 - Particle identification and π^0 reconstruction as described in Sections 6.2.1 and 8.2
 - Tagging efficiency described in Sections 6.2.2 and 8.2
 - The track finding efficiency described in Section 8.2
 - The number of B-meson pairs in the data set and the associated uncertainties described in Chapter 9 and Section 8.2
- The ρ meson line shape correction described in Section 5.1
- The systematic uncertainty from $\rho - \omega$ interference described in Section 8.2
- The choice of input parameters to the BLNP model and the associated systematic uncertainty as detailed in Section 8.2
- The choice of parameters used in the bremsstrahlung correction described in Section 6.1.2

The XRootD caching project described in Section 3.5 has been conducted together with Dr. Matthias Schnepf, Dr. Manuel Giffels, Dr. Max Fischer and Prof. Dr. Günter Quast and has been submitted to the proceedings of the ACAT2022 conference.

Contents

1	Introduction	1
2	Foundations	3
2.1	The CKM Matrix	3
2.2	Semileptonic B Decays	5
2.2.1	Inclusive Charmless Semileptonic Decays	5
2.2.2	Exclusive Charmless Semileptonic Decays	6
2.3	Current Experimental Status	7
3	The Belle II Experiment	9
3.1	B Mesons and Υ Resonances	9
3.2	The Belle II Experiment	11
3.2.1	The SuperKEKB Accelerator	11
3.3	The Belle II Detector	12
3.4	Data Sets	16
3.5	The Belle II Computing Grid and XRootD Caching	17
4	Tools	22
4.1	Boosted Decision Trees	22
4.2	The Belle II Analysis Software Framework	23
4.3	Particle Identification	23
4.4	Vertex Fitting	24
4.5	Full Event Interpretation	25
5	Simulated Samples	28
5.1	$B \rightarrow \rho \ell \nu_\ell$ Signal Samples	28
5.2	$B \rightarrow X_u \ell \nu_\ell$ Backgrounds	30
5.2.1	Exclusive $B \rightarrow X_u \ell \nu_\ell$ backgrounds	30
5.2.2	Inclusive $B \rightarrow X_u \ell \nu_\ell$ Backgrounds	31
5.2.3	Hybrid Model	32
5.3	Other Simulated Backgrounds	35

6	Reconstruction and Selection	38
6.1	Event Selections	38
6.1.1	Tag-Side Selection	38
6.1.2	Signal-Side Selection	40
6.1.3	Selections on the $\Upsilon(4S)$ Candidate	44
6.1.4	Light-Quark Pair Background Suppression	44
6.2	Monte Carlo Corrections for Selection Effects	50
6.2.1	Correction of Particle Reconstruction Efficiencies and Misidentification Rates	50
6.2.2	Tagging Efficiency Correction	51
6.3	Resulting Simulated and Recorded Samples	55
6.4	The $B \rightarrow D(\rightarrow \pi^\pm \pi^\mp) \ell \nu_\ell$ Control Channel	57
7	Signal Extraction Strategy	58
7.1	The Method of Maximum Likelihood	58
7.1.1	Parameter Estimation with Templates	59
7.1.2	Uncertainties Originating from Template Distributions	60
7.1.3	Uncertainty Estimation using the Likelihood	61
7.2	Implementation of the Fit	63
7.2.1	Fit Model in the $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ Analysis	63
7.2.2	Fit Model in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ Analysis	65
7.3	Fit Validation with Simulated Data	70
7.3.1	Validation using Simplified Statistical Models	70
7.3.2	Linearity Tests in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ Reconstruction Channel	74
7.3.3	Model Dependence of the Fit in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ Channel	74
8	Systematic Uncertainties	77
8.1	Correlation Matrices Estimation	77
8.2	Sources of Systematic Uncertainty	78
8.2.1	Uncertainties from Theoretical Models and Simulation	78
8.2.2	Uncertainties from Reconstruction Effects	86
8.2.3	Uncertainties from the Signal Extraction Procedure	86
9	Results	88
9.1	Fit to Data	88
9.2	Branching Fraction Determination	94
10	Conclusion	97
	Bibliography	101

Appendix A Two-dimensional Projections of the Fit Templates with Bin-wise Uncertainties	112
A.0.1 $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ Templates	112
A.0.2 $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ Templates	117
Appendix B Bin-by-bin Fit Result Figures and Comparison to Simulation	126
B.0.1 $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ Results	126
B.0.2 $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ Results	131
Appendix C Post-fit Nuisance Parameters for All Templates in the Fit	136
C.0.1 $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ Nuisance Parameters	136
C.0.2 $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ Nuisance Parameters	138
Appendix D Correlation Matrices between the FEI Calibration Factors	141
Appendix E Hybrid MC for Neutral $B \rightarrow X_u \ell \nu_\ell$ Decays	143
Glossary	145
Acronyms	147
List of Figures	150
List of Tables	158

Chapter 1

Introduction

High-energy physics seeks to determine the laws of fundamental interactions in nature and the constituents of matter. The widely-accepted theory which can describe three of the four fundamental interactions is commonly referred to as the Standard Model of particle physics (SM). The SM is enormously successful and, over the last 50 years, has shown excellent agreement with even high-precision measurements. However, even disregarding the lack of an explanation for gravity in the SM, it is far from complete. Unsettled questions range from explanations for astronomical observations such as dark matter or baryon asymmetry to the theories underlying the parameters used as inputs to the SM. To search for phenomena beyond our current knowledge and extend the SM, two major strategies have established themselves: Direct searches for new phenomena and precision measurements of the properties of the SM.

With only 19 dimensionless parameters and equivalently 19 degrees of freedom, the SM provides ample opportunity for over-determination. To this end, flavor physics, the study of elementary fermion transitions, has established itself as a powerful tool. Due to their large masses, B mesons and in particular their decays offer access to a large phenomenological landscape. Experiments primarily dedicated to B meson studies, namely the asymmetric B factories Belle and BaBar as well as the Large Hadron Collider beauty (LHCb) experiment, have successfully mapped this landscape over the last 25 years. Further advances require, in addition to better theoretical predictions and improved analysis techniques, vastly larger data sets and higher experimental resolution. This advancement is the goal of the Belle II experiment which aims to collect a data set approximately 50 times larger [1] than the one collected by the Belle experiment.

The magnitudes of the Cabibbo–Kobayashi–Maskawa (CKM) matrix elements* are parameters of the SM which can only be measured. Transitions from a b quark to a u quark are governed by the CKM matrix element $|V_{ub}|$. Measurements of $|V_{ub}|$ assign the largest relative uncertainty of all CKM elements, making increased precision in its determination essential. A key factor of current determinations of $|V_{ub}|$ is given by the treatment of hadronic processes in semileptonic decays with transitions from an up quark to a bottom quark. While $B \rightarrow \pi \ell \nu_\ell$

*The CKM matrix will be introduced in Chapter 2.

processes are the most explored among the transitions to a specific hadronic final state, additional insight can be gained by transitions involving ρ mesons. Previous determinations of their branching fractions disagree significantly with each other so further studies of charged and neutral $B \rightarrow \rho l \nu_\ell$ processes are warranted.

This work presents the first branching fraction measurements of $B \rightarrow \rho l \nu_\ell$ decays at the Belle II experiment. A new hadronic tagging method is employed in this determination and new effects are considered for their impact on the result and the systematic uncertainties.

In the next chapters of this work, I describe the foundations of the $B \rightarrow \rho l \nu_\ell$ decay and the current experimental status (Chapter 2), summarize the experimental setup at Belle II as well as a research project to improve its computing infrastructure (Chapter 3) and explain the tools essential for the presented analysis (Chapter 4). The succeeding chapters are dedicated to descriptions of the $B \rightarrow \rho l \nu_\ell$ analysis: Descriptions of the simulated processes forming the backbone of the analysis as well as the reconstruction and selection procedures can be found in Chapter 5 and Chapter 6, respectively. Chapter 7 details the signal extraction approach and the statistical foundations underpinning it. The sources of systematic uncertainty considered in this analysis, which form a main focus of this measurement, are given in Chapter 8. Finally, the results determined for the branching fraction of the $B \rightarrow \rho l \nu_\ell$ process are discussed in Chapter 9.

Chapter 2

Foundations

In this chapter, I give an overview of the theoretical and experimental foundations of this analysis. I briefly describe the mechanisms of weak decays which form the basis of flavor physics and outline inclusive and exclusive approaches to determine the rate of semileptonic B decays. At the end of the chapter, I briefly summarize the current experimental status of measurements of the $B \rightarrow \rho \ell \nu_\ell$ branching fraction and the extraction of $|V_{ub}|$ from it.

2.1 The CKM Matrix

With the discovery of the top quark in 1995, the flavor sector of the standard model as it is currently understood was completed. It contains six quarks grouped into three generations,

$$\begin{pmatrix} \text{up} \\ \text{down} \end{pmatrix} \quad \begin{pmatrix} \text{charm} \\ \text{strange} \end{pmatrix} \quad \begin{pmatrix} \text{top} \\ \text{bottom} \end{pmatrix},$$

with masses ranging from approximately $2 \text{ MeV } c^{-2}$ for the up quark to approximately $170 \text{ GeV } c^{-2}$ for the top quark. This hierarchy, which spans five orders of magnitude, is not yet understood and is one of the central open questions in the SM.

Weak interactions are one of the four fundamental interactions known today and are mediated by the W and Z vector bosons. The foundation for flavor-changing interactions between the quarks is given by the Kobayashi-Maskawa mechanism [2] introduced in 1973 to explain CP violation which had been observed almost a decade earlier in neutral kaon decays. At this point, mixing between up, down and strange quarks had already been described with a quark mixing mechanism by Nicola Cabibbo [3]. By introducing a third generation of quarks and describing mixing between all quark families with the CKM matrix, Kobayashi and Maskawa were able to introduce a CP-violating phase in the quark fields which cannot be absorbed by a quark-field phase redefinition.

Due to the unitarity of the 3×3 CKM matrix, it can be parametrized by three mixing angles θ_{ij} and the CP-violating phase δ . A standard choice [4] is

$$V = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \quad (2.1)$$

$$= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{-i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{13}s_{23} - s_{12}c_{23}s_{13}e^{-i\delta} & c_{23}c_{13} \end{pmatrix} \quad (2.2)$$

with $c_{ij} = \cos \theta_{ij}$ and $s_{ij} = \sin \theta_{ij}$. Here, θ_{12} is equivalent to the Cabibbo angle already introduced in the two-generation model.

While the CKM mechanism itself is well understood, its parameters are inputs to the SM and must be measured in experiments. Independent measurements of the magnitudes of the CKM matrix elements summarized in Ref. [5] find

$$|V| = \begin{pmatrix} 0.97373 \pm 0.00031 & 0.2243 \pm 0.0008 & 0.00382 \pm 0.0002 \\ 0.221 \pm 0.004 & 0.975 \pm 0.006 & 0.0408 \pm 0.0014 \\ 0.0086 \pm 0.0002 & 0.0415 \pm 0.0009 & 1.014 \pm 0.029 \end{pmatrix}.$$

As these measurements show, the CKM matrix exhibits a strong hierarchy with elements on the diagonal being close to unity and off-diagonal elements being small. This makes direct measurements of the magnitudes of V_{td} and V_{ub} especially challenging as the corresponding decays are heavily suppressed.

For the element V_{td} , B - \bar{B} oscillations offer an indirect approach which can provide precise measurements. For $|V_{ub}|$, no such indirect approach is known which makes the direct determination in B decays currently the only feasible method to measure $|V_{ub}|$. Consequently, current measurements of $|V_{ub}|$ give the largest fractional uncertainty out of all CKM elements.

In principle, both leptonic and semileptonic B decays can be used to measure $|V_{ub}|$. Due to their much smaller branching fractions, however, measurements of leptonic decays have not yet reached the precision on $|V_{ub}|$ obtained with semileptonic decays. Measurements with semileptonic decays are conducted either using an inclusive approach which encompasses all hadronic states with b and u quark content or using an exclusive approach in which a specific hadronic final state is required. As illustrated in Fig. 2.1, over time, values of $|V_{ub}|$ measured with these two approaches have developed tension [5] which indicates the need for further study.

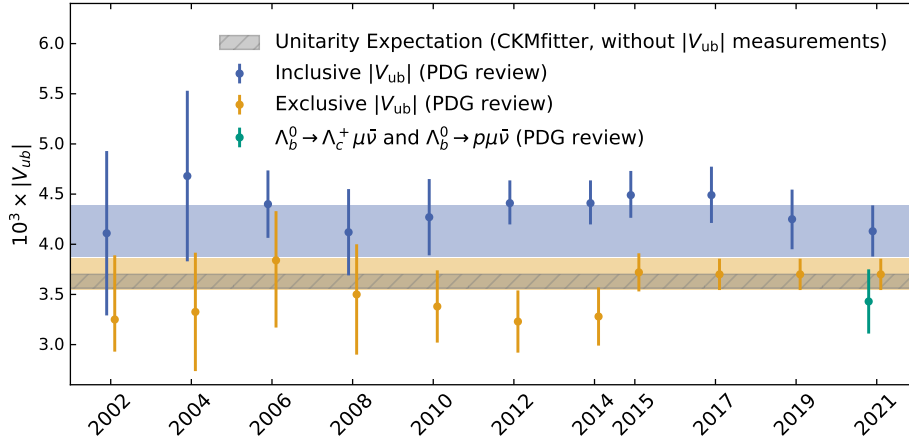


Figure 2.1: Inclusive and exclusive average of $|V_{ub}|$ over time as calculated by the Particle Data Group (PDG). Also given is the expectation from CKM unitarity determined in a fit performed by the CKMfitter group [6] without the input from $|V_{ub}|$ measurements.

2.2 Semileptonic B Decays

The effective Lagrangian for semileptonic $b \rightarrow u$ processes in the SM is given by

$$\mathcal{L}_{\text{eff}} = \frac{-4G_F}{\sqrt{2}} V_{ub} (\bar{u} \gamma_\mu P_L b) (\bar{\ell} \gamma^\mu P_L \nu) + h.c. \quad (2.3)$$

with $P_L = (1 - \gamma_5)/2$ and the Fermi constant G_F . Using the large mass difference between the b quark mass and the mass of the W boson m_W , the W boson is integrated out at leading order in \mathcal{L}_{eff} . Higher-order electroweak terms are suppressed by powers of G_F which is small with $G_F \approx 1.16637 \times 10^{-5} \text{ GeV}^{-2}$.

An important property in the description of semileptonic decays is that the electroweak part of the decay, which is calculated using perturbation theory, and the hadronic component, which must be treated with non-perturbative methods, effectively factorize. Further steps in calculating this hadronic component depend on whether the aim is to predict the inclusive decay rate to all hadronic final states or the exclusive decay rate to one specific hadronic final state as described above. In the following, both approaches are outlined following the more thorough descriptions in Ref. [7] and Ref. [5, Chapter 76].

2.2.1 Inclusive Charmless Semileptonic Decays

The total inclusive decay rate of $b \rightarrow u$ can be calculated in the framework of Heavy Quark Effective Theory (HQET) using the Heavy Quark Expansion (HQE) [8, 9, 10, 11] in analogy to $b \rightarrow c$ processes as described in the review in Ref. [5, Chapter 16] or in Ref. [12]. The total decay rate, however, is not directly accessible to measurements of $b \rightarrow u$ processes. These can typically only determine the partial rate in a region close to the maximum lepton energy where $b \rightarrow c$ processes are kinematically forbidden. The determination of a partial inclusive decay

rate requires the introduction of non-perturbative distribution functions referred to as “shape functions” for which the exact form is unknown. While different theoretical descriptions of the shape functions exist [13, 14, 15], the leading-order shape function can be obtained directly from measurements of any transition of a heavy to a light quark which allows its determination from $B \rightarrow X_s \gamma$ decays. To further constrain the moments of the shape functions in subleading order, a global fit of the HQE parameters m_b and μ_π^2 to $B \rightarrow X_c \ell \nu_\ell$ and $B \rightarrow X_s \gamma$ moments is performed in Ref. [16]. These two parameters are naturally identified with the mass of the b quark and the average kinetic energy in the B meson [17, Chapter 17].

2.2.2 Exclusive Charmless Semileptonic Decays

To describe the decay rates of exclusive $B \rightarrow X_u \ell \nu_\ell$ processes, the properties of the hadronic transition are encoded within form factors. In the following, I summarize the application of factors for $B \rightarrow \rho \ell \nu_\ell$ decays, more detailed descriptions are given in Refs. [18, 19]. A description of the more well-understood $B \rightarrow \pi \ell \nu_\ell$ process can be found in Ref. [5].

The hadronic matrix element for exclusive transitions to vector mesons V is given by

$$\langle V(p_V) | \bar{u} \gamma^\mu P_L b | \bar{B}(p_B) \rangle = c_V \sum_i \mathcal{T}_i^\mu F_i(q^2) \quad (2.4)$$

with the basis \mathcal{T}_i and the aforementioned form factors F_i which are dependent on the squared four-momentum transfer to the lepton system $q^2 = (p_B - p_V)^2$. c_V is the Clebsch-Gordan coefficient which for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays is $c_V = 1/\sqrt{2}$ and for $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ decays is $c_V = 1$.

As a weak interaction, in the traditional basis $\{A_P, V, A_0, A_1, A_{12}, T_1, T_2, T_{23}\}$ used e.g. in Refs. [18, 19], only the vector and axial vector components are expected to contribute to the $B \rightarrow \rho \ell \nu_\ell$ process.

It is convenient to define the helicity basis with the amplitudes

$$H_\pm(q^2) = \frac{2m_B |p_V| V(q^2)}{m_B + m_V} \pm (m_B + m_V) A_1(q^2), \quad (2.5)$$

$$H_0(q^2) = \frac{8m_B m_V A_{12}(q^2)}{\sqrt{q^2}}, \quad (2.6)$$

$$H_s(q^2) = \frac{2m_B |p_V| A_0(q^2)}{\sqrt{q^2}}. \quad (2.7)$$

In this basis, the decay rate for $B \rightarrow V \ell \nu_\ell$ processes is

$$\begin{aligned} \frac{d\Gamma}{dq^2} &= \frac{G_F^2 |V_{ub}|^2 c_V^2}{96\pi^3} |p_V| \frac{q^2}{m_B^2} \left(1 - \frac{m_\ell^2}{q^2}\right)^2 \\ &\times \left[\left(1 + \frac{m_\ell^2}{2q^2}\right) (H_+^2(q^2) + H_-^2(q^2) + H_0^2(q^2)) + \frac{3m_\ell^2}{2q^2} H_s^2(q^2) \right], \end{aligned} \quad (2.8)$$

where $|p_V| = \sqrt{\lambda(q^2)}/(2m_B)$ is the vector meson 3-momentum in the B meson rest frame with the Källén function $\lambda(q^2) = ((m_B + m_V)^2 - q^2)((m_B - m_V)^2 - q^2)$.

While for $B \rightarrow \pi \ell \nu_\ell$ decays, the F_i themselves can also be determined using lattice QCD, for $B \rightarrow \rho \ell \nu_\ell$ decays the only viable determination is via Light-Cone Sum Rules (LCSR). In this approach, correlation functions are determined in a perturbative manner in a kinematic region where perturbation theory is valid and then extended to the non-perturbative region with a dispersion relation [20]. LCSRs are most reliable in regions of q^2 below $14 \text{ GeV}^2/c^4$, this makes them complementary to lattice QCD which is most reliable in high q^2 regions.

2.3 Current Experimental Status

The $B \rightarrow \rho \ell \nu_\ell$ decay was first observed in 1996 by the CLEO collaboration in a measurement [21] in which the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ and $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ channels were combined assuming isospin symmetry. In total, Ref. [5] lists eight determinations of the branching fraction of $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ and eleven determinations of the branching fraction of $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$. Not all of these determinations are independent of each other, Fig. 2.2 shows the most recent and precise independent measurements from which the world average is determined. Figure 2.2 also shows that the most recent and precise determinations of $\mathcal{B}(B \rightarrow \rho \ell \nu_\ell)$ are in tension. The value measured by BaBar for $\mathcal{B}(B^0 \rightarrow \rho^- \ell^+ \nu_\ell)$ is 3σ below the value measured by Belle, making this tension significant. Similarly, a 2σ tension is seen between the measurements of $\mathcal{B}(B^+ \rightarrow \rho^0 \ell^+ \nu_\ell)$ by BaBar and Belle. As the two measurements from Belle and BaBar use different experimental methodology, additional determinations of both branching fractions should be able to determine if this discrepancy is a systematic effect and allow a more straightforward combination to a world average.

In addition to measuring the branching fraction, $B \rightarrow \rho \ell \nu_\ell$ can also be used to determine $|V_{ub}|$. The most precise determinations of $|V_{ub}|$ from exclusive decays use $B \rightarrow \pi \ell \nu_\ell$ decays, which offer both theoretical and experimental advantages. There are, however, also multiple determinations of $|V_{ub}|$ from $B \rightarrow \rho \ell \nu_\ell$ decays which offer an additional avenue to test for consistency. The determination in Ref. [19], however, consistently finds values of $|V_{ub}|$ below those determined from $B \rightarrow \pi \ell \nu_\ell$ which warrants further investigation.

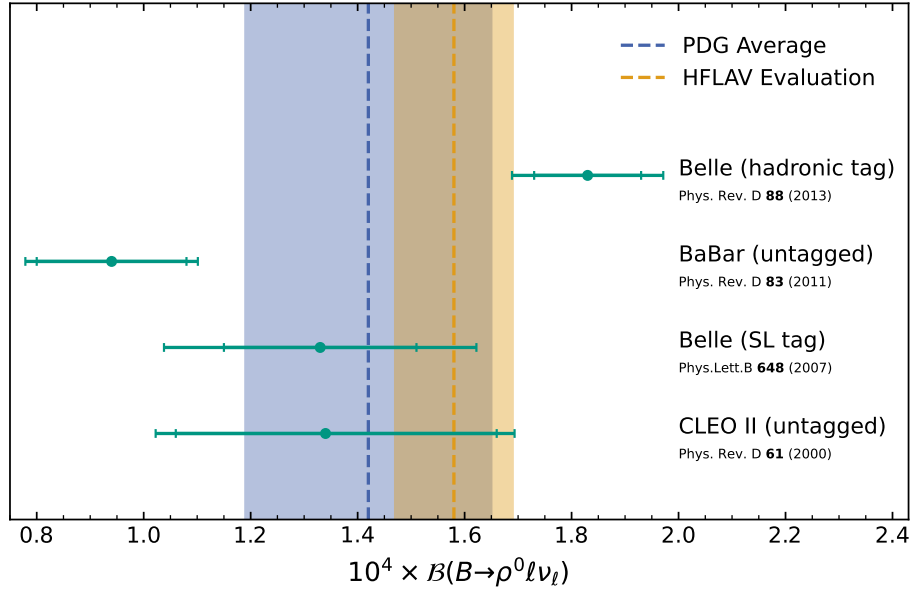
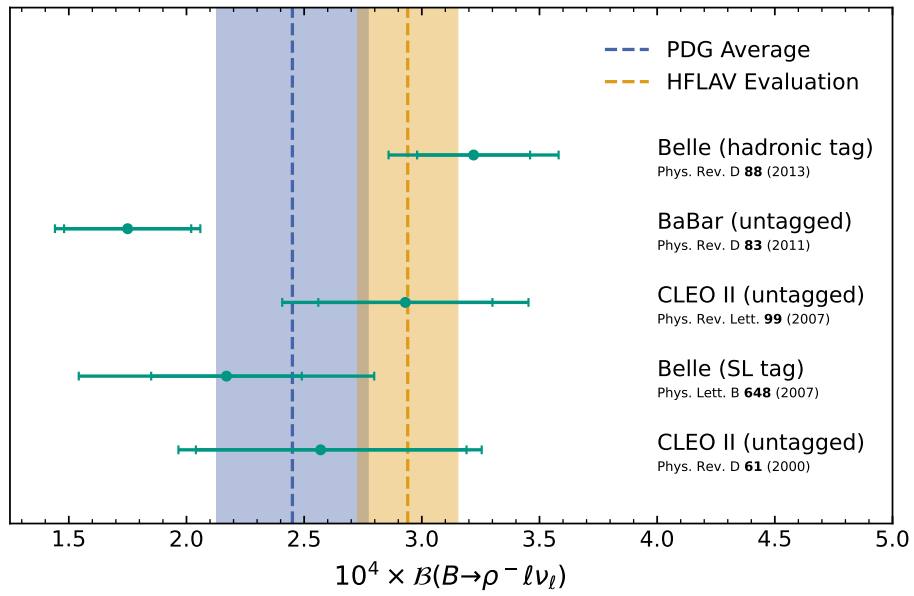
(a) $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ (b) $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$

Figure 2.2: Comparison of statistically independent, previous determinations of $\mathcal{B}(B \rightarrow \rho \ell \nu_\ell)$ together with the PDG average and the Heavy Flavor Averaging Group (HFLAV) evaluation. The PDG average is independent between the charged and neutral decay mode while the HFLAV evaluation uses isospin relations between the modes.

Chapter 3

The Belle II Experiment

In this chapter I present the experimental setup at the B Factories and give an overview of the fundamental concepts relevant for this analysis. I briefly describe the SuperKEKB accelerator and the Belle II detector, for a detailed description see Ref. [22].

3.1 B Mesons and Υ Resonances

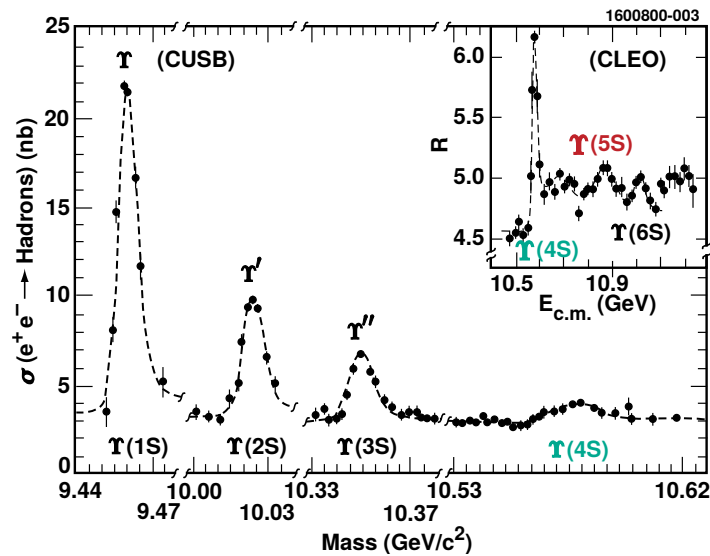


Figure 3.1: The bottomonium resonances $\Upsilon(1S)$, $\Upsilon(2S)$, $\Upsilon(3S)$ and $\Upsilon(4S)$ in the total hadronic cross section, given in nb as measured at the Cornell Electron Storage Ring (CESR) by the Columbia University-Stony Brook (CUSB) detector [23]. The inset figure shows additional results for $\Upsilon(4S)$, $\Upsilon(5S)$ and $\Upsilon(6S)$ obtained by the CLEO Collaboration [24]. These are given as the ratio $R = \sigma_{\text{had}}/\sigma_{\mu\mu}$. This plot was initially published in Ref. [25]. This figure is adapted from [26].

The bound states formed by two charm or bottom quarks of the same flavor can be described with non-relativistic methods due to the high quark mass and, in analogy to the

positronium state formed by electron and positron, are referred to as (heavy) quarkonia. The state formed by a $b\bar{b}$ quark pair is called “bottomonium.”

At electron-positron colliders operating at energies far from the Z boson mass, bottomonium states are created via virtual photons which have the quantum numbers $J^{PC} = 1^{--}$. This limits the $b\bar{b}$ states created in this process to ones with $^{2S+1}L_J = ^S3_1$ [25]. Six of these so-called $\Upsilon(nS)$ resonances, shown in Fig. 3.1, have been measured at energies between 9.46 to 10.99 GeV [5].

The resonances $\Upsilon(1S)$, $\Upsilon(2S)$, and $\Upsilon(3S)$ can only decay via $\bar{O}kubo-Zweig-Iizuka$ (OZI)-suppressed [27, 28, 29] strong force decays or via the weak or electromagnetic interaction. For Υ states above the “open-bottom threshold,” i.e. the energy needed to create a pair of B mesons, the OZI suppression is lifted. This makes the decay to two B mesons the dominating decay mode for $\Upsilon(4S)$ resonances with a branching fraction of more than 96% [5]. As the mass of the $\Upsilon(4S)$ resonance is only 20 MeV above the open-bottom threshold, the $B\bar{B}$ pair is produced almost at rest in the frame of the $\Upsilon(4S)$. Because B mesons provide many interesting physics cases, there are specialized experiments that operate at or near the Υ resonances. These experiments are referred to as “ B Factories.” Examples for this type of experiment design are the Belle, BaBar, and Belle II experiments. While the LHCb experiment [30] at the Large Hadron Collider [31] is also frequently considered a B factory, it is excluded from the following description due to the fundamental differences between lepton and hadron colliders.

The initial physics motivation for the construction of the first generation of B Factories, namely Belle and BaBar, was the investigation of CP violation as predicted by the CKM mechanism [2]. To precisely measure neutral B meson oscillations which are directly related to the amount of CP violation, one has to precisely measure the distance between the two decay vertices of an entangled pair of neutral B mesons. Since the difference between the decay lengths of two B mesons in their center-of-mass frame is only on the order of a few $10\ \mu\text{m}$, the B Factories have asymmetric beam energies to boost the bottomonium system with respect to the inertial reference frame of the detector. This increases the distances between the B decay vertices such that they can be resolved with current detector technologies.

The exceptional properties of the $\Upsilon(4S)$ resonance and the clean environment at Belle and BaBar made these experiments ideal to perform other precision measurements of theory predictions in the B meson sector with a focus on decays involving neutral and invisible particles. Building on these achievements as described in Ref. [32] for Belle, the Belle II experiment aims to increase both the precision and scope of the Belle physics program. In the flavor sector, a key element of these improvements is an increase of the data taking rate to reduce statistical uncertainties. The following sections describe the changes necessary to increase this rate and the data set obtained so far.

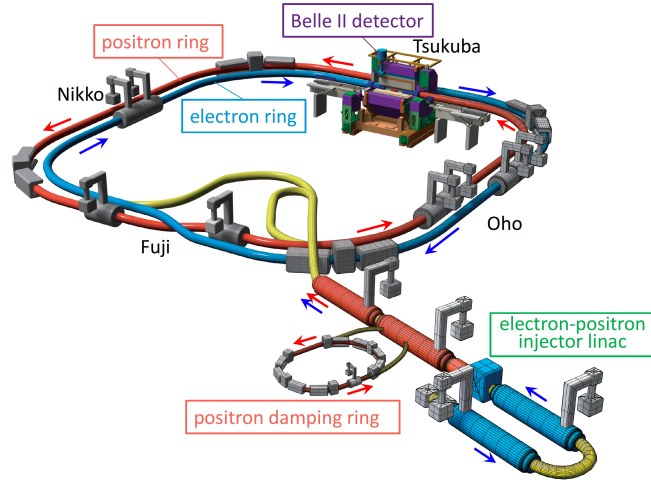


Figure 3.2: Illustration of the SuperKEKB accelerator complex at High Energy Accelerator Research Organization (KEK). The labels “Tsukuba,” “Nikko,” “Fuji,” and “Oho” refer to straight sections of the accelerator, the Belle II experiment is located in the Tsukuba section where the beams collide. Illustration taken from Ref. [33].

3.2 The Belle II Experiment

The Belle II experiment consists of the SuperKEKB accelerator and the Belle II detector. The former is described in detail in Ref. [33], the latter in Ref. [22]. Here, brief descriptions of both are given.

3.2.1 The SuperKEKB Accelerator

The SuperKEKB collider at the High Energy Accelerator Research Organisation (KEK) in Tsukuba, Japan (shown in Fig. 3.2) provides the particle beams used for the Belle II detector. It is an upgrade of the KEKB accelerator, which provided e^+e^- collisions to the Belle experiment, with a primary goal to considerably increase the instantaneous luminosity and thereby the rate at which Belle II can record data. SuperKEKB has been constructed in the tunnel first used for the TRISTAN collider and consists of two storage rings, one High-Energy Ring (HER) for 7 GeV electrons and one Low-Energy Ring (LER) for 4 GeV positrons. The two rings cross at two points, one of them, referred to as the Interaction Point (IP), is surrounded by the Belle II detector.

A single linear accelerator is used to fill the two SuperKEKB rings with electrons and positrons. While the electrons are injected directly from this linear accelerator into the HER, the positrons to fill the LER are created by striking a tungsten target with electrons.

The beams in the HER and the LER collide inside the Belle II detector, the energy for each beam collision or “event” is described by

$$E_{\text{cms}} = 2\sqrt{E_{\text{HER}}E_{\text{LER}}}. \quad (3.1)$$

Table 3.1: Key machine parameters of KEKB and SuperKEKB. The KEKB parameters are those achieved while the SuperKEKB parameters are the design parameters.

Ring	KEKB		SuperKEKB		Units
	LER	HER	LER	HER	
Beam Energy	3.5	8.0	4.0	7	GeV
β_x	1200	1200	32	25	mm
β_y	5.9	5.9	0.27	0.30	mm
σ_x	147	170	10.1	10.7	μm
σ_y	940	940	48	62	nm
Beam current	1.64	1.19	3.60	2.60	A

In contrast to hadron colliders such as the Large Hadron Collider (LHC), E_{cms} is also the exact energy of each interaction at electron-positron colliders which allows defining a complete center-of-mass system (cms) of the collision. This makes lepton colliders particularly useful for missing-energy studies such as semileptonic decays and searches for dark matter candidates.

The luminosity L at electron positron colliders can be expressed with the proportionality

$$L \propto \left(1 + \frac{\sigma_y}{\sigma_x}\right) \left(\frac{I_{\pm}}{\beta_y}\right) \quad (3.2)$$

where σ_x and σ_y are the beam dimensions in x and y direction at the IP, I_{\pm} is the current of the electron and positron beams, and β_y is the vertical beta function at the IP. The nanobeam collision scheme, originally proposed for the never-built SuperB accelerator [34], reduces β_y by a factor of 20 by colliding the beams with a large Piwinski angle as described in Ref. [35]. Due to this reduction in beam size and a planned increase of the beam current by a factor of two, the design luminosity of SuperKEKB is $8 \times 10^{35} \text{ cm}^{-2} \text{ s}^{-1}$, i.e. an increase by a factor of 40 compared to KEKB. Other design parameters of SuperKEKB compared to those achieved by KEKB are given in Table 3.1. As of 8th June, 2022, SuperKEKB holds the world record for peak instantaneous luminosity with $4.65 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ [36].

The desired increase of luminosity compared to KEKB also motivates the decrease of beam asymmetry to a Lorentz boost of $\beta\gamma = 0.28$. As described in Ref [34], with higher energy in the LER, lower emittance growth due to intrabeam scattering is expected. Similarly, the losses due to synchrotron radiation are expected to decrease with lower energy in the HER. The loss of vertex resolution due to the lower boost of the $\Upsilon(4S)$ system is expected to be compensated by improved vertex detectors and a 33% smaller beam pipe at the IP in Belle II.

3.3 The Belle II Detector

The Belle II detector, a general-purpose 4π detector, consists of a cylindrical structure around the IP. It can be subdivided into a part parallel to the beam axis, the ‘‘barrel,’’ and two disks perpendicular to the beam, the ‘‘endcaps’’. Belle II presents a major upgrade of the Belle

detector in which every subdetector received significant upgrades or was replaced entirely. This upgrade allows the detector to have better or at least equal performance than the Belle detector in the presence of strongly increased background levels due to much higher luminosity.

New inner vertex detectors, an upgraded main tracking detector, new particle identification detectors, new scintillators in the K_L^0 and muon system, and new electronics in the calorimeter enable better vertex and track reconstruction, better separation between charged pions and kaons, and higher resistance against misreconstructed neutral particles.

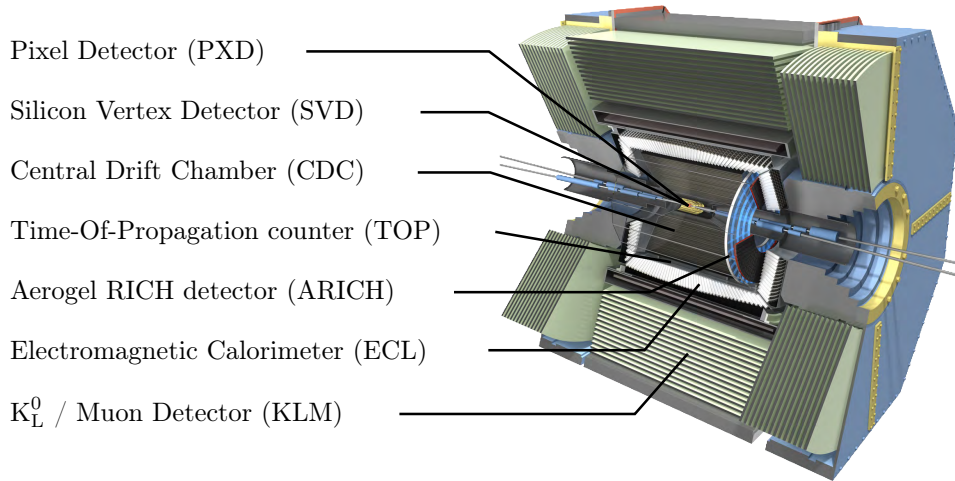


Figure 3.3: The Belle II detector with its seven subdetectors. Adapted from Ref. [37].

The Belle II detector, which is illustrated in Fig. 3.3, consists of seven subdetectors: Three tracking detectors, a calorimeter, a K_L^0 and muon detection system and two particle identification (PID) detectors utilizing Cherenkov radiation. The three innermost detectors of the experiment, the tracking detectors, are used to reconstruct the trajectories, or “tracks,” which charged particles take through the detector. Between the calorimeter and the K_L^0 and muon detection system, a superconducting solenoid is placed which provides the 1.5 T magnetic field in which the particle tracks are bent.

The Belle II coordinate system is defined by the z axis in the laboratory frame pointing along the central axis of the solenoid in the direction of the electron beam. As the detector is symmetric along the z axis, it is advantageous to use cylindrical coordinates to describe it. The polar angle θ , as well as the longitudinal and transverse directions, are defined with respect to the z axis.

The following paragraphs briefly summarizes the functions of each detector in the configuration used to record the data set for the presented analysis. A more thorough description of the initial design for Belle II can be found in Ref. [22], with additional information on the detector as installed given in Ref. [1].

Pixel Detector The two innermost active layers of the detector at radii of 14 mm and 22 mm are formed by the Pixel Detector (PXD) [38]. The PXD is a silicon detector which can be read

out in individual segments (“pixels”) with sizes between $50\ \mu\text{m} \times 55\ \mu\text{m}$ and $50\ \mu\text{m} \times 85\ \mu\text{m}$. This fine segmentation for distances closer than roughly 40 mm from the beam is necessary to reduce the occupancy of each pixel, i.e. the fraction of activated readout channels out of all PXD channels in each event at Belle II, to levels at which detector hits can still be assigned to particle tracks. Previously existing implementations of silicon pixel detectors at e.g. LHC experiments require a large material budget for each sensor and the attached readout electronics. While this is feasible at the high energies at the LHC, the multiple scattering effects at SuperKEKB’s energy introduce difficulty in the accurate determination of B decay vertices. For the PXD, Depleted Field Effect Transistor (DEPFET) sensors are used to avoid this issue. The DEPFET technology, which locates the electronics outside the active detector, allows sensors to be as thin as $75\ \mu\text{m}$. In the data taking period relevant for the presented analysis, only six out of the twelve modules of the outer PXD layer had been installed. During Belle II’s first long shutdown in 2022 and 2023, a new detector with the properties described above but two complete layers has been installed, fulfilling the initial design of the detector.

Silicon Vertex Detector The PXD is followed by a dual-side silicon strip detector, the Silicon Vertex Detector (SVD) [39]. It reaches from an inner radius of 38 mm to an outer radius of 140 mm and covers an angular region of $17^\circ < \theta < 150^\circ$. To reach this angular acceptance while keeping the number of sensors to a minimum, the sensors are slanted in the forward region of the detector. In addition to providing vertex information for B decays together with the PXD, the SVD is also the only subdetector able to vertex K_S^0 mesons decaying outside the PXD and to reconstruct tracks of particles with low transverse momenta which do not reach the surrounding tracking detector.

Central Drift Chamber The main tracking detector in Belle II, the Central Drift Chamber (CDC) [40], reaches from a radius of 16 cm to a radius of 113 cm. It covers the same polar angle as the SVD, $17^\circ < \theta < 150^\circ$. As multiple scattering dominates at charged particle momenta of 1 to $2\ \text{GeV}\ c^{-1}$ which are typical at Belle II, a low-density detector is essential. The CDC is filled with a gas mixture of 50 % helium and 50 % ethane, both gasses with small atomic numbers, which minimizes this effect. 14 336 instrumented wires with a diameter of $30\ \mu\text{m}$ are strung through the detector to determine the ionization caused by charged particles traversing the gas. The wires are arranged along the z axis in 32 “axial layers” and tilted by up to 74 mrad to it in 24 “stereo layers”. In this configuration, the detector can be used to determine both transversal and longitudinal information. In addition to providing particle trajectories, the CDC also contributes to particle identification by measuring the energy loss of a particle over the distance which it has traveled.

Particle Identification Detectors Belle II is equipped with two subdetectors dedicated to charged particle identification, the Time-Of-Propagation counter (TOP) in the barrel and the Aerogel Ring-Imaging Cherenkov detector (ARICH) in the forward endcap [41].

The TOP consists of 16 quartz bars along the z axis, each 2.6 m long, 45 cm wide and 2 cm thick. It is installed at a distance of approximately 120 cm from the IP and covers an angular region of 31° to 128° . When charged particles at high energies pass the quartz, they create photons due to the Cherenkov effect. The photons are reflected internally in the bars and propagate to the end of the bar. There, the information contained in the characteristic Cherenkov ring is reconstructed from positional and temporal measurements to determine the particle speed. This requires a determination of the particle production time with a resolution of about 50 ps. The second detector dedicated to PID, the ARICH, covers an angular region of 14° to 30° and also utilizes the Cherenkov effect. The ARICH is designed to distinguish charged particle species at momenta from 0.4 GeV up to 4.0 GeV and is constructed of 4 cm thick aerogel tiles. To capture the location and intensity of photons produced in each aerogel tile, 73×73 mm hybrid avalanche photon detectors with 144 readout channels are used.

Electromagnetic Calorimeter The Electromagnetic Calorimeter (ECL)'s main purpose is the detection of photons and neutral hadrons which shower in this subdetector and are reconstructed as "clusters," i.e. localized energy depositions. In addition to the detection of neutral particles, the ECL is also used to distinguish between electrons and charged hadrons. The former ones deposit their entire energy in the calorimeter while for the latter ones a large part of the energy passes through the ECL undetected. The ECL has an angular acceptance of 12.4° to 155.1° , with gaps at 31.4° to 32.2° and 128.2° to 130.7° where the end caps and the barrel meet. The active material is composed of 8736 thallium-doped caesium iodide crystals, each 30 cm long which corresponds to 16.1 radiation lengths. These crystals with 29 distinct shapes were previously used in the Belle detector. At the end of each crystal, a photo sensor is mounted to measure the scintillation light produced in the crystal. The upgrade of the ECL mainly concerns the calorimeter electronics which must be able to handle much higher background rates. To distinguish between photons produced in the current events and scintillation remnants from the previous event, the readout electronics are redesigned to finely sample the wave forms produced in the photo sensors.

K_L^0 and muon detector The outermost subdetector, the K_L^0 and muon detector (KLM), is a sampling calorimeter in which the absorber material is formed by iron plates which also act as the magnetic flux return of the solenoid. The active components of the detector in the outer barrel are glass electrode resistive plate chambers (RPCs). While the RPCs can handle the particle flux in this region, their long dead time makes them unsuitable for the inner barrel and the end caps in Belle II. The KLM therefore uses scintillator strips, wavelength-shifting fibers and silicon photomultipliers in these two regions. This presents a change from Belle, where the entire KLM was instrumented with RPCs. The KLM covers an angular region of 25° to 40° and is reached by particles with momenta above 600 MeV.

Table 3.2: Overview of the data set recorded by the Belle II experiment between 2019 and 2022. The first uncertainty is statistical, the second is systematic. The methodology with which the luminosity is measured is described in Ref [42]. The main data set for B flavor physics is recorded at the $\Upsilon(4S)$ resonance, with a supporting off-resonance data set recorded at beam energies 60 MeV below the resonance. The two “scan” data sets consist of data recorded at various beam energies and are needed to study heavy quarkonia and measure the mass and width of the $\Upsilon(4S)$ resonance.

Beam energy	Luminosity / fb^{-1}
$\Upsilon(4S)$	$361.601 \pm 0.021 \pm 2.327$
$\Upsilon(4S) - 60 \text{ MeV}$	$42.279 \pm 0.007 \pm 0.272$
$\Upsilon(4S)$ scan	$0.078 \pm 0.001 \pm 0.001$
$\Upsilon(5S)$ scan	$19.661 \pm 0.004 \pm 0.118$
Total	$423.673 \pm 0.022 \pm 2.543$

3.4 Data Sets

For the presented analysis, the sizes of both recorded and simulated data sets present limitations on the precision with which the parameters of interest, the $B \rightarrow \rho \ell \nu_\ell$ branching fractions, can be determined. Here and throughout this thesis, **recorded data** refers to events measured and stored by the Belle II detector while **data** in general refers to both simulated and recorded samples.

Recorded data sample Since the end of its commissioning phase in 2019, the Belle II experiment has collected a data set with an integrated luminosity of $(432 \pm 3) \text{ fb}^{-1}$. This is competitive with the data set size recorded by the BaBar collaboration and corresponds to approximately 50% of the data set recorded by the Belle collaboration [17]. A detailed breakdown is given in Table 3.2, the total data set size over time is visualized in Fig. 3.4.

Simulated data sample In addition to recorded data samples, the analysis presented in this thesis also uses samples simulated with the methods described in Chapter 5. The simulated $B\bar{B}$ samples correspond to an integrated luminosity of 3 ab^{-1} , the simulated $q\bar{q}$ samples correspond only to 1 ab^{-1} . To limit systematic uncertainties from the description of $B \rightarrow X_u \ell \nu_\ell$ signal and background decays, a large data sample describing 100 million resonant $B \rightarrow X_u \ell \nu_\ell$ decays with $X_u \in \{\rho, \omega, \eta, \eta'\}$ and 50 million nonresonant $B \rightarrow X_u \ell \nu_\ell$ decays is used. Additionally, a simulated sample of 15 million resonant $B \rightarrow X_u \ell \nu_\ell$ decays with $X_u \in \{f_0(500), f_0(980), f_2(1270), \rho^0(1450)\}$ is used to describe this possible background.

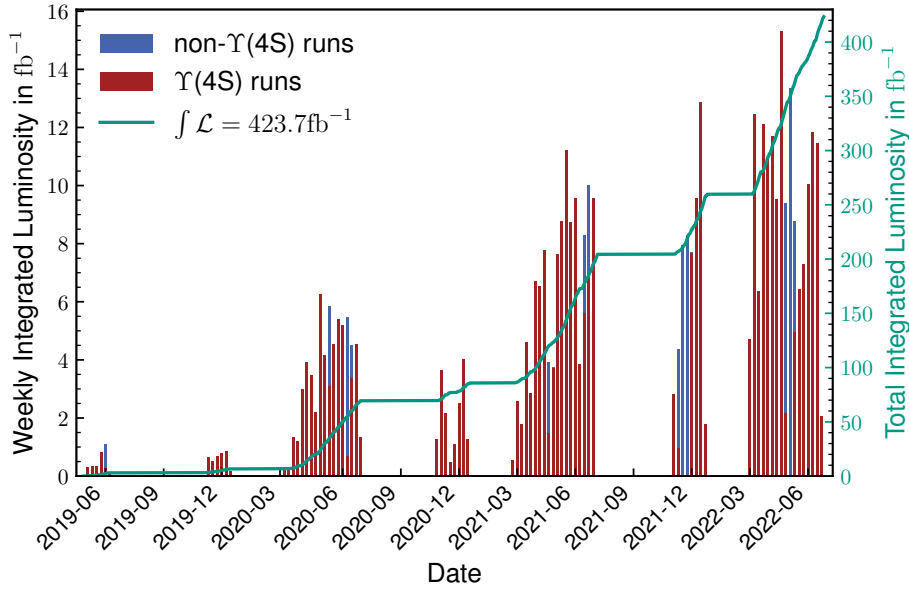


Figure 3.4: The weekly and total integrated luminosity recorded by the Belle II experiment since 2019. The label “ $\Upsilon(4S)$ runs” refers to data taking periods in which SuperKEKB operated with a cms energy of $\sqrt{s} \approx 10.58$ GeV, the label “non- $\Upsilon(4S)$ runs” refers to data taking periods at which the accelerator operated 60 MeV below the $\Upsilon(4S)$ resonance or scanned various energy points. Values taken from Ref. [43].

3.5 The Belle II Computing Grid and XRootD Caching

The following section has been adapted from a conference paper submitted to the proceedings of the 21st International Workshop on Advanced Computing and Analysis Techniques in Physics Research.

In modern, global particle physics collaborations, the sophistication of computational methods and the size of recorded data sets are continually growing to measure ever-smaller effects. These two demands, together with the distributed nature of these collaborations, make grid-based computing infrastructure a logical conclusion. In the course of this work, a caching approach has been investigated which decouples file storage from computing resources to increase efficiency in the Belle II computing grid.

For centralized production of simulated data as well as user analysis, Belle II uses a computing grid spanning approximately 30 computing centers. The grid decouples workload requirements from individual computing sites which increases reliability and accessibility for all members of the collaboration. To manage this diversity of resources, Belle II has selected the DIRAC middleware [44] for workload management and the data management system Rucio [45] to distribute data sets among grid storage endpoints — so-called “Storage Elements” — worldwide.

Grid-based approaches in general average out the computing demands of large collaborations by distributing workloads between the resources. This presents challenges for workloads

requiring specific data sets as workflow management systems only submit computing tasks, so-called “jobs”, to sites which have local copies of these data sets. This concept is referred to as “data locality.” This requirement for data locality, which decreases network strain caused by transferring data over wide-area networks, significantly reduces the number of eligible sites for individual jobs bound to specific input data. While systems such as Rucio can be used to increase the amount of replicas of popular data sets and thus the number of sites capable of executing jobs which require them, this method requires identification of popular files first. Additionally, this approach also assumes static availability of CPU resources at computing. However, this is not always guaranteed for e.g. resources not explicitly dedicated to HEP analysis, so-called “opportunistic resources” [46]. A consequence of the imperfect distribution of data sets and thus jobs are high numbers of waiting jobs at individual sites as depicted in Fig. 3.5.

Transparent, dynamic caches at individual sites present an alternative solution for this problem. They can be used to dynamically add copies of popular data sets to storage systems which do not have to fulfill the requirements on redundancy, reliability, and long-term commitment usually imposed on Storage Elements. In addition to increasing the utilization of resources already integrated into the grid, lightweight caching solutions could also allow sites which currently do not meet the reliability requirements for grid sites to contribute storage and CPU resources to the Belle II collaboration.

The XRootD project The XRootD project [47, 48] provides a data transfer protocol as well as server and client implementations to efficiently transfer files between computing sites connected via wide-area networks. The XRootD protocol is particularly suited to transfer large files typically found in HEP environments and allows partial transfer of supported file types, so-called “streaming.”. Most notably, the XRootD protocol has been adopted by the A Large Ion Collider Experiment (ALICE) collaboration as the primary data transfer protocol. Via plugin interfaces in both the XRootD client and server, additional protocols such as HTTP are supported. This is desirable for application in the Belle II grid: While XRootD protocol support exists incidentally at multiple computing sites, all sites must support HTTP.

Disk-based proxy caches with the XCache concept are also supported via the XRootD plugin interface. As shown in Refs. [49, 50], XCaches provide opportunities to reduce network bottlenecks and increase the CPU utilization of jobs. Both the XRootD server and the associated XCache plugin are relatively lightweight and require little configuration. Deployment on a single server is straightforward, with the option to expand to multiple machines via XRootD’s regular cluster manager mechanism if the demand cannot be satisfied by a single machine anymore. Via the XRootD streaming mechanism, the XCache can provide files already once they are partially downloaded, introducing almost no delay compared to access without the cache.

To redirect selected file access to the XCache, a XRootD client-side plugin is available as part of the XRootD software package. The redirection is completely transparent to the user

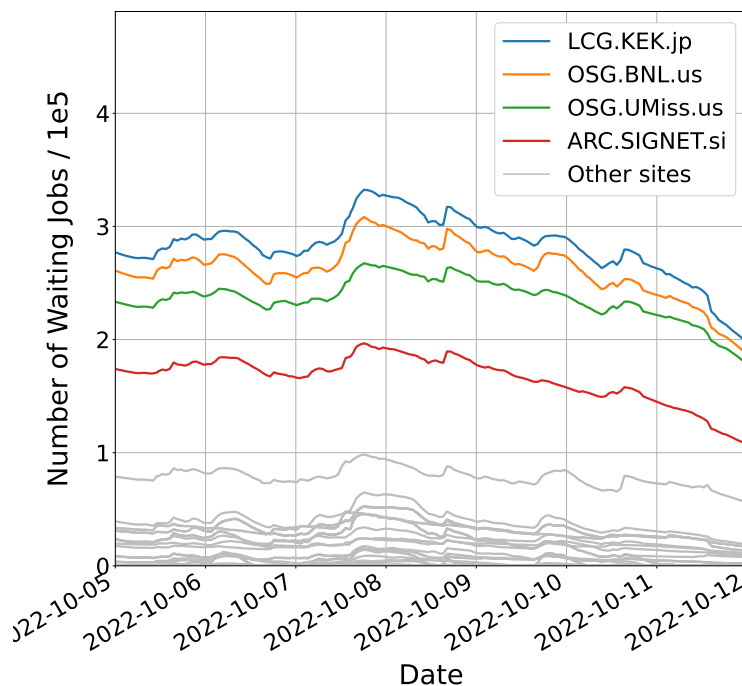


Figure 3.5: Number of waiting jobs at different Belle II grid computing sites, captured during one week. Unequal resource demands lead to significantly more than 100,000 jobs waiting at four sites while the remaining sites have much fewer waiting jobs. This uneven amount of waiting jobs can reduce the efficiency of the grid.

as well as the DIRAC and Rucio systems.

Implemented Setup An XCache instance is deployed to a node at the WLCG Tier1 / Belle II Raw Data Center “GridKa” which allows access to several opportunistic computing resources [46]. These resources provide up to 2200 CPU cores to Belle II and are referred to as `LCG.KIT-TARDIS.de`. On these opportunistic computing resources, all read access via the XRootD client is redirected by the client-side plugin mentioned above. No caching is performed for write access via XRootD or other protocols. This setup is illustrated in Fig. 3.6.

The 40-core server node hosting the XCache is equipped with 256 GB of memory and a 100 Gbit s^{-1} network interface. Via this network interface, the machine has access to a `ceph` distributed storage cluster [51] of which 500 TB is available as storage space for the cache.

To test the functionality of the XCache instance and the associated performance monitoring, all read accesses to the GridKa storage back-end by Belle II jobs on `LCG.KIT-TARDIS.de` is redirected to the XCache instance at GridKa. This does not add any performance or job load distribution advantages as no network bottlenecks are removed between the opportunistic resources and the storage. It does, however, allow evaluating the file reuse rate of typical jobs

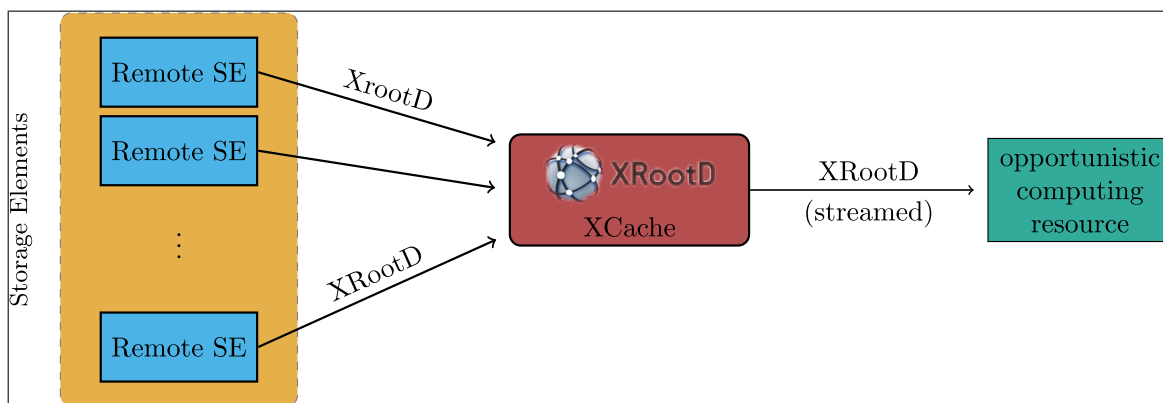


Figure 3.6: Schematic depiction of file access through an XRootD XCache. The label “Remote SE” refers to Storage Elements located at a computing sites not physically close to the server on which the XCache is hosted.

in a Belle II production environment, both for centrally launched MC productions as well as user analysis workloads.

In a second step, the XCache also caches data located at the Institut Jožef Stefan (IJS) in Ljubljana, Slovenia. Here, actual performance benefits are to be expected, however not enough data on this has been collected yet as caching has only been enabled for roughly two weeks at the time of writing.

While caching the GridKa storage back-end is completely transparent to DIRAC, caching data located at remote sites requires changes in the central configuration. To match jobs to the individual sites, DIRAC uses information on the Storage Elements connected to each computing site. Amending the configuration allows jobs at the `LCG.KIT-TARDIS.de` computing site to access data located at the IJS.

To evaluate the performance of the XCache, the network monitoring streams provided by XRootD are captured and unpacked using the `xrootdlib` library [52]. The unpacked information is forwarded to an `elasticsearch` database [53] where it is collected for analysis.

Results Over a period of 14 months, the XCache is operated in a Belle II production environment to demonstrate the utility of the technology. In the last two weeks of operation, the XCache also started caching data from a remote site, however these data points have been excluded from the following results as they do not allow independent statistical analysis yet. Key performance metrics are collected in Table 3.3. Here the computing jobs are divided into three categories: Data from simulation (MC) production jobs, analysis jobs using recorded data and analysis jobs using Monte Carlo (MC). The first category makes up the bulk of jobs submitted to the Belle II grid, an effect which is increased on the `LCG.KIT-TARDIS.de` site due to the specifics of its configuration. MC production jobs use only a limited set of input files, necessary to accurately describe beam-induced backgrounds. Due to the high number of jobs and a comparatively small number of input files, the cache hit rate for these file requests is high. The second category, analysis jobs using real data, shows a lower cache hit rate.

Table 3.3: Cache performance metrics, collected over 14 months and separated into types of jobs.

Job type	Written into cache (TiB)	Read from cache (TiB)	Average access count	Cache hit rate	Total hits in %
MC production	18.72	573.10	30.5	96.7 %	98.89
Data analysis	2.37	5.97	2.4	57.5 %	0.36
MC analysis	11.37	15.47	1.4	27.6 %	0.76

The total contribution to the amount of cache hits is small, but the cache hit rate itself is above 50%. This is most likely due to the limited and universal nature of all recorded data so far. The last category, analysis jobs on simulated data, contains a varied set of simulated input files. While data sets are centrally produced, they can be highly analysis specific. This explains the low cache hit rate.

For MC production jobs, clear benefits are observed by employing the XCache technology as average access counts above 30 indicate high re-use of data sets. For user analysis jobs on both recorded and simulated data, no final conclusion on the usefulness of caching data can be drawn yet. As the XCache storage is far from being filled with approximately 6.5% utilization and the data set sizes for Belle II are expected to grow rapidly with the amount of collected data, higher cache hit rates and average access counts are to be expected.

Chapter 4

Tools

In this chapter, I introduce the methods and tools used in the presented analysis. I introduce Boosted Decision Trees (BDTs), the Belle II software framework as well as particle identification, B tagging, and vertex fitting methods employed at Belle II.

4.1 Boosted Decision Trees

During the last 20 years, machine learning methods have become indispensable in multivariate analysis due to their relative ease of use and ability to exploit complex correlations. While both neural network classifier and decision tree classifier methods are commonly used, the latter are often preferred due to their robustness and good performance without complex tuning procedures. They also offer straightforward access to feature importance estimation. An introduction to BDTs is given in Ref. [54], here only the key concepts are summarized.

A single decision tree can be thought of as a series of selections, each optimized on a figure-of-merit using a labeled data set until a separation maximum on a statistically independent data set is reached. The depth of the decision tree is then given by the maximum number of consecutive selections and is a parameter which must be optimized for each classification task. The approach with only one deep decision tree is straightforward but susceptible to statistical fluctuations in the training data set and typically not very powerful. A combination of relatively shallow decision trees, a “forest”, on the other hand, has shown much better performance. Two different approaches are popular to create this combination: Bagging [55] and boosting [56]. In the bagging approach, the training data set is resampled with replacement to create (often smaller) subsets of the data from which each decision tree is optimized. In the boosting approach, decision trees are optimized iteratively by assigning weights to the data in each iteration such that previously misclassified data samples have larger importance. Gradient boosting [57] and stochastic gradient boosting [58] build on these two concepts by applying the concept of gradient descent to decision trees. The resulting approach is referred to as Stochastic Gradient Boosted Decision Trees (SGBDTs).

As described below, this work uses decision-tree classifiers as part of the B tagging method,

the background suppression, and the particle identification. In two of these applications, the FastBDT software package [59] is used which implements SGBDTs in a fast and cache-optimized way. FastBDT is robust to missing data, supports preprocessing of data with equal-frequency binning, and provides several feature importance estimation methods.

4.2 The Belle II Analysis Software Framework

The Belle II Analysis Software Framework (basf2) [60] contains the majority of software necessary to process data at Belle II. It is described in detail in Ref. [61] and is both used for real time data taking and post-processing of the data for specific analyses. The bulk of the software is written in the C++ language and builds on the popular ROOT software framework. To increase the flexibility of this approach, the framework is separated into individual modules, each dedicated to a specific functionality. Using the Python scripting language, these modules are both configured and arranged in almost arbitrary order.

The *analysis* subpackage is of particular importance for the presented measurement. The analysis methods provided here operate mostly on particle candidates, i.e. particle hypotheses assigned to tracks, clusters or combinations thereof. This recombination is usually not unambiguous within an event and typically results in more than one candidate when recombining the detector objects to e.g. a $\Upsilon(4S)$ candidate. Therefore, care must be taken to either consider the multiplicity in each event in the signal extraction or reduce it to unity.

For each candidate in basf2 one can also define a collection of tracks and clusters in the event which do not belong to the candidate. This Rest of the Event (ROE) is particularly useful in suppressing background in which particles in background processes are missed in the reconstruction or in which no B meson is produced.

4.3 Particle Identification

Particle identification in Belle II uses observables from all subdetectors except the PXD. The TOP and ARICH detectors provide indirect measurements of the particles' velocity, the tracking detectors measure the energy loss per distance and the ECL calorimeter measures the deposited energy. The KLM allows discriminating between pions and muons by the depth which they reach in the detector.

To combine the output of these detectors to high-level quantities, Belle II uses two main approaches. The first approach, described in Ref [1], is based on the assumption that the observables provided by each subdetectors are independent between the detectors and that thus the different particle hypotheses can be described by independent likelihood functions. These likelihood functions $\mathcal{L}^d(\mathbf{x} | i)$, defined for each particle hypothesis i as joint probability density functions (PDFs) based on a set of observables \mathbf{x} in each detector d , are determined using either simulation, well-measured physical processes, or analytic descriptions. The global likelihood $\mathcal{L}(\mathbf{x} | i)$ is then the product of the individual $\mathcal{L}^d(\mathbf{x} | i)$ for each hypothesis.

Leveraging Bayes' theorem, the probability of identifying a particle candidate in a hypothesis i against all other hypotheses j is then given by

$$P(i | \mathbf{x}) = \frac{\mathcal{L}(\mathbf{x} | i)}{\sum_j \mathcal{L}(\mathbf{x} | j)} \quad (4.1)$$

which, according to the Neyman-Pearson lemma, gives the optimal discrimination under the assumptions outlined above.

In addition to this likelihood product approach, basf2 also provides a second particle identification method [62]. A drawback of the likelihood product approach described above is that the likelihoods for the ECL are only determined in terms of E/p , the ratio of energy deposited in the calorimeter and the particle momentum. While this variable is very powerful for high-momentum particles, discrimination for lower-momentum particles is improved by combining several low-level features of the calorimeter shower. These variables describe the lateral and longitudinal shower shape development as well as the quantized sensor wave forms and are highly correlated, making them ideally suitable for combination in a BDT. To add the particle identification information from other subdetectors to this combination, their likelihoods are used directly. In contrast to the likelihood product approach, this method is also able to consider potential correlations between the individual likelihoods and thus between the subdetectors.

4.4 Vertex Fitting

When combining two or more final-state particle candidates to e.g. a ρ meson, the momentum estimate obtained by adding the individual final-state particle momenta does not optimally exploit the available information. Instead, under the assumption that these particles originate from a single decay, a χ^2 minimization algorithm can be used to determine the most likely momentum of the composite particle candidate. Besides improving the momentum and mass resolution of composite particles, these vertex fitting algorithms are useful to suppress candidates composed of unrelated final-state particles. The best-fit χ^2 value gives a measure on the probability that a set of particles originates from a single decay. Additional assumptions can be made on particle masses and the production vertex in this determination which allows even better background discrimination with this approach. Several algorithms have established themselves at the B Factories for this task [17, Chapter 6], however those in use at the Belle experiment do not optimally consider the full decay chain in e.g. a $B \rightarrow \rho^0 (\rightarrow \pi^\pm \pi^\mp) \ell \nu_\ell$ decays. With these algorithms, a vertex fit is first performed for the $\rho^0 \rightarrow \pi^\pm \pi^\mp$ decay and subsequently for the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decay. With TreeFitter, Belle II introduces a new vertex fitter which is presented in Ref. [63]. This algorithm uses a Kalman filter to fit the entire decay chain, an approach that was first introduced in BaBar [64]. In this global fit algorithm, for particle species with flight lengths above $1 \mu\text{m}$, i.e. an order of magnitude below the minimum vertex resolution of the detector, separate vertices are assumed for production and

decay. In addition to fitting tracks, the algorithm can also incorporate the energy deposited in calorimeter clusters in the reconstruction if additional information on the particle species from which they originate is given as a constraint.

4.5 Full Event Interpretation

To constrain B meson decays with one or more neutrinos, a common approach at B Factories is to use the recoil from the second B meson. This is only feasible due to the very well-known initial state, the fact that the $B\bar{B}$ pair is produced without additional particles and the almost perfect enclosure of the IP by the detector [17, Chapter 7]. In the context of this work, this method is referred to as “B tagging”. The B meson decaying in the process of interest is called B_{sig} , the second B meson is called B_{tag} . The B_{tag} can be reconstructed both in semileptonic and hadronic decay channels. While the former offer much larger branching fractions in each channel than the latter, the additional neutrino produced in the decay weakens the constraint considerably. Therefore, only hadronic reconstruction is described in the following. Analyses conducted by the Belle collaboration have historically used two different methods for B_{tag} reconstruction. The first one is an algorithm in which fixed selections are used to reconstruct B mesons in several $B \rightarrow D^{(*)}X$ decay modes. The second approach, presented in Ref. [65], increases the number of hadronic reconstruction channels. It then uses machine learning methods to determine a probability metric for the decay to be reconstructed correctly at each reconstruction stage. This second algorithm is named “Full Reconstruction” and was previously used in a hadronically tagged analysis of $B \rightarrow \rho\ell\nu_\ell$ decays with the Belle experiment [66].

The Full Reconstruction algorithm’s successor is named Full Event Interpretation (FEI). It has been developed for the Belle II experiment and is part of the basf2 framework. The remainder of this section aims to give an overview of the algorithm and its usage at Belle II.

The FEI algorithm uses approximately 200 SGBDTs classifiers from the FastBDT package to recombine final-state particles to B mesons. The BDT classifiers are arranged in six stages, each combining particles to reconstruct progressively heavier mesons. Figure 4.1 shows the relationships between the different intermediate and final-state particles.

In detail, the six reconstruction and classification stages of the FEI at the $\Upsilon(4S)$ resonance are:

- Stage 1: Construction of long-lived charged (e^- , μ^- , π^+ , K^+ , p) and neutral (γ and K_L^0) particles from tracks, calorimeter and KLM clusters, and pre-vertexed V0 objects
- Stage 2: Reconstruction of J/ψ , π^0 , and Λ candidates
- Stage 3: Reconstruction of K_S^0 and Σ^+ candidates
- Stage 4: Reconstruction of different D as well as Λ_c^+ candidates
- Stage 5: Reconstruction of different D^* candidates

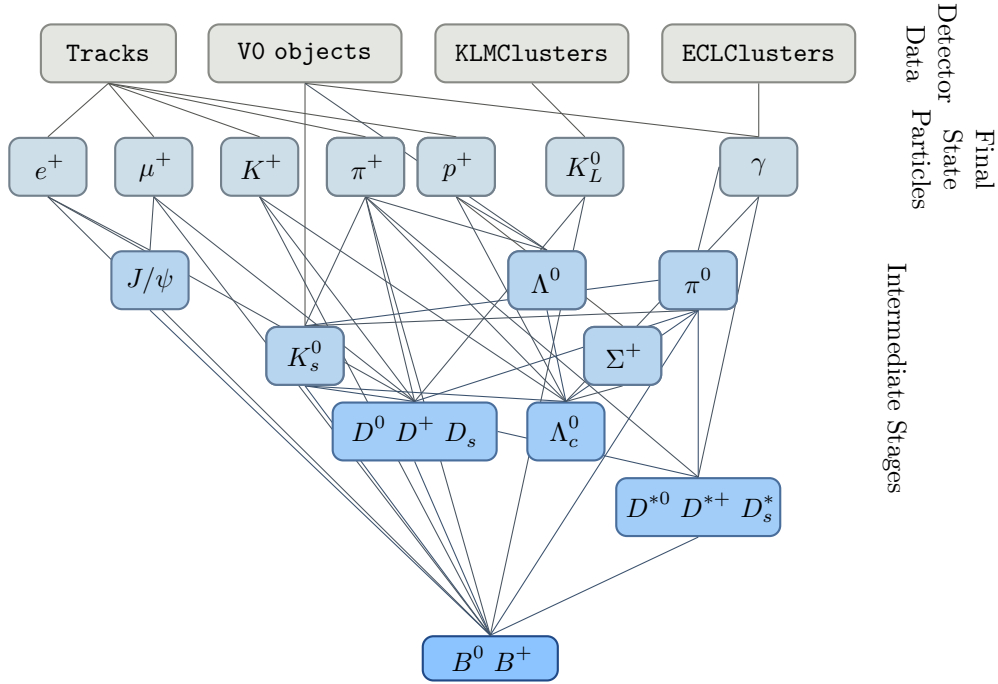


Figure 4.1: The hierarchical nature of the Full Event Interpretation. Particle candidates are reconstructed in six stages and a FastBDT classifier is applied at each stage to rank the candidates. Initial candidates are reconstructed from tracks and clusters, intermediate particles from combinations of candidates from previous stages. The classifier in each stage uses the classifier output of the previous stage as input. Adapted from Ref. [67].

- Stage 6: Reconstruction of B^+ and B_d^0 candidates

In each of the FEI's reconstruction stages, all particle candidates from the previous stages are combined to heavier hadrons. After this recombination, the vertexing algorithm used in Belle [68] is applied to determine the probability of the recombined particles originating from a common decay vertex.

Next, a SGBDT classifier from the FastBDT package is applied to the recombined particle candidates and the 5–20 candidates (the exact value depends on the particle species) with the highest classifier output are propagated to the next stage. This classifier uses features specific to each stage of the decay, however a strong constraint can always be obtained from the production vertex of the decay. This approach, usually called a “best-candidate selection”, differs from the candidate selection in the Full Reconstruction algorithm in which specific classifier output thresholds are defined for each decay channel. The best-candidate selection distributes available resources more evenly between the different events and decay channels since no single event can generate large amounts of possible candidates which would have to be fit in the next stage. The top-level classifier which is used to assign a signal probability to B meson candidates provides 20 candidates for analysis use. While it is possible to simply use the candidate with the highest classifier output, this is not always the best strategy since

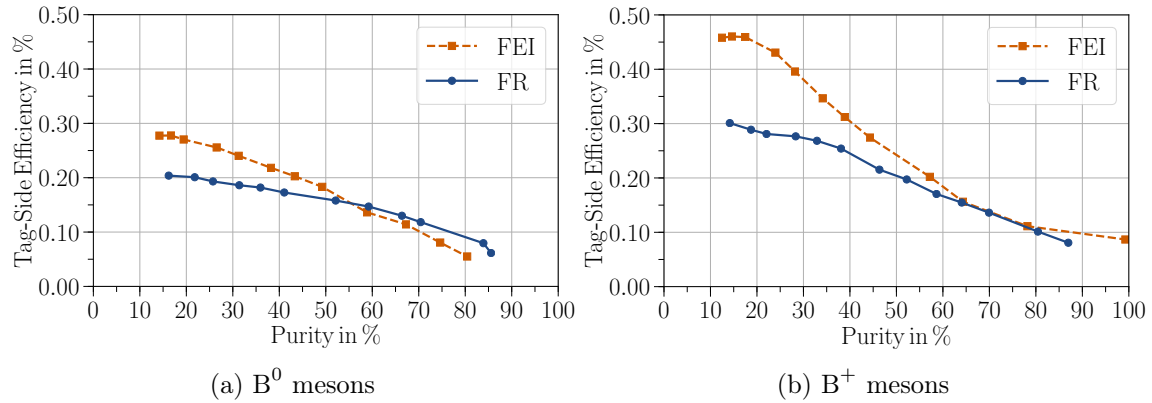


Figure 4.2: Receiver operating characteristic of the Full Reconstruction and the Full Event Interpretation applied to B^0 meson and B meson candidates in recorded data. The tag-side efficiency and purity are measured by fitting the beam-constrained mass spectrum. Taken from Ref. [67].

it does not exploit the information of the signal side. An analysis-specific signal-side selection can be used to discard wrong B_{tag} candidates or to change their ranking. However, this then also requires an analysis-specific calibration for these specific selection criteria as described in Section 6.1.1.

The improved classification method, the best-candidate selection, and the additional channels lead to larger tag-side efficiency and higher purity of the Full Event Interpretation compared to the Full Reconstruction. This is also illustrated in Fig. 4.2 which shows that the FEI out-performs its predecessor in most cases [67]. This is especially true for regions of high efficiency in which most analyses set their working point.

Chapter 5

Simulated Samples

In the presented analysis, simulated samples are used for three main purposes: To design and validate the selection strategy, to determine the reconstruction efficiency of this selection, and to derive expected probability density functions (PDFs) for signal and background expectations needed in the signal extraction. Various models are combined to achieve a simulated description that incorporates as much knowledge about the signal process and relevant background processes as possible. In the following, these will be described in order along with approaches to combine the samples where necessary.

The following description is mainly focused on the simulation of relevant physical processes occurring immediately after the e^+e^- collision. While interactions with the detector (which are modeled using GEANT4 [69]) and its response (which is modeled using basf2) must be simulated in addition to these “primary” processes, no specific effects of this simulation have been identified as of paramount importance for the presented analysis.

5.1 $B \rightarrow \rho \ell \nu_\ell$ Signal Samples

As described in Section 2.2, the semileptonic $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ and $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ decay processes cannot be described with perturbative methods and instead rely on combinations of LCSR predictions and experimental spectra such as the one in Ref. [19]. The results of this combination, illustrated in Fig. 5.1a, can be given in various parametrizations, the one chosen in Ref. [19] originates from Ref. [18] and expands the form factors $F_i(q^2)$ introduced in Section 2.2 as

$$F_i(q^2) = P_i(q^2) \sum_k \alpha_k^i (z(q^2) - z(0))^k, \quad (5.1)$$

where $P_i(q^2) = (1 - q^2/m_B^2)^{-1}$ is the inverse Blaschke factor with the mass of the B resonance m_B .

This definition, which is referred to as Bharucha–Straub–Zwicky (BSZ) parametrization, expands around a parameter

$$z(q^2, t_0) = \frac{\sqrt{t_+ - q^2} - \sqrt{t_+ - t_0}}{\sqrt{t_+ - q^2} + \sqrt{t_+ - t_0}}, \quad (5.2)$$

which is defined such that $|z| < 0.1$ for the pair production threshold $t_+ = (m_B + m_\rho)^2$ and the z origin $t_0 = (m_B + m_\rho)(\sqrt{m_B} - \sqrt{m_\rho})^2$ with the ρ mass m_ρ [19].

Using the Monte Carlo generator EvtGen [70] and the BCL EvtGen model in basf2, the fitted form factors are then used to generate simulated events in which one of the two B mesons is forced to decay via the signal process $B \rightarrow \rho \ell \nu_\ell$.

While the kinematic properties of all particles involved in the decay are mainly governed by the decay models in EvtGen, the mass spectrum of particles with non-negligible widths is given by a relativistic Breit-Wigner (BW) function [5, Chapter 50].

According to the EvtGen user guide [71], the amplitude of the relativistic Breit-Wigner function used to describe the ρ line shape is given by

$$BW(M(\pi\pi)) = \frac{|\mathbf{p}_\pi|}{M(\pi\pi)^2 - m_\rho^2 + iM(\pi\pi)\Gamma(M(\pi\pi))} \frac{B(|\mathbf{p}_\pi|)}{B(|\mathbf{p}'_\pi|)}, \quad (5.3)$$

where $M(\pi\pi)$ is the invariant mass of the two pions, m_ρ is the nominal ρ mass, $|\mathbf{p}_\pi|$ is the pion momentum in the ρ rest frame, and $|\mathbf{p}'_\pi|$ is the pion momentum assuming the nominal ρ mass, again in the frame of the ρ .

The decay width is given by

$$\Gamma(M(\pi\pi)) = \Gamma_0 \left(\frac{|\mathbf{p}_\pi|}{|\mathbf{p}'_\pi|} \right)^3 \frac{m_\rho}{M(\pi\pi)} \left(\frac{B(|\mathbf{p}_\pi|)}{B(|\mathbf{p}'_\pi|)} \right)^2, \quad (5.4)$$

with the nominal ρ width Γ_0 and a Blatt-Weisskopf barrier factor [72, 73]

$$B(x) = \frac{1}{\sqrt{1 + R^2 x^2}} \quad (5.5)$$

with scale parameter $R = 3(\text{GeV}/c)^{-1}$.

This description is congruent with the one given in e.g. Ref. [5, Chapter 50]. Closer investigation, however, reveals that the line shape implemented in EvtGen does not completely correspond to Eq. (5.3). Instead, the approximation

$$iM(\pi\pi)\Gamma(M(\pi\pi)) \rightarrow im_\rho\Gamma(M(\pi\pi)) \quad (5.6)$$

has been used. This common approximation, which is valid for narrow resonances, induces a shift of the resonant peak as illustrated in Fig. 5.1b.

To correct this behavior in all simulated samples, a $B \rightarrow \rho \ell \nu_\ell$ sample without detector interactions or reconstruction effects is generated with EvtGen and histogrammed in fine bins of the generated dipion mass $M(\pi\pi)_{\text{gen}}$. Then, $M(\pi\pi)_{\text{gen}}$ -dependent weights w_i are calculated

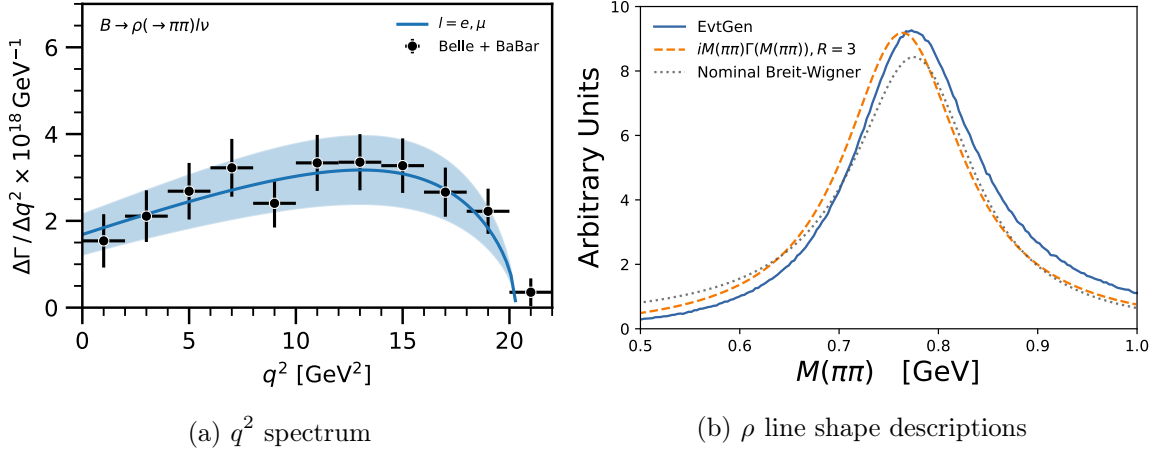


Figure 5.1: Left: Differential decay rates for the fit of BSZ coefficients to $B \rightarrow \rho \ell \nu_\ell$ data performed in Ref. [19]. Figure adapted from Ref. [19].

Right: Line shapes of the ρ meson in the EvtGen simulation and modeled according to Ref. [5, Chapter 50] by Ref. [74].

from the ratio of the histogrammed EvtGen model and the expected distribution.

In the nominal simulated samples, ρ decays from which ρ candidates are reconstructed are identified in both signal and background samples. Based on $M(\pi\pi)_{\text{gen}}$, the correct weight is determined and applied to the $B \rightarrow \rho \ell \nu_\ell$ candidate.

5.2 $B \rightarrow X_u \ell \nu_\ell$ Backgrounds

5.2.1 Exclusive $B \rightarrow X_u \ell \nu_\ell$ backgrounds

In addition to $B \rightarrow \rho \ell \nu_\ell$ decays, several other $B \rightarrow X_u \ell \nu_\ell$ modes contribute to the selected events. These processes can be divided into two categories: Those with measured branching fractions and those which could exist but for which no branching fraction has been established so far.

The first category consists of $B \rightarrow \pi \ell \nu_\ell$, $B \rightarrow \omega \ell \nu_\ell$, $B \rightarrow \eta \ell \nu_\ell$, and $B \rightarrow \eta' \ell \nu_\ell$ decays which (with the small but notable exception of the G-parity suppressed [75] $\omega \rightarrow \pi^\pm \pi^\mp$ decay mode) can only contribute via missed or mistakenly assigned final-state particles. For $B \rightarrow \pi \ell \nu_\ell$ and $B \rightarrow \omega \ell \nu_\ell$ decays, form factor descriptions from Ref. [76] and Ref. [19] are obtained by combined fits of Lattice Quantum Chromodynamics (LQCD) predictions and experimental data, analogous to the determination for $B \rightarrow \rho \ell \nu_\ell$ outlined above. To describe $B \rightarrow \eta \ell \nu_\ell$ and $B \rightarrow \eta' \ell \nu_\ell$ decays, form factors from LCSR calculations given in Ref. [77] are used.

For the second category of $B \rightarrow X_u \ell \nu_\ell$ background decay modes, four processes with $X_u \in \{f_0(500), f_0(980), f_2(1270), \rho^0(1450)\}$ which are all expected to decay to two pions and can therefore not only contribute via misreconstruction, are considered. These are simulated with the generic updated Isgur–Scora–Grinstein–Wise (ISGW2) model for semileptonic decays

Table 5.1: Branching fractions of exclusive $B \rightarrow X_u \ell \nu_\ell$ processes used in the simulated sample. The bottom four exclusive processes are completely unmeasured, the values here only divide the difference between the $B^+ \rightarrow \pi^\pm \pi^\mp \ell \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ branching fractions assuming a $2\times$ larger fraction for $B \rightarrow f_2(1270) \ell \nu_\ell$. For the $B^0 \rightarrow \pi^- \pi^0 \ell \nu_\ell$ branching fraction the value from PYTHIA fragmentation is used.

Process	$\mathcal{B}(B^+) \times 10^{-4}$	$\mathcal{B}(B^0) \times 10^{-4}$
$B \rightarrow \rho \ell \nu_\ell$	1.58 ± 0.11	2.940 ± 0.211
$B \rightarrow \pi \ell \nu_\ell$	0.78 ± 0.03	1.50 ± 0.06
$B \rightarrow \omega \ell \nu_\ell$	1.19 ± 0.09	-
$B \rightarrow \eta \ell \nu_\ell$	0.39 ± 0.05	-
$B \rightarrow \eta' \ell \nu_\ell$	0.23 ± 0.08	-
$B \rightarrow \pi \pi \ell \nu_\ell$	2.27 ± 0.39	9.51
$B \rightarrow f_0(500) \ell \nu_\ell$	0.16	-
$B \rightarrow f_0(980) \ell \nu_\ell$	0.16	-
$B \rightarrow f_2(1270) \ell \nu_\ell$	0.31	-
$B \rightarrow \rho^0(1450) \ell \nu_\ell$	0.16	-
$B \rightarrow X_u \ell \nu_\ell$	38.4 ± 4.8	3.52 ± 0.44

described in Ref. [78]. As no reliable branching fraction estimates are available for this second category, for simulation studies their branching fractions are set to arbitrary values such that their sum together with the branching fraction expectation of the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ signal process is consistent with the total branching fraction of $B \rightarrow \pi \pi \ell \nu_\ell$ decays given in Ref. [79]. The branching fractions used in the simulated sample are listed in Table 5.1.

5.2.2 Inclusive $B \rightarrow X_u \ell \nu_\ell$ Backgrounds

The sum of currently known and measured $B \rightarrow X_u \ell \nu_\ell$ decay modes described above can only account for approximately 25% of the measured inclusive $B \rightarrow X_u \ell \nu_\ell$ branching fraction, i.e. the fraction of semileptonic B decays in which a $b \rightarrow u$ quark transition is involved. To simulate the remaining contributions which are expected to contribute significantly to misreconstructed background events, the inclusive Bosch–Lange–Neubert–Paz (BLNP) [13] description is used with m_b and μ_π^2 HQE parameters obtained in a global fit in Ref. [16]. Additional input parameters are three scales μ_h , μ_i , and $\bar{\mu}$ which describe the transition from the non-perturbative low-energy regime to the high-energy regime of the heavy quark. The triple-differential prediction of the $B \rightarrow X_u \ell \nu_\ell$ decay rate from BLNP is given in bins of the mass of the hadronic system M_X , the lepton momentum in the B frame E_ℓ^B and the invariant mass of the dilepton system q^2 . Unfortunately, the model cannot describe the resonant structures dominating in low ($M_X < 1.4$ GeV) regions of the hadronic mass. The parton-level description provided by this model is then fragmented and hadronized to at least two mesons using the PYTHIA [80] general purpose Monte Carlo generator. This hadronization process partly governs the $M(\pi\pi)$ distribution and introduces a step-like feature at 0.91 GeV, approximately the combined mass of a ρ meson and a π meson. The existence

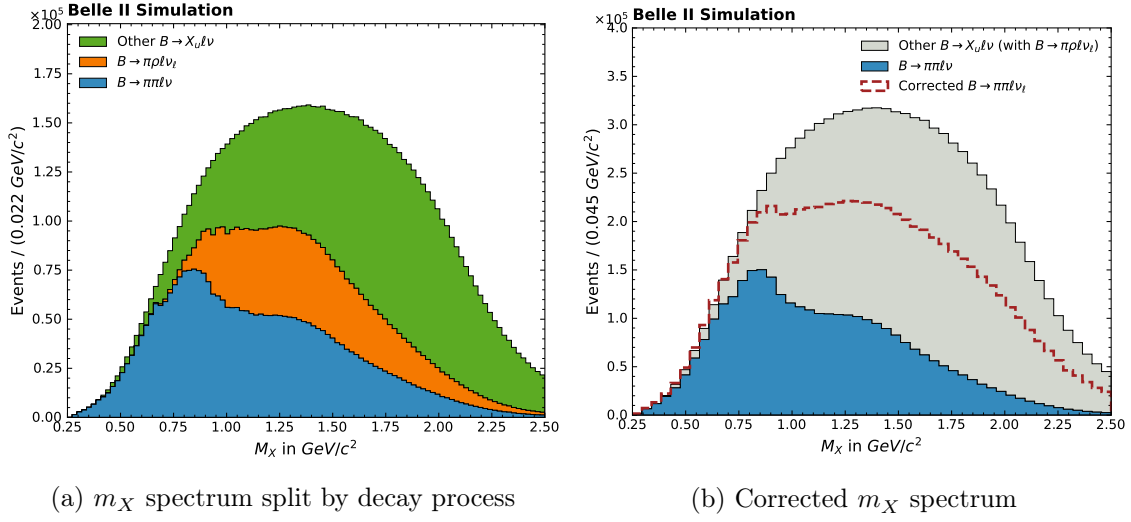


Figure 5.2: The spectrum of the hadronic mass m_X as generated by the BLNP model and fragmented by PYTHIA. As illustrated in the left figure, the $B \rightarrow \pi \pi \ell \nu_\ell$ distribution exhibits a step-like feature at the kinematic threshold of $B \rightarrow \pi \rho \ell \nu_\ell$ production. As this process is not measured, the spectrum is corrected as illustrated on the right.

or magnitude of this step, illustrated in Fig. 5.2a, is not supported by measurements. It is therefore partially removed by finely binning the M_X spectrum and reweighting the $M(\pi\pi)$ distribution such that it would account for an additional 50% of the $B \rightarrow X_u \ell \nu_\ell$ branching fraction in each of the M_X bins. The reweighting does not change the overall branching fraction of the nonresonant $B \rightarrow \pi \pi \ell \nu_\ell$ background. The new distribution is illustrated in Fig. 5.2b and is equivalent to interpolating between the $B \rightarrow \pi \pi \ell \nu_\ell$ distribution from PYTHIA and a model in which the X_u system can only decay to two charged pions. An appropriate systematic uncertainty is assigned for the effect, this is described in Section 8.2.1.

The predicted nonresonant $B \rightarrow \pi \pi \ell \nu_\ell$ branching fraction from this approach is not consistent with the total $B \rightarrow \pi \pi \ell \nu_\ell$ branching fraction given in Ref. [79]. To account for this, the nonresonant $B \rightarrow \pi \pi \ell \nu_\ell$ branching fraction is re-weighted such that it accounts for 50% of the total unmeasured $B \rightarrow \pi \pi \ell \nu_\ell$ branching fraction (i.e. the total branching fraction after subtracting the contribution from $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$), a reduction by approximately one order of magnitude. All other $B \rightarrow X_u \ell \nu_\ell$ processes simulated with the inclusive model are then also weighted to maintain the overall inclusive branching fraction given in Ref. [81]. After the signal extraction as described in Section 7.2.2, the background $B \rightarrow X_u \ell \nu_\ell$ contributions are re-weighted to the best-fit values for the template normalizations.

5.2.3 Hybrid Model

To combine the inclusive BLNP prescription and the various measured resonances in the $M_X < 1.4 \text{ GeV}$ region, a hybrid modelling technique for $b \rightarrow u$ transitions has been proposed in Ref. [82] which has been employed in e.g. Ref. [83].

In this approach, a simulated sample without reconstruction or selection effects is used to determine event-wise weights based on the lepton momentum in the B-frame E_l^B , the squared momentum transfer q^2 and the mass of the hadronic system M_X . The sample is divided in three-dimensional bins of these quantities such that the combined model predicts the number of events as

$$H_i = R_i + w_i I_i, \quad (5.7)$$

with the total number of events H_i , the number of resonant events R_i , the number of inclusive events I_i and the weight $w_i = (I_i - R_i)/I_i$ such that $H_i = I_i$. The bin edges used for this procedure are given by

$$\begin{aligned} M_X &= [0.0, 1.4, 1.6, 1.8, 2.0, 2.5, 3.0, 3.5] \text{ GeV}/c^2, \\ q^2 &= [0.0, 2.5, 5.0, 7.5, 10.0, 12.5, 15.0, 25.0] \text{ GeV}^2/c^4, \\ E_l^B &= [0.0, 0.5, 1.0, 1.25, 1.5, 1.75, 2.0, 3.0] \text{ GeV}. \end{aligned}$$

As in Ref. [83], these three-dimensional bins are motivated by theory as they correspond to the intervals for which the inclusive BLNP prediction is valid. In contrast to Ref. [83], which uses different parameters for the BLNP model, adjacent bins are merged if the exclusive decays eclipse the inclusive model which prevents discontinuities in the weighting procedure.

Figure 5.3 displays the effect of applying the weights determined for charged $B \rightarrow X_u \ell \nu_\ell$ decays back to the sample from which they are calculated. Figures illustrating the same for neutral $B \rightarrow X_u \ell \nu_\ell$ decays can be found in Chapter E.

As illustrated in Fig. 5.3c, the reweighting procedure produces an unphysical step at the bin edge at 1.4 GeV. Analyses must consider the effects of this step, either by not extending templates across it in studies of resonant shapes such as the one presented or by integrating over this range in the signal extraction in inclusive studies such as the one presented in Ref. [81].

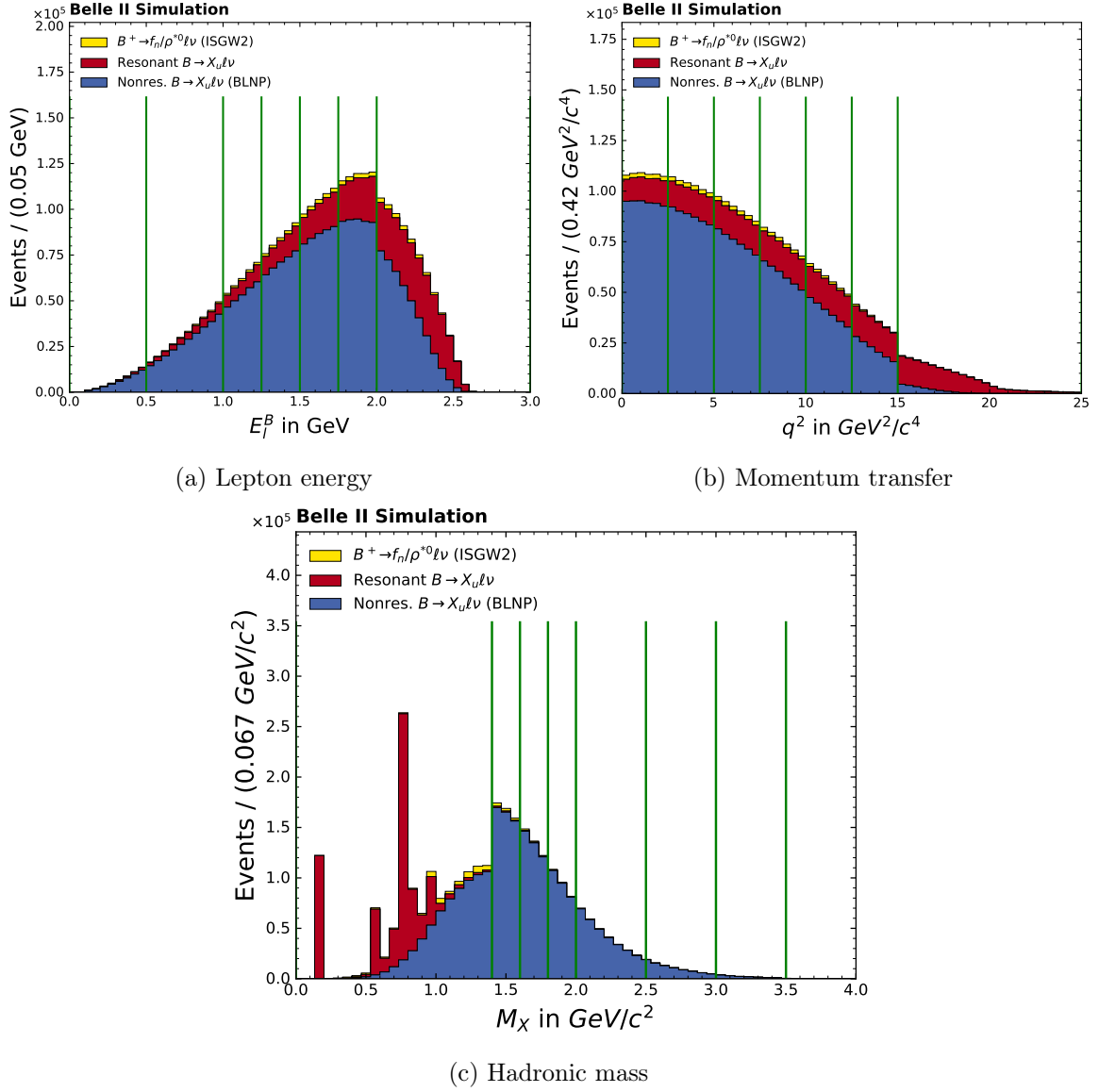


Figure 5.3: Hybrid MC model combination for charged $B \rightarrow X_u \ell \nu_\ell$ decays projected to the reweighting variables with bins of the reweighting procedure indicated by green vertical lines. Shown are the lepton energy in the B frame E_l^B , the momentum transfer q^2 , and the mass of the hadronic system M_X .

5.3 Other Simulated Backgrounds

B \rightarrow X_clν_ℓ decays In addition to the semileptonic B decays involving u quarks described above, important backgrounds originate from CKM-favoured semileptonic decays involving c quarks such as e.g. B \rightarrow D l ν_ℓ. While the decay kinematics of these processes are distinctly different from those of B \rightarrow X_u l ν_ℓ decays, their much higher branching fractions require accurate modeling in the extraction of $\mathcal{B}(B \rightarrow \rho l \nu_\ell)$. Fortunately, the current degree of knowledge over individual B \rightarrow X_c l ν_ℓ decays allows describing approximately 90% of the inclusive B \rightarrow X_c l ν_ℓ branching fraction [84]. The remaining 10% of B \rightarrow X_c l ν_ℓ decays (the “gap”) must be described using so-far unmeasured processes such as B \rightarrow D η l ν_ℓ. The branching fraction values with uncertainties are given in Table 5.2, the prescription to calculate the branching fraction of these individual processes combines absolute measurements, branching fraction ratios and knowledge of isospin relations between the involved hadrons. The approach is described in more detail in Refs. [81, 85]. To correct the simulated sample such that these branching fractions are reproduced, the weight w_B is calculated as

$$w_B = \frac{\mathcal{B}_{\text{new}}(B \rightarrow X_c l \nu_\ell)}{\mathcal{B}_{\text{old}}(B \rightarrow X_c l \nu_\ell)}, \quad (5.8)$$

with the branching fraction $\mathcal{B}_{\text{old}}(B \rightarrow X_c l \nu_\ell)$ in the simulated sample and the branching fraction $\mathcal{B}_{\text{new}}(B \rightarrow X_c l \nu_\ell)$ to which the sample is updated. The weights w_B are then applied to the appropriate decays identified in the simulated sample.

In addition to the branching fraction, the accurate simulation of D^(*) decays also requires form factors in the Boyd–Grinstein–Lebed (BGL) parametrization as input to the BGL EvtGen model in basf2. For D*, these are determined in Ref. [86], for D they are given in Ref. [87].

D meson decays Besides accurate descriptions of semileptonic B decays, the signal extraction approach in the presented analysis relies on knowledge of the distribution formed by the invariant mass of two pions. In B \rightarrow X_c l ν_ℓ background processes, this distribution is influenced by the branching fractions and kinematic distributions of individual D^(*) meson decays. However, in contrast to B decays, the branching fractions of D decays are relatively well-understood and comparatively few assumptions and compromises must be made to create a consistent description. To be able to describe the kinematic distributions, however, additional information on the resonant structure of individual three- and four-body decays is needed. The approach used for the simulated backgrounds is described in detail in Ref. [88], here a brief overview is given. If available, the resonant structures as measured by several collaborations with Dalitz plot approaches (see Refs. [89, 90, 91, 92, 93, 94, 95, 96, 97, 98]), are used directly in the simulation. Frequently, no such measurements are available (which is the case for all D decays to more than three hadrons). In this case, the sum of non-interfering resonances for which branching fraction measurements exist or can be deduced using e.g. isospin relations, is used to describe the kinematic properties in the decay as accurately as possible. This treatment can of course not account for interference effects. For some final

Table 5.2: Branching fractions of exclusive $B \rightarrow X_c \ell \nu_\ell$ processes determined in Ref. [85] such that they sum to the measured inclusive $B \rightarrow X_c \ell \nu_\ell$ branching fraction given in the last row. The bottom two exclusive processes involving η mesons are completely unmeasured so 100% uncertainties must be assigned.

Process	$\mathcal{B}(B^+) \times 10^{-3}$	$\mathcal{B}(B^0) \times 10^{-3}$
$B \rightarrow D \ell \nu_\ell$	24.10 ± 0.71	22.40 ± 0.66
$B \rightarrow D^* \ell \nu_\ell$	55.02 ± 1.15	51.14 ± 1.08
$B \rightarrow D_1^\ell \nu_\ell$	6.63 ± 1.09	6.16 ± 1.01
$B \rightarrow D_0^* \ell \nu_\ell$	4.20 ± 0.75	3.90 ± 0.70
$B \rightarrow D_1' \ell \nu_\ell$	4.20 ± 0.90	3.90 ± 0.84
$B \rightarrow D_2^* \ell \nu_\ell$	2.93 ± 0.32	2.73 ± 0.30
$B \rightarrow D \pi \pi \ell \nu_\ell$	0.62 ± 0.89	0.58 ± 0.82
$B \rightarrow D^* \pi \pi \ell \nu_\ell$	2.16 ± 1.02	2.01 ± 0.95
$B \rightarrow D_s K \ell \nu_\ell$	0.30 ± 0.14	-
$B \rightarrow D_s^* K \ell \nu_\ell$	0.29 ± 0.19	-
$B \rightarrow D \eta \ell \nu_\ell$	3.77 ± 3.77	4.09 ± 4.09
$B \rightarrow D^* \eta \ell \nu_\ell$	3.77 ± 3.77	4.09 ± 4.09
$B \rightarrow X_c \ell \nu_\ell$	108.00 ± 4.00	101.00 ± 4.00

states, the branching fractions of the measured resonances cannot account for the measured total branching fraction. To still maintain overall consistency in the simulation model, the remainder is filled using decays in which the final-state particles are distributed evenly in phase space.

While the overall description of the D decays with this approach strikes a good balance for most analyses at Belle II, in the presented analysis specific focus must be placed on decays involving ρ resonances. One such case is the resonant process $D^0 \rightarrow K^- a_1(1260)$. In the Belle II simulated sample, the branching fraction of the entire $D^0 \rightarrow K^- a_1(1260)$ decay is set to 7.8%. Together with the $a_1(1260) \rightarrow \rho^0 \pi^+$ branching fraction, which is set to 30.9% (estimated from [99]), this gives an overall branching fraction of 2.4% for the $D^0 \rightarrow K^- a_1(1260)(\rightarrow \rho^0 \pi^+)$ process. This value disagrees by approximately 80% with the branching fraction of $(4.32 \pm 0.32)\%$ for the $D^0 \rightarrow K^- a_1(1260)(\rightarrow \rho^0 \pi^+)$ process given in Ref. [5]. A second process for which a discrepancy is found is the nonresonant decay $D^0 \rightarrow K^- \pi^+ \pi^+ \pi^-$. The simulated sample uses a branching fraction of 0.88%, while Ref. [5] gives a value of (1.81 ± 0.07) for this process. To correct for these two discrepancies, candidate-based weights are also used here to enhance the contribution from $D^0 \rightarrow K^- a_1(1260)$ and $D^0 \rightarrow K^- \pi^+ \pi^+ \pi^-$ in the simulated samples. To avoid double-counting calibration factors, the weights are only applied to events in which at least one of the pions used to reconstruct the ρ candidate originates from the mismodeled D decay. Discrepancies between simulated and recorded data in D mesons used to reconstruct the FEI tag are already corrected with specific calibration factors as detailed in Section 6.2.2.

Light-quark pair interactions A second major class of processes relevant as $B \rightarrow \rho \ell \nu_\ell$ backgrounds is light-quark pair production from e^+e^- interactions, also referred to as continuum or $q\bar{q}$ background. As the resulting hadrons are not produced in resonance, EvtGen cannot be used directly. Instead, a combination of two MC generators is used first: Initially, the KKMC generator [100] is used to produce quark pairs while taking into account Initial State Radiation (ISR), Final State Radiation (FSR), and the interference between them. The general-purpose MC generator PYTHIA then produces a parton shower from the quark pair which is then hadronized. The decay of unstable hadrons produced in this hadronization process is then again simulated using EvtGen. This procedure is able to reproduce key features of the off-resonance data sets, however due to the phenomenological nature of the PYTHIA simulation it is highly dependent on many input parameters which must be determined from recorded data. At the current stage of the experiment, this determination has not yet been finalized, so simulated continuum data is expected to be less reliable than $B\bar{B}$ simulated data.

While decay processes in these two categories are expected to contribute most to the irreducible backgrounds in this analysis and therefore require the most precise description, in total a much larger number of processes is simulated for this and other analyses. At present, simulated B decay samples allow decays in approximately 1300 channels for neutral mesons and 1100 channels for charged ones. In addition, processes which produce no hadrons in the primary interaction (such as $e^+e^- \rightarrow \tau^-\tau^+$ which is simulated using TAUOLA [101]) are simulated using dedicated MC generator software. A detailed overview of all MC generators used in Belle II is given in Ref. [102].

Chapter 6

Reconstruction and Selection

6.1 Event Selections

In addition to selections applied to whole collision events at Belle II, the candidate-based analysis approach described in Section 4.2 allows disentangling the two B mesons in each $\Upsilon(4S)$ event entirely, reconstructing and selecting them almost independently. In the following, I will first describe the selections applied to the tag-side B meson candidate, then those applied to the signal-side one.

The decay signatures for both $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ are illustrated in Fig. 6.1 together with the decay of a tag-side B meson.

6.1.1 Tag-Side Selection

Compared to its predecessor algorithm, the FEI algorithm described in Section 4.5 offers several advantages. Among these are more configurability and the potential for increased reconstruction efficiency as demonstrated in Ref. [67]. However, as the FEI algorithm currently exhibits significantly different performance between simulated and measured data, calibration studies using well-known physical properties must be made to evaluate this performance difference and apply appropriate correction factors to simulated data. To ensure the correctness of these calibration factors for analyses such as the ones presented in this work, the selections used in calibration must be replicated exactly. This reduces the absolute reconstruction efficiency considerably as the tag-side selection presently cannot be optimized for specific signal channels. There are, however, calibration factors available for three working points on the efficiency-purity-curve of the FEI algorithm output. As the signal processes under investigation are CKM-suppressed and therefore have small branching fractions, a high-efficiency working point is chosen by requiring a value of at least 0.001 for the BDT classifier output of the B_{tag} meson. To further enhance the ratio of correctly reconstructed to incorrectly reconstructed B_{tag} candidates, the basic kinematic properties of B mesons at B Factories are considered: Correctly reconstructed B candidates have invariant masses corresponding to the well-known mass of B mesons and a total energy equal to the beam energy $E_{\text{CMS, beam}}$ in the

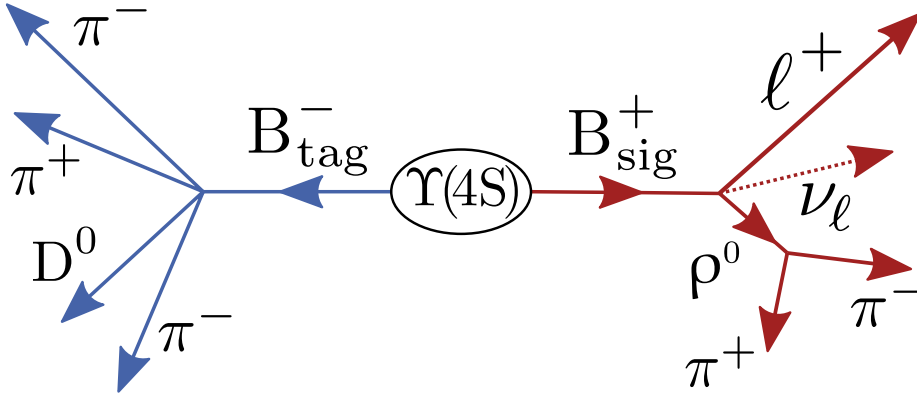


Figure 6.1: Signature of a B_{sig} candidate decaying in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ channel on the right and a representative B_{tag} decay on the left. In the $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ channel, both the ρ meson on the signal side and the D meson on the tag side must be charged instead of neutral.

frame of the $\Upsilon(4S)$ (i.e. 50% of the energy available in the collision).

While these quantities could be directly used in their naive form, two engineered variables, the beam-constrained mass m_{bc} and the energy difference ΔE are preferred which both rely on precise knowledge of $E_{\text{CMS, beam}}$.

The beam-constrained mass is defined as

$$m_{\text{bc}} = \sqrt{E_{\text{CMS, beam}}^2 - \mathbf{p}_{\text{CMS, B}}^2} \quad (6.1)$$

with the beam energy $E_{\text{CMS, beam}}^2$ and the momentum of the B_{tag} candidate $\mathbf{p}_{\text{CMS, B}}^2$, while ΔE is simply

$$\Delta E = E_{\text{CMS, B}} - E_{\text{CMS, beam}} \quad (6.2)$$

As both B mesons are produced almost at rest, these two variables are not very strongly correlated but achieve their advantages over the naive kinematic definitions by exploiting the resolution properties of the detector: ΔE depends strongly on the resolution of the detector calorimeter and the particle hypotheses used to recombine the B candidate from its daughter particles, any wrong combination shifts ΔE away from zero. The complementary variable m_{bc} , on the other hand, is by definition independent of the final-state particle hypotheses. This makes it ideal to distinguish combinatorial backgrounds (such as e.g. those from the hadronization of light-quark pairs) from B candidates reconstructed using the correct detector objects. The latter are expected to produce a peak at the nominal B meson mass while the former are spread out uniformly within the allowed kinematic range. To increase the purity of B_{tag} candidates in the selection, an m_{bc} value above $5.27 \text{ GeV } c^{-2}$ and an energy difference value in the range $-0.15 < \Delta E < 0.1 \text{ GeV}$ are required. With these selections applied, considerable background from light-quark production remains which can be suppressed by exploiting the topological properties of $\Upsilon(4S) \rightarrow B\bar{B}$ decays. A useful quantity to describe aspects of the event topology is thrust. The thrust axis T_C of a collection C of N particle

momenta is here defined as the unit vector to which the projection of these momenta yields the largest value,

$$\mathbf{T}_C = \max_{|\mathbf{T}|=1} \left(\sum_i^N \mathbf{T}_C \cdot \mathbf{p}_i \right). \quad (6.3)$$

The angle T_B, T_O between the thrust axes calculated for the collection of momenta comprising the B_{tag} candidate T_B and for the collection of all other particle momenta in the event T_O is then expected to be distributed uniformly for B decays at rest and to be strongly collimated for decays of lighter, boosted particles. A selection of $\cos(T_B, T_O) < 0.9$ is therefore highly effective in suppressing light-quark backgrounds at a loss of approximately 10% reconstruction efficiency.

6.1.2 Signal-Side Selection

Following the B_{tag} reconstruction, first ρ meson candidates and then B meson candidates are formed by recombining the momenta and masses of final-state and composite particles. I first describe the selections applied to final-state particle candidates, then those applied to composite ones.

To reconstruct the signal process $B \rightarrow \rho \ell \nu_\ell$, one stable, light lepton ($\ell = e^-, \mu^-$) and a charged or neutral ρ meson which decays in the detector is needed. The ρ meson must therefore itself be reconstructed from recombined or stable particles.

Final state particles The particle species which are considered to be stable in the time frame relevant for Belle II are referred to as Final State Particles (FSPs). Candidates for these FSPs are created by assigning hypotheses to the particle trajectories (“tracks”) in the tracking detectors and energy depositions (“clusters”) in the calorimeters. To select particle candidates likely to be electrons and muons from the initial collision, I use the track impact parameters at the point of closest approach (POCA) as defined in Ref. [103] with respect to the measured IP. Distances of at most $|dz| < 5$ cm in z -direction and $dr < 2$ cm in transverse direction select tracks originating close to the IP and reject backgrounds from e.g. beam interactions. To ensure high-quality predictions of the track momenta and vertices and increase the reliability of particle identification algorithms, only tracks with a transverse momentum above 200 MeV and within the CDC acceptance are considered as lepton candidates. Electrons and muons are distinguished from other charged particles using the BDT-based lepton identification variables described in Ref. [62]. The chosen selections retain 97.5% of electrons and 91% of muons originating from $B \rightarrow \rho \ell \nu_\ell$ decays. If multiple lepton candidates are reconstructed in this way, the candidate with the highest value of its respective identification variable is used.

For tracks considered as charged pion candidates, the same requirements on the vertex and polar angle are used. Additionally, the track must have 20 hits in the CDC. To distinguish pions from kaons and light leptons, a selection using the output of the likelihood-based particle identification algorithm [62] is used.

Table 6.1: Momentum regions p_e^{\min} , p_e^{\max} and parameters of the bremsstrahlung correction as determined by optimizing the momentum resolution in simulated data.

$p_e^{\min}/\text{GeV } c^{-1}$	$p_e^{\max}/\text{GeV } c^{-1}$	$\beta/^\circ$	$E_{\gamma,\max} / \text{GeV}$
0.2	0.6	7.84	0.09
0.6	1.0	4.22	0.9
1.0	-	3.62	1.2

Photon candidates suitable to create π^0 meson candidates are created from clusters in the electromagnetic calorimeter, which are required to have energies of at least 80 MeV in the forward endcap of the calorimeter, 30 MeV in its barrel region or 60 MeV in the backward endcap. Additionally, the weighted sum of the number of calorimeter crystals hit by the reconstructed photon is required to be above 1.5 to suppress backgrounds from minimum ionizing particles passing through the calorimeter. To also allow vetoing clusters created by electrons in the ECL, tracking information from the CDC is used. As this information is only reliable for tracks which pass the CDC, only photon candidates from a region of the ECL corresponding to the angular acceptance of the CDC are selected.

Bremsstrahlung correction Electrons produced in Belle II lose energy via bremsstrahlung when interacting with dense structures. To correct for these losses and the resulting lower electron momenta, the electron candidates are corrected by the energy of calorimeter clusters assigned to the electron track. This assignment is based on geometrical reasoning, with calorimeter clusters considered as bremsstrahlung photons if they fall within a maximum angle β from the electron track and pass a maximum energy selection $E_{\gamma,\max}$. The parameters of these selections given in Table 6.1 are determined for three momentum regions independently by minimizing the difference between generated and reconstructed electron momenta. Calorimeter clusters used in the bremsstrahlung correction are not considered as π^0 daughters in the $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ reconstruction.

Recombined particles Two photon candidates passing the criteria given above are recombined to a π^0 candidate, if the invariant mass of the combination is in the range 120 to 145 MeV and the lab-frame angle between the photon satisfies $\cos \phi_{\gamma\gamma}^{\text{lab}} > 0.4$. This second criterion is derived using the kinematics of the signal process and is useful to suppress background from the combination of unrelated calorimeter clusters: Photons originating from the decay of relatively high-momentum π^0 's are expected to be highly collimated with 98% falling into the chosen range.

The π^0 candidates are then combined with charged pion candidates to form ρ^\pm candidates. These are accepted if the combination falls within the range 0.48 to 1.07 $\text{GeV } c^{-2}$, corresponding to an interval of approximately twice the ρ meson width, $\Gamma = 0.149 \text{ GeV } c^{-2}$, around the ρ^+ mass of $m_\rho = 0.775 \text{ GeV } c^{-2}$ [5].

Similarly, ρ^0 candidates are recombined from two pion candidates of opposite charge, here in a window of 0.46 to 1.34 GeV c^{-2} which is consistent with the selection used in Ref. [79]. Additionally, a vertex fit with the TreeFitter algorithm described in Section 4.4 is applied to the ρ^0 candidates. This increases the resolution of the reconstruction vertex by utilizing all available information and allows rejection of combinatorial background candidates where no common vertex exists.

The cms four-momenta of the reconstructed meson and lepton are combined into a pseudoparticle Y . Assuming that only a single neutrino is missing in the reconstruction, the angle between this pseudoparticle Y and the B meson momentum inferred using the beam conditions is given by

$$\cos(\theta_{BY}) = \frac{E_B^* E_Y^* - m_B^2 - m_Y^2}{2|\mathbf{p}_B^*||\mathbf{p}_Y^*|}. \quad (6.4)$$

For correctly reconstructed particle candidates, one can assume $|\cos(\theta_{BY})| < 1.0$ while for misreconstructed particles the incorrect masses and momenta can give values $|\cos(\theta_{BY})| > 1.0$. In the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ reconstruction channel, Y candidates are accepted if they pass this requirement. To account for resolution effects of the detector and to avoid strong dependence of the result on the description of this variable in the simulated sample, this requirement is loosened in the $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ channel, accepting candidates with $-3.0 < \cos(\theta_{BY}) < 1.0$.

A second vertex fit using TreeFitter is performed for B_{sig} candidates in both reconstruction channels and again, candidates with failed vertex fits are rejected. To restrict the number of degrees of freedom in the case of $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ decays and thus allow a vertex fit, the mass of the neutral pions is constrained to $m_{\pi^0} = 135$ MeV from Ref. [5].

I define the four-momentum of the signal B -meson in the cms frame as follows:

$$p_{B_{\text{sig}}} \equiv (E_{B_{\text{sig}}}, \mathbf{p}_{B_{\text{sig}}}) = \left(E_{\text{CMS, beam}}^2, -\mathbf{p}_{B_{\text{tag}}} \right), \quad (6.5)$$

using the energy of the beam in the cms as the energy of the B_{sig} and the recoil of the B_{tag} as the B_{sig} three-momentum. I then define the ‘‘missing’’ four-momentum as

$$p_{\text{miss}} \equiv (E_{\text{miss}}, \mathbf{p}_{\text{miss}}) = p_{B_{\text{sig}}} - p_Y, \quad (6.6)$$

with the square of the missing momentum defined as the squared missing mass $M_{\text{miss}}^2 \equiv p_{\text{miss}}^2$ which, due to resolution effects and misreconstruction, extends to negative values. For correctly reconstructed semileptonic decays such as $B \rightarrow \rho \ell \nu_\ell$, in which at most one massless particle is missed, M_{miss}^2 is expected to peak at zero.

Other determinations of the (squared) missing mass such as the one described in Ref. [66] rescale the B_{tag} momentum using a constraint derived from the beam energy and the known B meson mass ($M_{\text{miss, rescaled}}^2$) or, alternatively, also take the energy from the B_{tag} ($M_{\text{miss, simple}}^2$). With the energy and momentum resolution currently achieved by Belle II, these two alternative approaches, as illustrated in Fig. 6.2, show worse resolution than the definition given in

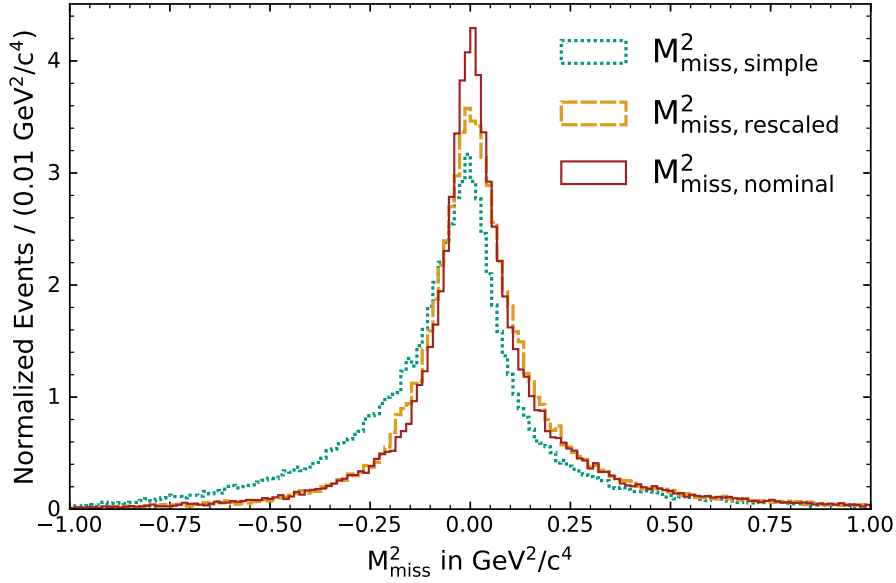


Figure 6.2: Variants of variables to estimate the squared missing mass of the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ signal process in the event, determined from simulation. $M_{\text{miss}, \text{simple}}^2$ uses no beam energy information, $M_{\text{miss}, \text{rescaled}}^2$ rescales the missing momentum to the expected momentum of a B meson and $M_{\text{miss}, \text{nominal}}^2$ follows Eq. (6.5).

Eq. (6.5).

For the presented analysis, a M_{miss}^2 region in the range -1.0 to $2.0 \text{ GeV}^2/c^4$ is considered which retains 97 % of $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays and 92 % of $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ decays.

Additionally, minimum requirements on the missing energy (or equivalently a maximum requirement on the negative energy difference ΔE) of the B_{sig} candidate are imposed. Optimizing for maximum signal retention and background rejection, minimum missing energies of 300 MeV and 100 MeV are required in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ and $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ reconstruction channels, respectively.

At this stage, in approximately 30 % of events more than one B^+ candidate can be reconstructed. To select one of them, a random selection is applied which has been shown to have only negligible impact on the momentum and mass resolution crucial in the signal extraction. In the B^0 reconstruction channel, only 26 % of events contain more than one candidate, however the maximum multiplicity in the data set is much higher due to the contributions from events in which many π^0 candidates can be reconstructed. In this reconstruction channel, a better resolution in momentum and invariant mass compared to random choice can be achieved by selecting the best candidate by the quality of the vertex fit applied to the B^0 candidate.

6.1.3 Selections on the $\Upsilon(4S)$ Candidate

After constructing hypotheses for both B_{tag} and B_{sig} mesons, they are combined to $\Upsilon(4S)$ meson candidates. This allows applying an additional, powerful constraint: Correctly reconstructed candidates should have no additional, well-reconstructed tracks in the event which is referred to as “completeness constraint”. To ensure that this criterion does not reject indiscriminately, well-reconstructed tracks are defined to have been reconstructed inside the CDC with a minimum transverse momentum of 200 MeV and the same impact parameter criteria as for pions and leptons ($dr < 2$ cm; $|dz| < 5$ cm).

To similarly suppress backgrounds from neutral particles in the event, analyses of semileptonic decays at Belle [66] used the sum of energy depositions in the calorimeter to which no tracks and no particle hypotheses in the reconstruction chain are assigned. While this variable, which is often referred to as E_{extra}^n , is similarly useful in Belle II analyses, describing it in simulated data is more challenging. Data taken at Belle II often has, due to the larger instantaneous luminosity, much larger energy depositions from interactions within the particle beam which are challenging to simulate. The presented analysis therefore avoids selections on E_{extra}^n .

6.1.4 Light-Quark Pair Background Suppression

To suppress continuum background, several topological variables have established themselves at Belle II. Most of these exploit the jet-like distribution of final-state particles in light-quark pair events compared to the more uniform distribution expected from $B\bar{B}$ events. This is schematically illustrated in Fig. 6.3. The variables engineered for this purpose can be grouped into three categories: Thrust variables, CLEO cones and (modified) Fox-Wolfram moments. Variables in each of these groups are calculated using both the B_{tag} and the B_{sig} candidate; while these are expected to be highly correlated, reconstruction differences can still provide additional information between both variants. To combine the chosen variables to a single selection quantity, a FastBDT [59] multivariate classifier is trained for each $B \rightarrow \rho\ell\nu_\ell$ reconstruction channel. In the following, each variable group and the multivariate classifier are briefly explained.

Thrust variables In addition to the angle T_B, T_O between the thrust of a B candidate and the thrust of all other momenta in the event already defined above, a second angle T_B, T_z can be defined between the thrust of the B candidate and the beam (z) axis. To distinguish between B decays and continuum background, one exploits conservation of angular momentum in the $\Upsilon(4S) \rightarrow B\bar{B}$ decay: The spin-1 $\Upsilon(4S)$ decays to two spin-0 B mesons, to conserve angular momentum the B mesons must be distributed with $\sin^2\theta_B$ with respect to the beam axis [17, Chapter 9]. However, as B mesons decay almost at rest, T_B, T_z is expected to be distributed uniformly for $B\bar{B}$ events. The $e^+e^- \rightarrow f\bar{f}$ process to two spin-1/2 fermions results in a $1 + \cos^2\theta_B$ distribution with respect to the beam axis, here T_B, T_z follows the distribution

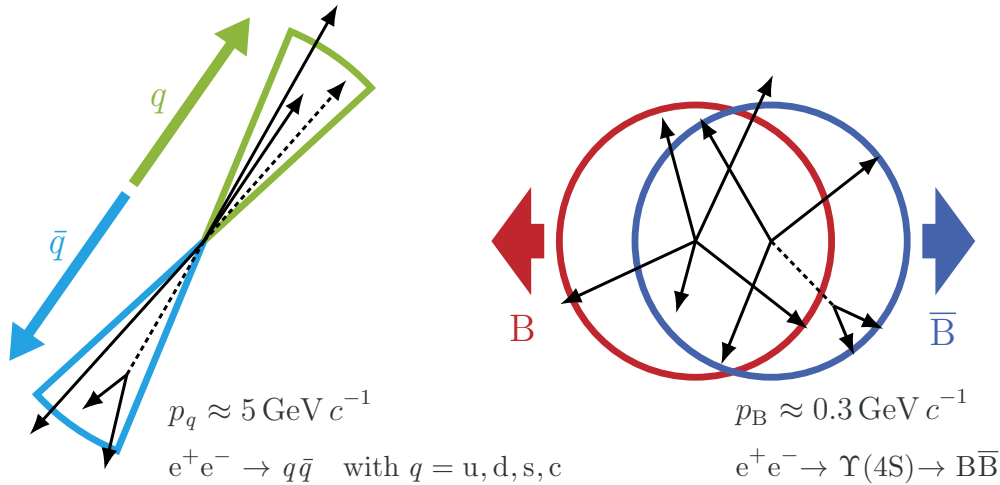


Figure 6.3: Illustration of the topological differences between events with light-quark pairs and events with $B\bar{B}$ production in the cms. Final-state particles from light-quark pairs are highly boosted and produce a jet-like signature. $B\bar{B}$ pairs are produced almost at rest which distributes the final-state particles uniformly in the cms. Adapted from Refs. [84, 104].

due to the high momenta of the fermions from the initial interaction. In addition to the angles between thrust vectors, the magnitudes of the thrust as defined in Eq. (6.3) for a B candidate or for all final-state particles not assigned to a B candidate provide some measure of discrimination between $\Upsilon(4S)$ events and background.

CLEO Cones To exploit differences in the distribution of B candidate’s daughter particles beyond thrust, the CLEO collaboration introduced variables describing the sum of particle momenta in 10° intervals from the thrust axis [105]. As these nine “CLEO cones” use particle momenta not assigned to specific hypotheses in the $\Upsilon(4S)$ reconstruction, they are filtered with selections on detector regions and thresholds on the energy of calorimeter clusters. Particle tracks are selected using the same selections applied for the completeness constraint described above. Calorimeter clusters are required to have energies of at least 100 MeV in the forward endcap of the calorimeter, 90 MeV in its barrel region or 160 MeV in the backward endcap to be considered as π^0 daughters. Similar to π^0 daughter candidates, only clusters within the acceptance of the CDC are considered. As not all CLEO cone variables show good agreement between simulated and recorded data, only those listed in Table 6.2 are considered in this analysis.

Fox-Wolfram moments A more theory-motivated approach to rotationally invariant descriptions of the event shape is given by Fox-Wolfram moments [106]. For a collection of N particles with momenta \mathbf{p}_i , angle θ_{ij} between each pair of them and the Legendre polynomials

P_k , the k -th Fox-Wolfram moment H_k is given by

$$H_k = \sum_{i,j}^N \frac{|\mathbf{p}_i||\mathbf{p}_j|}{s} P_k(\cos \theta_{ij}) \quad (6.7)$$

with the cms energy \sqrt{s} .

In practice, the ratio $R_k = H_k/H_0$ is used to discriminate $B\bar{B}$ and $q\bar{q}$ events, a common choice is R_2 . In general, it holds that for even k , $q\bar{q}$ events produce a peak at one, for odd k a peak at zero. Noticeable disagreement between data and current Belle II $B\bar{B}$ and $q\bar{q}$ simulation is observed in the R_2 variable which is attributed either to deficiencies in the $q\bar{q}$ simulation described in Section 5.3 or beam background effects. The R_2 variable is therefore not used in the continuum suppression classifier.

Instead, modified Fox-Wolfram moments, also referred to as Kakuno Super Fox–Wolfram (KSFW) variables and first introduced at Belle [17, Chapter 9] are used. These exhibit better data-MC agreement in the subset given in Table 6.2. For completeness, all of these variables, which were originally used to construct a Fisher discriminant [107] for continuum suppression, are described here.

The modified Fox-Wolfram moments are composed of two subsets: The H_k^{oo} variable subset using only information from the ROE of the B candidate and the H_{pk}^{so} variable subset which combines information from both final-state particles used to construct the B candidate and the ROE.

The H_k^{oo} variables are defined as

$$H_k^{\text{oo}} = \sum_{ij} Q_i Q_j |\mathbf{p}_i||\mathbf{p}_j| P_k(\cos \theta_{ij}) \quad (6.8)$$

which runs over the particle pairs (i, j) in the ROE with the momenta p , the angle between them θ_{ij} and the Legendre polynomials P_k . For odd k , the $Q_{i,j}$ are the particle charges for the particles i and j , for even k the product $Q_i Q_j$ is +1. This makes the k even H_k^{oo} variables equivalent to Fox-Wolfram moments calculated using only the ROE of a candidate.

The H_{pk}^{so} variable subset combines more information into the definition: The p subscript identifies whether the variable is calculated using only charged particles ($p = 0$), neutral particles ($p = 1$) or a pseudoparticle representing the sum of missing particle momenta ($p = 2$) which is calculated using beam-energy constraints.

H_{pk}^{so} is given by

$$H_{pk}^{\text{so}} = \sum_{ij} Q_i Q_j |\mathbf{p}_j| P_k(\cos \theta_{ij}) \quad (6.9)$$

with j running over the momenta of ROE particles of the group indicated by p and i running over the daughters of the B candidate (which are not necessarily final-state particles). $Q_{i,j}$ are again the particle charges for odd k and $p = 0$, for even k the product $Q_i Q_j$ is +1 and for $p = 1, 2$ it is zero.

To normalize the KSFV variables such that they are independent of the energy difference ΔE , the H_{pk}^{SO} are scaled by H_{max} and the H_l^{OO} variables are scaled by H_{max}^2 with

$$H_{\text{max}} = \frac{1}{-2\Delta E} \quad (6.10)$$

where ΔE is calculated using the B meson candidate.

BDT classifier Using only well-modelled input features as given in Table 6.2, two SGBDT classifiers from the FastBDT package [59] with 200 decision trees and otherwise default hyperparameter settings are trained to distinguish signal events from $q\bar{q}$ events. The independent data sets used to train the multivariate analysis (MVA) classifier consist of simulated continuum data equivalent to 500 fb^{-1} and signal MC containing approximately 1.9×10^6 events. Both samples are weighted to account for the unequal number of events such that selections on the classifier output approximate the probability of selecting a signal event in a sample only containing signal and continuum events. The performance of the classifiers as evaluated on an independent data set is illustrated in Fig. 6.4. A classifier output \mathcal{P}_{CS} above 0.6 is required for events to pass the selection in both $B \rightarrow \rho l \nu_l$ reconstruction channels. The agreement between simulated and recorded data for the classifier output in the sideband is illustrated in Fig. 6.4. This figure shows that the simulation without systematic uncertainties cannot completely reproduce the data in \mathcal{P}_{CS} but illustrates that the chosen working point is not expected to influence the signal efficiency significantly.

Table 6.2: Variables used to train the Continuum Suppression MVA method. A check mark indicates that the variable is derived from the corresponding candidate, a cross mark indicates that it is only used for the other B meson candidate in the event.

Variable index	B _{tag}	B _{sig}
Thrust variables		
T_B, T_O	✓	✓
T_B, T_z	✓	✓
CLEO cones		
Cone 3	✓	✓
Cone 4	✓	✓
Cone 5	✓	✓
Cone 6	×	✓
Cone 7	×	✓
Cone 8	×	✓
Cone 9	✓	×
KSFV variables		
H_4^{oo}	✓	✓
H_{02}^{so}	✓	×
H_{04}^{so}	✓	✓
H_{12}^{so}	×	✓
H_{14}^{so}	✓	✓
H_{22}^{so}	×	✓
H_{24}^{so}	×	✓

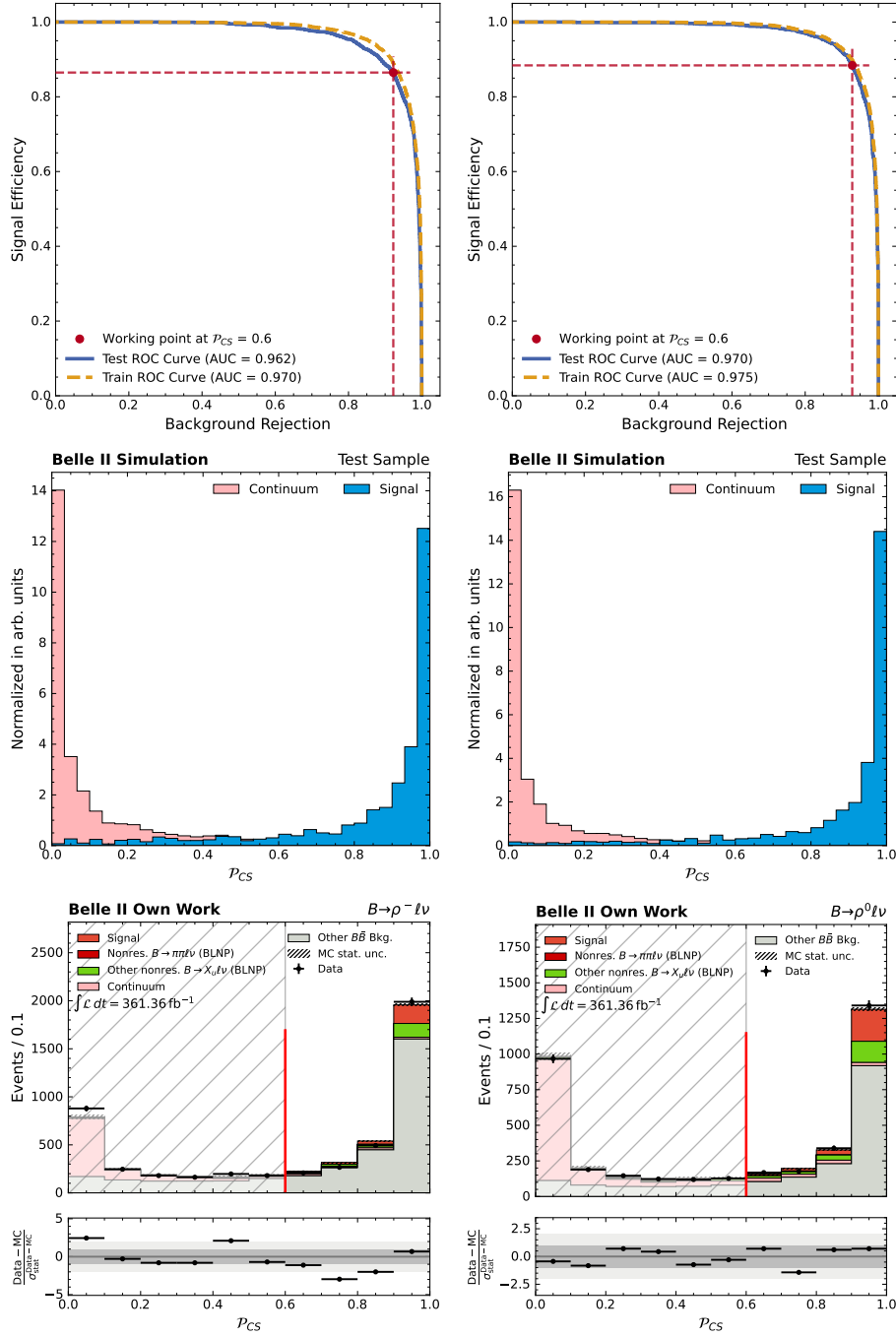


Figure 6.4: Properties of the MVA method used to reject continuum backgrounds in the (left) $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ and (right) $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ reconstruction. Shown are (top) the receiver operating characteristic (ROC) curve for test and training samples together with the AUC value, (middle) the separation of continuum background and $B \rightarrow \rho \ell \nu$ signal in the test sample and (bottom) the agreement between Data and MC distributions. The hatched region in the bottom corresponds to rejected data, with the red line indicating the selection of $\mathcal{P}_{CS} > 0.6$.

6.2 Monte Carlo Corrections for Selection Effects

In addition to the corrections of the simulated samples due to changes of the underlying physics model or updates of measurement inputs described in Chapter 5, several other corrections originating in the reconstruction and selection procedure must be applied. These corrections compensate for different performance of the complex methods used to identify particle species and reconstruct π^0 and B_{tag} candidates in simulated and recorded data.

6.2.1 Correction of Particle Reconstruction Efficiencies and Misidentification Rates

Both the BDT-based lepton identification algorithm as well as the likelihood-based algorithm to identify charged pions exhibit different efficiencies and misidentification rates between recorded and simulated data. These differences depend on the momentum and polar angle of the associated tracks and are applied as weights to simulated particle candidates based on the momentum and angle of their final-state constituents. Efficiency factors are applied where correctly-identified particle candidates are used in the $\Upsilon(4S)$ reconstruction. Factors for misidentification rates, on the other hand, are only applied to simulated candidates in which final-state particles are misidentified.

To provide efficiency correction factors for the lepton identification algorithm which span a wide range of particle momenta p_{lab} and polar angles θ , four analyses of well-known physical processes ($J/\psi \rightarrow \bar{\ell}\ell$, $e^+e^- \rightarrow \bar{\ell}\ell(e^+e^-)$, $e^+e^- \rightarrow e^+e^-(\gamma)$ and $e^+e^- \rightarrow \mu^+\mu^-\gamma$) are combined [108]. Each of these processes dominates in certain regions of the $p_{\text{lab}} - \theta$ parameter space such that combining them minimizes uncertainties and maximizes coverage. A tag and probe method is used to determine the efficiency in data in each analysis with the “tag” sample selected using the invariant mass of the two lepton candidate tracks and event shape variables. The “probe” sample is then a subset of the tag sample in which a lepton identification requirement is added. Using the number of events $N_{\text{pass}}^{\text{sig}}$ which pass the lepton identification requirement and the number of events $N_{\text{fail}}^{\text{sig}}$ which do not, the efficiency ϵ is given by

$$\epsilon = \frac{N_{\text{pass}}^{\text{sig}}}{N_{\text{pass}}^{\text{sig}} + N_{\text{fail}}^{\text{sig}}}. \quad (6.11)$$

The calibration factors determined with this approach are close to unity, approximately 91% of candidates reconstructed in both $B \rightarrow \rho\ell\nu_\ell$ channels have calibration factors in the interval $[0.95, 1.05]$.

In addition to corrections to the efficiency, the simulated sample must also be corrected for the rates of charged pions and kaons misidentified as leptons. These correction factors are typically further away from unity than those necessary for efficiency differences with only approximately 23% (27%) of $B^+ \rightarrow \rho^0\ell^+\nu_\ell$ ($B^0 \rightarrow \rho^-\ell^+\nu_\ell$) correction factors for misidentified pions in the interval $[0.95, 1.05]$. However, as only roughly 5% of leptons in both reconstruction channels are misidentified, the overall impact of this correction is much smaller. The number

of pions misidentified as leptons is estimated using the decay of neutral kaons to pions and the decay of tau leptons to one or three charged pions, again using a tag and probe method. To identify the number of charged kaons misidentified as leptons, the decay $D^* \rightarrow D^0 (\rightarrow K^- \pi^+) \pi^+$ is used.

The same physical processes used to determine efficiency and misidentification rate corrections for lepton candidates are also used to correct efficiency and misidentification rates for the pion identification algorithm. Only a weak selection on the hadron identification variable is used in the $B \rightarrow \rho \ell \nu_\ell$ reconstruction so the impact of these corrections (which are applied twice in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ reconstruction channel to account for the two charged pions) is small compared to those applied for the lepton identification.

For all particle identification corrections, statistical uncertainties (which are assumed to be uncorrelated between $p_{\text{lab}} - \theta$ bins) and systematic uncertainties (which are assumed to be fully correlated between $p_{\text{lab}} - \theta$ bins) are provided.

While no particle identification algorithms are used to select the π^0 candidate in the $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ recombination, the less well-defined nature of photon candidates reconstructed from energy depositions in the calorimeter still requires corrections for the π^0 reconstruction efficiency. These are determined from $D^0 \rightarrow K^- \pi^+ \pi^0$ decays by comparing the D^0 mass peak in data and simulation and from $\tau \rightarrow 3\pi^\pm \pi^0 \nu_\tau$ decays by doing the same for the π^0 mass peak. The correction factors applied to simulated data originate only from the D^0 determination, with an additional systematic uncertainty assigned from the difference between both methods. This systematic uncertainty is assumed to be uncorrelated to the systematic uncertainties estimated in the D^0 method. In addition to a correction on the π^0 reconstruction efficiency, a second calibration is needed to correct for a discrepancy in the calorimeter cluster energy between simulated and recorded data [109]. This effect systematically shifts the reconstructed mass of π^0 mesons, it is corrected for with a calibration factor derived in an analysis of $B \rightarrow \bar{D}^0 (\rightarrow K^+ \pi^- \pi^0) \pi^+$ processes.

6.2.2 Tagging Efficiency Correction

The FEI tagging method described in Section 4.5 exhibits different performance in recorded and simulated data which must be corrected for with calibration factors

$$\epsilon_{\text{FEI},i} = \frac{N_{\text{MC},i}}{N_{\text{Data},i}} \quad (6.12)$$

for FEI reconstruction channels or combinations of reconstruction channels i .

To estimate the size of the factors $\epsilon_{\text{FEI},i}$ for different requirements on the FEI classifier output, two independent analyses with different methods are conducted, one using an inclusive reconstruction of semileptonic decays and one using a partial reconstruction approach in $B \rightarrow D^{(*)} \pi$ decays. Then, a χ^2 parameter estimation is performed to combine values from both analyses.

Semileptonic calibration method In the calibration analysis using semileptonic decays described in Ref. [110], a B_{tag} candidate selected according to the description given in Section 6.1.1 is combined with a well-identified lepton with high momentum ($p_\ell^* > 1.0 \text{ GeV } c^{-1}$) in the B rest frame. Backgrounds arise here from hadronic B decays, in which a hadron decays to a lepton or is misidentified as one, as well as $q\bar{q}$ processes. A template fit similar to the one described in Section 7.1.3 but using the lepton momentum in the B rest frame is performed to determine the number of signal and background events in data. Shape-changing systematic uncertainties are incorporated also via bin-wise nuisance parameters. While a single template is used to describe all background contributions, the signal template is split into four sub-templates in which the hadron can be a ground-state D meson, a D^* meson, a D^{**} meson, or a hadron containing a u quark. The fraction with which each of these contribute to the signal yield is constrained to the value determined from simulation but within systematic uncertainties arising from their respective branching fractions. These uncertainties from $B \rightarrow X\ell\nu_\ell$ branching fractions also dominate the overall relative systematic uncertainties of the study, presenting a relative uncertainty of 2.7% on the calibration factor.

Hadronic calibration method In the second calibration analysis using hadronic B decays described in Ref. [111], the B_{tag} candidate is combined with the pion candidate with the highest momentum in the event. Then, the recoil momentum of this combination, which for $B \rightarrow D^{(*)}\pi^\pm$ decays should have a peak at the masses of D and D^* mesons, is used to estimate the number of these decays. The maximum likelihood fit here uses Crystal Ball functions [112] combined with a normal distribution to describe the two signal peaks with the mean parameter of the D peak left floating and the mean of the D^* peak shifted by the known mass difference between the two mesons. Backgrounds arise from partially reconstructed $B \rightarrow D^{(*)}\rho$ decays, misreconstructed $B\bar{B}$ decays and $q\bar{q}$ processes. While the latter two are described using a common exponential function with the slope parameter of the fit left free, a Gaussian distribution with all shape parameters determined from simulation is used to describe the peaking $B \rightarrow D^{(*)}\rho$ backgrounds. The systematic uncertainties on the calibration factors derived with this method are again dominated by branching fraction uncertainties, in this case those of the $B \rightarrow D\pi^\pm$ process. This source of uncertainty dominates the overall systematic uncertainties of 2.5% for charged and 3.5% for neutral B_{tag} calibration factors by contributing 1.9% and 3.1% uncertainties, respectively.

Combination To combine the two calibration factors determined for each B_{tag} reconstruction channel subset, a χ^2 fit is used which is described in detail in Ref. [113]. In this fit, systematic uncertainties common to all FEI channels but unique to each calibration approach are treated as nuisance parameters constrained to the values determined in each separate analysis. Additional nuisance parameters are introduced for discrepancies between the two calibration methods in some FEI channels which cannot yet be attributed to specific causes.

The nuisance parameters are constrained to

$$\sigma_{\text{discrep}} = \sqrt{(\epsilon_{X\ell} - \epsilon_{D\pi})^2 - \sigma_{X\ell}^2 - \sigma_{D\pi}^2}, \quad (6.13)$$

with calibration factors $\epsilon_{X\ell}$ and $\epsilon_{D\pi}$ as well as their respective uncertainties $\sigma_{X\ell}$ and $\sigma_{D\pi}$. The calibration factors from the two independent analyses as well as the factor combinations are illustrated in Fig. 6.5.

In addition to combined, FEI-channel dependent central values and uncertainties for the calibration factor, this approach also provides a correlation matrix between the factors which is used in the signal extraction of the $B \rightarrow \rho\ell\nu_\ell$ analysis as detailed in Section 8.2.2.

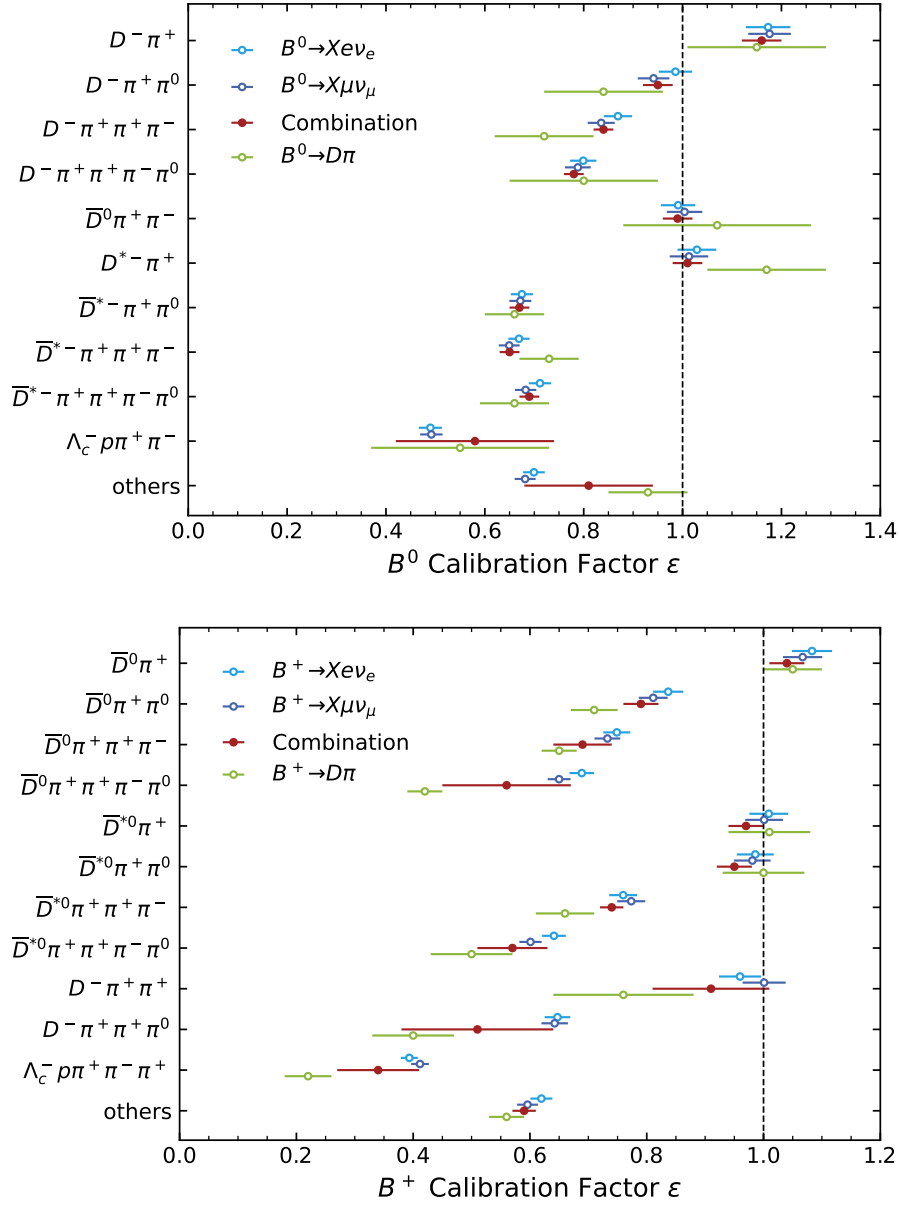


Figure 6.5: Channel-dependent calibration factors [114, 110, 113] for (top) neutral and bottom (charged) B_{tag} candidates to correct for efficiency differences of the FEI in recorded and simulated data. Shown are factors derived in two independent analyses as well as a combination of them which takes into account correlated uncertainties and differences between the two calibration approaches. Empty markers correspond to the individual measurements while solid markers give the combined factors. The dashed line indicates unity with which calibration factors of several well-modelled FEI reconstruction channels are consistent.

6.3 Resulting Simulated and Recorded Samples

The reconstruction efficiencies and expected number of events in the $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ reconstruction channel are given in Table 6.3, for the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ reconstruction channel these values are given in Table 6.4.

Figure 6.6 shows the signal-extraction distributions in recorded and simulated data. A distinct overestimation in simulated data is visible in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ reconstruction channel for $0.8 \text{ GeV } c^{-2} < M(\pi\pi) < 0.88 \text{ GeV } c^{-2}$. This effect is attributed to interference between ρ and ω mesons which is not simulated and for which a systematic uncertainty is introduced in Section 8.2.1. Figure 6.6 also allow defining a signal region for both $B \rightarrow \rho \ell \nu_\ell$ channels with $|M_{\text{miss}}^2| < 1.0 \text{ GeV}^2/c^4$ and a sideband region with $1.0 \text{ GeV}^2/c^4 < M_{\text{miss}}^2 < 2.0 \text{ GeV}^2/c^4$ in which no $B \rightarrow \rho \ell \nu_\ell$ decays are expected.

Table 6.3: Expected number of events and reconstruction efficiencies in the reconstructed $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ sample. The reconstruction efficiencies are given with uncertainties purely originating from the size of the MC sample.

Template	Expected N_{Events}	Rec. Efficiency $\times 10^3$
$B^0 \rightarrow \rho^- \ell^+ \nu_\ell$	272.8	1.26 ± 0.01
Nonresonant. $B^0 \rightarrow \pi^- \pi^0 \ell \nu_\ell$	105.6	0.32 ± 0.01
Other Backgrounds	2789.3	
Continuum	77.1	

Table 6.4: Expected number of events and reconstruction efficiencies in the reconstructed $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ sample. The reconstruction efficiencies are given with uncertainties purely originating from the size of the MC sample.

Template	Expected N_{Events}	Rec. Efficiency $\times 10^3$
$B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$	322.4	2.61 ± 0.01
Nonresonant. $B^+ \rightarrow \pi^\pm \pi^\mp \ell \nu_\ell$	24.6	0.83 ± 0.01
Other Backgrounds	1651.3	
Continuum	101.1	
$B \rightarrow f_0(500) \ell \nu_\ell$	7.4	1.22 ± 0.02
$B \rightarrow f_0(980) \ell \nu_\ell$	8.4	1.38 ± 0.02
$B \rightarrow f_2(1270) \ell \nu_\ell$	12.4	1.02 ± 0.01
$B \rightarrow \rho^0(1450) \ell \nu_\ell$	2.3	0.37 ± 0.01

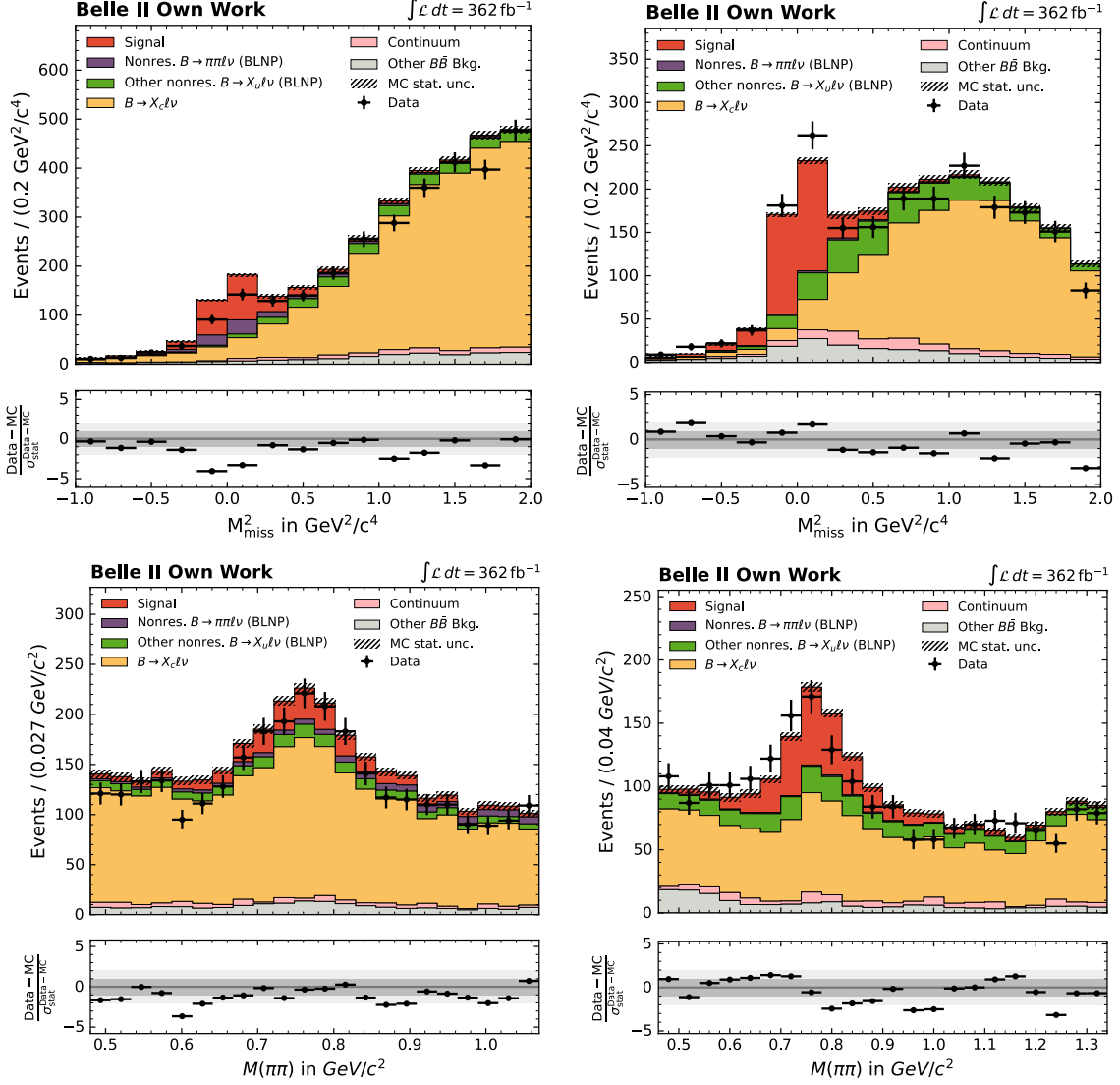


Figure 6.6: Signal-extraction variables (top) M_{miss}^2 and (bottom) $M(\pi\pi)$ for the (left) $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ and (right) $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ reconstruction channels in recorded and simulated data with all normalizations set to the expectation in simulation as described above. The $B \rightarrow X_c \ell \nu_\ell$ component exhibits distinctly peaking behavior at the nominal ρ mass in $M(\pi\pi)$ as ρ mesons are reconstructed in background events. While the simulated data can reproduce the necessary structures in data, the comparison also indicates that substantial systematic uncertainties must be assigned to the simulated sample to e.g. account for the overestimation above the ρ mass in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ channel. The comparison in the $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ channel also shows that either the signal or the nonresonant $B \rightarrow \pi\pi \ell \nu_\ell$ component are overestimated.

6.4 The $B \rightarrow D(\rightarrow \pi^\pm \pi^\mp) \ell \nu_\ell$ Control Channel

To test whether the FEI calibration factors can be applied to low-multiplicity B decay channels and whether the π^\pm momenta are reproduced accurately in the simulated sample, the $B \rightarrow D(\rightarrow \pi^\pm \pi^\mp) \ell \nu_\ell$ process is used. While semileptonic decays to D mesons (possibly via excited D^* states) have a high ($\approx 9.6\%$) branching fraction, the D^0 decay to two charged pions has been measured only with a branching fraction of $(1.45 \pm 0.02) \times 10^{-3}$ such that the overall branching fraction of this process is comparable to the branching fraction expected for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$. Due to the much smaller D width, reconstructing it is however much more straightforward than reconstructing a ρ meson. Therefore, it presents a good opportunity for crosschecks. The same selection as in the signal process is used in this test, the only selections which are removed are on the invariant mass of the pion pair and the continuum suppression classifier output. The agreement between recorded and simulated data is excellent as illustrated in Fig. 6.7.

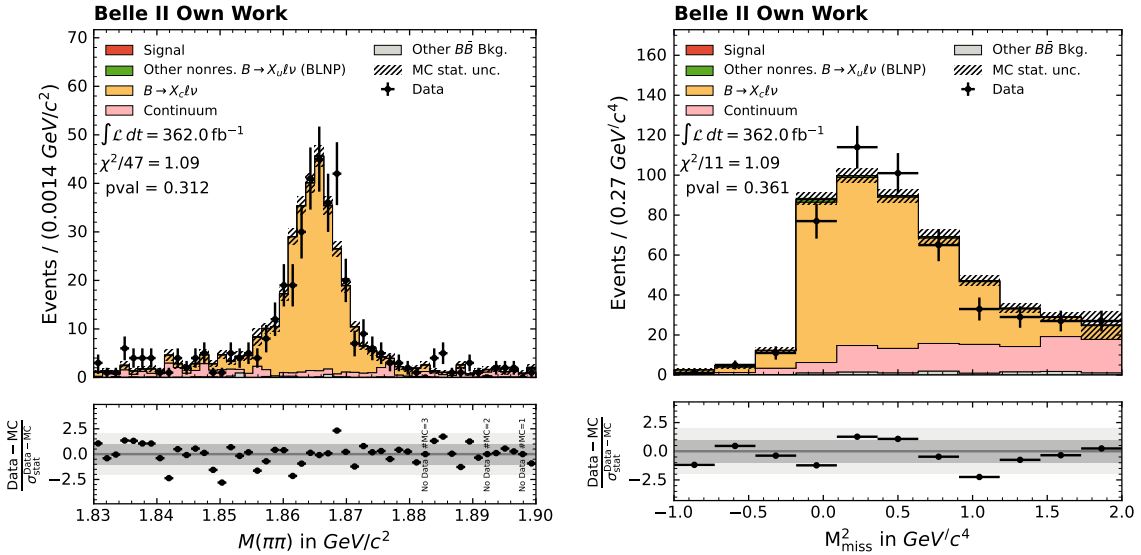


Figure 6.7: Agreement between recorded and simulated data for $B \rightarrow D(\rightarrow \pi^\pm \pi^\mp) \ell \nu_\ell$ decays in very fine bins of (left) $M(\pi\pi)$ and (right) M_{miss}^2 . For $M(\pi\pi)$, an additional selection with $|M_{\text{miss}}^2| < 1.0 \text{ GeV}^2/c^4$ is applied to ensure only well-reconstructed decays.

Chapter 7

Signal Extraction Strategy

The measured events, which remain after the reconstruction and selection steps described in the previous chapter, are in a relation with the values of interest, which we would like to measure: The total branching fractions of the processes $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ and $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$. In the following chapter, an approach to estimate the number of signal events in the data sample, from which the branching fraction can be determined, is presented.

7.1 The Method of Maximum Likelihood

The following is a brief introduction to the maximum likelihood technique for parameter estimation, more detailed descriptions can be found e.g. in Ref. [115, Chapter 6] or Ref. [116].

The method of maximum likelihood provides a way to estimate one or more true parameters $\boldsymbol{\theta}$ using a composite hypothesis $f(x_i; \boldsymbol{\theta})$ of a PDF and n measurements x_i .

The core concept in this approach is the likelihood function

$$\mathcal{L}(\boldsymbol{\theta}) = \prod_{i=1}^n f(x_i; \boldsymbol{\theta}), \quad (7.1)$$

which simply gives the combined probabilities to measure values x_i given the PDF f and the independent parameter values $\boldsymbol{\theta}$. Maximizing this overall probability by varying the parameters of the PDF is therefore a straightforward approach to obtain an estimate $\hat{\boldsymbol{\theta}}$ of the true parameter $\boldsymbol{\theta}$. For most PDFs, analytic descriptions of the likelihood maxima are not available and numerical methods are used to find the largest local maximum for each likelihood. To simplify this evaluation, instead of maximizing the likelihood, the negative logarithm of the likelihood (the “negative log-likelihood”) is minimized. Due to the monotonous nature of the logarithm this maintains the minima but reduces the product of $f(x_i, \boldsymbol{\theta})$ to a sum which is advantageous for iterative minimization algorithms.

The extended Maximum Likelihood method In the form described above, the size of the measured data sample N is assumed to be equal to the expectation for the total number

of events ν . Often, this is however also subject to statistical fluctuations, which one wants to incorporate in the likelihood function. This is achieved by extending the likelihood with a Poisson distribution with mean ν for the number of measurements

$$\mathcal{L}(\boldsymbol{\theta}) = e^{-\nu} \frac{\nu^N}{N!} \prod_{i=1}^n f(x_i; \boldsymbol{\theta}). \quad (7.2)$$

Assuming ν to be independent of the other parameters $\boldsymbol{\theta}$, this yields the same estimators $\hat{\boldsymbol{\theta}}$ as the likelihood without this extension. Even in this case the extended maximum likelihood is useful to treat problems in which the parameters θ_i sum to unity (such as when estimating a branching fraction) as the formalism directly gives estimates for the number of signal and background events without redefining one of the parameters θ_n as a function of all other parameters $\theta_{i \neq n}$. Instead, one defines the parameters $\mu_i = \nu \theta_i$ as the number of expected events of type i and redefines the PDF $f(x_i; \boldsymbol{\theta})$ as a sum of signal and background contributions. For the simple case of a single signal and a single background component, the PDF is then given by

$$f(x_i; \boldsymbol{\mu}) = \frac{\mu_s}{\mu_s + \mu_b} f_s(x_i) + \frac{\mu_b}{\mu_s + \mu_b} f_b(x_i), \quad (7.3)$$

with the expected number of events defined as $\nu = \mu_s + \mu_b$.

The binned Maximum Likelihood The method of maximum likelihood does not require the data to be binned, i.e. discretized over a variable of interest. However, the advantages of unbinned fits shrink quickly with larger data sets while the required evaluation time (an important property for iterative minimization) scales often linearly with each available sample. As large data sets are a common occurrence in high-energy particle physics, data and PDFs are often histogrammed.

The likelihood in this binned case for a fixed number of measurement N and a number of bins B is then defined as

$$\mathcal{L}(\boldsymbol{\theta}) = N! \prod_{i=1}^B \frac{P_i(\boldsymbol{\theta})^{n_i}}{n_i!} \quad (7.4)$$

with n_i the number of measurements in bin i and the integral over each bin

$$P_i(\boldsymbol{\theta}) = \int_{x_i^{\text{low}}}^{x_i^{\text{high}}} f(x; \boldsymbol{\theta}) dx \quad (7.5)$$

and the lower and upper bin edges x_i^{low} and x_i^{high} .

7.1.1 Parameter Estimation with Templates

In many particle physics analyses, the goal is not only to estimate e.g. the normalization parameter of a single signal process but also to estimate contributions from background processes. To obtain these estimates, the combined PDF $f(x, \boldsymbol{\theta})$ is defined as the sum of

individual PDFs $f_i(x)$ which are already known but have unknown normalization parameters θ_i :

$$f(x, \boldsymbol{\theta}) = \sum_{i=1}^m \theta_i f_i(x), \quad (7.6)$$

As often no analytical description of these processes (which encompass complex detector interactions) exist, one must resort to approximations. These binned approximations (“templates”) are derived using MC simulation and include effects from primary physical processes (e.g. particle decays), the finite resolution of the detector and non-linear analysis techniques such as multi-variate methods.

As the templates are available in discrete form already, the binned likelihood is

$$\mathcal{L} = \prod_i^{\text{bins}} \mathcal{P}(n_i | \nu_i(\boldsymbol{\theta})), \quad (7.7)$$

with the Poisson distribution \mathcal{P} and n_i (ν_i) the number of measured (expected) events in bin i .

In this case, the number of expected events ν_i in a bin corresponds to the sum of contributions from all templates in this bin:

$$\nu_i = \sum_k^{\text{templates}} f_{ik} \eta_k \quad (7.8)$$

with η_k as the total number of events from template k and f_{ik} the fraction of events of bin i in template k .

Parameter Limits While ideally, all parameters θ_i in the likelihood can vary from negative to positive infinity to determine a best-fit value, not all parameter values have straightforward physical interpretation. To limit parameters from reaching undesired values, a parameter transformation,

$$\theta \rightarrow \theta' = \arcsin \left(2 \frac{\theta - a}{b - a} - 1 \right), \quad (7.9)$$

is performed such that the new parameter θ' can take on any value while the parameter θ is constrained between the lower upper limits a and b .

7.1.2 Uncertainties Originating from Template Distributions

The normalized templates derived using simulation are subject to uncertainties originating from, among others, the limited size of the simulated sample, reconstruction effects or unknown properties of the underlying physical processes. These additive uncertainties are frequently evaluated separately by e.g. varying each unknown parameter within its uncertainties, deriving new templates and estimating the parameter again. This method of separate treatment has

its advantages but does not optimally exploit the available information: The distribution of measured events contains information on these parameters which can be used to constrain them in the parameter estimation itself.

To do this, the template likelihood is extended by adding a nuisance parameter θ_{ik} for each bin i and template k of the fit to the likelihood and modifying the template shape according to the nuisance parameter value. Here, a nuisance parameter is one for which a prior exists which can be used to constrain it to some value. For simplicity, this prior is assumed to be normally distributed, an assumption which is valid in the large sample limit.

This extension of the likelihood function is achieved by multiplying it with a multivariate Gaussian distribution \mathcal{N} for each template k :

$$\mathcal{L} = \prod_i^{\text{bins}} \mathcal{P}(n_i | \nu_i(\theta)) \times \prod_k^{\text{templates}} \mathcal{N}(\theta_k | 0, \Sigma_k) \quad (7.10)$$

where Σ_k is the correlation matrix given by the systematic uncertainties and template statistics. This penalty term increases the negative log-likelihood, when the parameters θ_k deviate from zero.

The corresponding modification of the template shape is achieved by incorporating θ_{ik} in f_{ik} :

$$f_{ik} = \frac{\eta_{ik}(1 + \theta_{ik}\epsilon_{ik})}{\sum_j^{\text{bins}} \eta_{jk}(1 + \theta_{jk}\epsilon_{jk})}, \quad (7.11)$$

where η_{ik} and ϵ_{ik} are the number of events and the uncertainty in bin i and template k .

Incorporating the nuisance parameters into the template in this manner ensures that deviations from zero in the nuisance parameters in each bin do not influence the overall normalization. This allows disentangling additive, i.e. shape-changing, and multiplicative, i.e. efficiency-changing, influences of each uncertainty source.

A disadvantage of this approach to introducing systematic uncertainties in the fit is that it offers no direct mapping of the individual sources of uncertainties to the nuisance parameters which means that deviations from the expectation for a nuisance parameter cannot be attributed to a single source. The correlation matrix Σ_k accounts for all (assumed to be uncorrelated) sources of systematic uncertainty which cannot be decoupled trivially.

The nuisance parameter method of including additional information on partially-known parameters in the likelihood model is not limited to the template shape: In principle, arbitrary combinations of parameters in the fit can be treated as nuisance parameters in the same way.

7.1.3 Uncertainty Estimation using the Likelihood

To determine the uncertainty on a parameter of interest p once an estimate \hat{p} on the value of this parameter has been obtained from the optimum of the constructed likelihood \mathcal{L} , two options present themselves:

In the large sample limit, the uncertainty on p can be determined using the second

derivative of the log-likelihood at $p = \hat{p}$,

$$\sigma(\hat{p}) = \sqrt{\frac{d^2 \log(\mathcal{L}(p))}{d^2 p}}^{-1}, \quad (7.12)$$

as derived in Ref. [117] using the Rao-Cramér-Fréchet inequality for a sufficient, efficient statistic. For multiple parameters p, p' at $p = \hat{p}, p' = \hat{p}'$ this becomes [17]

$$\sigma(\hat{p}) = \sqrt{\frac{\partial^2 \log(\mathcal{L}(p, p'))}{\partial p \partial p'}}^{-1}, \quad (7.13)$$

Alternatively, one can use Wilks' theorem [118] to estimate the uncertainty on p using the interval given by scanning the log-likelihood ratio

$$\lambda(p) = \log \frac{\mathcal{L}(p)}{\mathcal{L}(\hat{p})} \quad (7.14)$$

from -0.5 to 0.5, i.e. by varying p until these two values are reached.

In the presence of nuisance parameters θ , this becomes the profile likelihood ratio,

$$\lambda(p) = \log \frac{\mathcal{L}(p, \hat{\theta})}{\mathcal{L}(\hat{p}, \hat{\theta})}, \quad (7.15)$$

where the scanned parameter p is varied from its best-fit values \hat{p} and $\hat{\theta}$. The best-fit values $\hat{\theta}$ for all nuisances must here be re-determined for each scanned p value.

The two approaches to estimate the uncertainty are equivalent in the large sample limit. Given limited statistics, the likelihood-ratio approach (which in this case can give asymmetric uncertainties) is preferred.

Likelihood minimization Determining the global minimum of the likelihood function is usually not possible, certainly not using the likelihood construct described above. Instead, heuristic methods are used to determine the smallest local maximum. For the presented work, the MIGRAD algorithm [119] which is part of the `iminuit` [120] software package has been chosen to minimize the negative log-likelihood. This algorithm uses estimates of the gradient calculated during the minimization procedure, which allows it to reach minima faster but requires the likelihood function to be differentiable everywhere.

In addition to MIGRAD, the `iminuit` package also contains the algorithms HESSE and MINOS, implementations of the two approaches to estimate the parameter uncertainty outlined above.

7.2 Implementation of the Fit

Using this statistical framework, two independent signal extraction procedures are designed for $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$. The fit models are implemented using an adapted version of the `TemplateFitter` Python package [121] which constructs the likelihood function as described, minimizes it using `iminuit` and creates figures illustrating the result.

Motivated by experiences using a previous data set [122] in which significant nonresonant $B \rightarrow \pi\pi\ell\nu_\ell$ background to the $B \rightarrow \rho\ell\nu_\ell$ process was suspected, I use two variables to extract the signal yield. These are the missing mass M_{miss}^2 , which presents a measure of reconstruction quality, and the dipion invariant mass $M(\pi\pi)$, which allows distinguishing between nonresonant and resonant $B \rightarrow \pi\pi\ell\nu_\ell$ processes.

7.2.1 Fit Model in the $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ Analysis

In the $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ signal extraction, I construct four two-dimensional templates describing distributions in M_{miss}^2 and $M(\pi\pi)$, each corresponding to a category of decays in the reconstructed sample, as described below.

Signal template The $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ signal template is constructed using a large sample of events simulated according to the prescriptions given in Section 5.1. All events in which a $\Upsilon(4S)$ candidate can be reconstructed are used to describe this PDF, independent of the correct identification of tracks or clusters in the reconstruction. This has been shown to decrease the expected statistical uncertainty slightly as misreconstructed events still produce peaking structures in M_{miss}^2 and $M(\pi\pi)$.

Nonresonant background template A template describing $B^0 \rightarrow \pi^- \pi^0 \ell \nu_\ell$ decays, in which the two pions do not form a resonance, is created by using simulated events in which decays predicted by the BLNP model described in Chapter 5 fragment into a charged and a neutral pion. This background is treated separately from all other backgrounds as it is expected to have the same peaking M_{miss}^2 distribution as the signal process. The branching fraction of this process is completely unknown, therefore its normalization in the fit cannot be constrained to an expectation. The fragmentation model cannot describe interference with the signal process. To prevent unphysical values of the yield parameter and decrease the overall uncertainty, the template normalization is limited to positive values using the parameter transformation procedure described above.

$B\bar{B}$ background template All other $B\bar{B}$ background is described by a single template. The simulated events used to construct this template are mostly $B \rightarrow X_c \ell \nu_\ell$ decays due to the similar signatures and high branching fractions, however a significant contribution is also expected from $B \rightarrow X_u \ell \nu_\ell$ processes with non-dipion final states. A breakdown of individual modes and effects contributing to the template is given in Fig. 7.2a. The large contribution

labeled “ π^0 fake” is assigned if the π^0 candidate cannot be matched to a generated neutral pion. This can indicate that one of the calorimeter clusters used to reconstruct the π^0 candidate originates from beam interactions or photons radiated elsewhere in the decay. Common to these decays is that they are only reconstructed due to combinatorial effects in which e.g. a final-state particle is missed or misidentified. Therefore, these processes contribute mostly in the region with $M_{\text{miss}}^2 > 1.0 \text{ GeV}^2 c^{-4}$. Consequently, no peaking structures around $0 \text{ GeV}^2 c^{-4}$ in M_{miss}^2 are expected. The normalization of this template is left unconstrained in the fit as the efficiency of the FEI for misreconstructed decays (in which often a final-state particle candidate from the B_{sig} is instead used as the B_{tag}) has not yet been studied in detail.

Continuum template Events expected from light-quark processes are treated via a separate template which is constructed using samples simulated as described in Chapter 5. For these background processes, additional information from a separate data set is available to constrain the normalization of the $q\bar{q}$ template. Collision data corresponding to an integrated luminosity of 43 fb^{-1} recorded at a beam energy 60 MeV below the $\Upsilon(4S)$ mass is used. In this data set, variables such as the beam-constrained mass M_{bc} must be rescaled by the ratio of center-of-mass energies to correctly apply the same selections as in the nominal data set.

Using this additional data, the normalization of the simulated continuum template is scaled by a factor

$$f_{qq} = \frac{N_{or}}{N_{MC}} \times \frac{L_{MC}^{\text{int}}}{L_{or}^{\text{int}}} \times \left(\frac{E_{or}^*}{E_{MC}^*} \right)^2, \quad (7.16)$$

where the subscript *or* indicates off-resonance data and the subscript *MC* indicates the values used in simulation. This takes into account that the total hadronic cross section of e^+e^- collisions scales quadratically with the collision energy. The normalization of the $q\bar{q}$ template is constrained within a Gaussian distribution with mean and standard deviation corresponding to the scaled number of events in the off-resonance data set.

In $M(\pi\pi)$, I choose eleven equidistant bins between 0.48 GeV and 1.07 GeV, i.e. two times the ρ^+ width Γ_{ρ^+} . Finer bins increase the dependency on the simulated $M(\pi\pi)$ description, which especially for decays used to describe the $B\bar{B}$ background template uses several approximations. Additionally, bins in which less than 5 to 10 events are expected can lead to asymmetric uncertainties on the parameter of interest as the uncertainties in these bins become non-Gaussian.

In M_{miss}^2 , I choose non-equidistant bins to balance these effects with the need for a detailed description of the signal peak: For the sparsely populated region $-1.0 \text{ GeV}^2/c^4 < M_{\text{miss}}^2 < -0.4 \text{ GeV}^2/c^4$, a single bin is used to avoid empty bins in two dimensions. In the signal-rich region from $-0.4 \text{ GeV}^2/c^4 < M_{\text{miss}}^2 < 1.6 \text{ GeV}^2/c^4$, the bin width is chosen as $0.2 \text{ GeV}^2/c^4$. In the mostly background-dominated region in the M_{miss}^2 range 1.6 to $2.0 \text{ GeV}^2/c^4$, a single bin is used to reduce dependency on the description of $B \rightarrow X_c \ell \nu_\ell$ background processes. In total, this amounts to 12 bins in M_{miss}^2 . The fit templates and used bins in two dimensions

are illustrated in Fig. 7.1.

7.2.2 Fit Model in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ Analysis

Similar to the $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ signal extraction, four main two-dimensional templates in M_{miss}^2 and $M(\pi\pi)$ are defined for the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ signal extraction which are expected to describe almost all recorded and reconstructed events. These templates are one $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ signal template, one template describing nonresonant $B \rightarrow \pi\pi\ell\nu_\ell$ decays (here with two charged pions), one generic background template describing mostly $B \rightarrow X_c\ell\nu_\ell$ decays but also other decays not peaking in M_{miss}^2 , and a $q\bar{q}$ template. Also here, the normalization of the $q\bar{q}$ template is constrained using data recorded at energies below the $\Upsilon(4S)$ resonance. The composition of the combined background template is illustrated in Fig. 7.2b. In contrast, however, to the $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ reconstruction channel, additional unmeasured resonant backgrounds could contribute to the $B^+ \rightarrow \pi^\pm \pi^\mp \ell \nu_\ell$ final state. As precise modeling of these processes is necessary to describe the $M(\pi\pi)$ spectrum, four additional templates for the other resonant $B \rightarrow \pi\pi\ell\nu_\ell$ background modes, simulated as described in Section 5.2.1, are considered in the fit with floating normalizations which are again constrained to positive values. To achieve good separation between signal and background contributions while reducing uncertainties from limited MC statistics in the background template, non-equidistant bins are again used in M_{miss}^2 . However, owing to the better resolution in M_{miss}^2 originating from a decay signature without neutral particles, I choose even narrower bins than for the $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ fit. For $-1.0 \text{ GeV}^2/c^4 < M_{\text{miss}}^2 < -0.1 \text{ GeV}^2/c^4$, again a single bin in M_{miss}^2 with a width of $0.9 \text{ GeV}^2/c^4$ is used. In the range $-0.1 \text{ GeV}^2/c^4 < M_{\text{miss}}^2 < 0.2 \text{ GeV}^2/c^4$, I choose bins with a width of $0.1 \text{ GeV}^2/c^4$. For $0.2 \text{ GeV}^2/c^4 < M_{\text{miss}}^2 < 1.6 \text{ GeV}^2/c^4$, again wider bins with a width of $0.2 \text{ GeV}^2/c^4$ are chosen, and a single bin is used for M_{miss}^2 between 1.6 GeV^2 and 2.0 GeV^2 . Also here, the binning is chosen to avoid bins with less than 5 expected events total and to reduce dependence on the statistics of the simulated sample in high regions of M_{miss}^2 .

In $M(\pi\pi)$, eleven equidistant bins between $0.46 \text{ GeV } c^{-1}$ and $1.34 \text{ GeV } c^{-1}$ are used. This is consistent with the bins used in Ref. [79] to allow comparison with this result. An illustration of the fit templates is given in Figs. 7.3 and 7.4, the parameters of the signal extraction are provided in Table 7.1.

Constraint on the total $B \rightarrow \pi\pi\ell\nu_\ell$ branching fraction A preliminary analysis of the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ branching fraction presented in Ref. [123] has shown that significant systematic uncertainties are expected to originate from its dependence on the $B \rightarrow \pi\pi\ell\nu_\ell$ background model.

To reduce this dependence, an additional input is used in the form of a result for the total branching fraction of $B \rightarrow \pi\pi\ell\nu_\ell$ decays measured by the Belle collaboration [79]. This result is used to constrain the sum of $B \rightarrow \pi\pi\ell\nu_\ell$ modes in the signal extraction by adding a

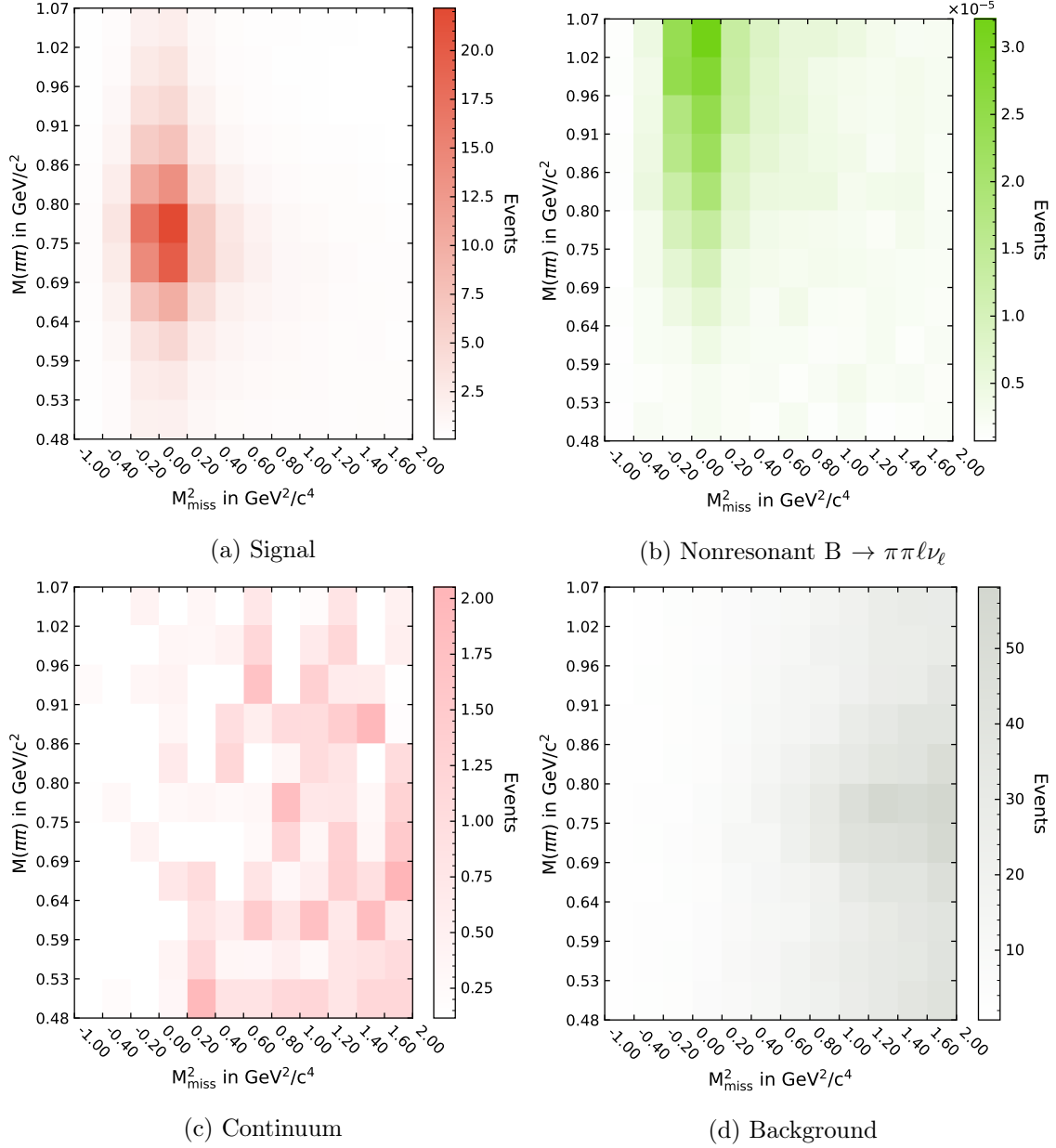


Figure 7.1: Fit templates in the $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ reconstruction channel in M_{miss}^2 and $M(\pi\pi)$ for (from top left) signal, nonresonant $B \rightarrow \pi^+ \ell \nu_\ell$, $q\bar{q}$ and $B\bar{B}$ background processes. The normalization is set to the expectation in simulation. Illustrations of the templates with uncertainties are given in Chapter A.

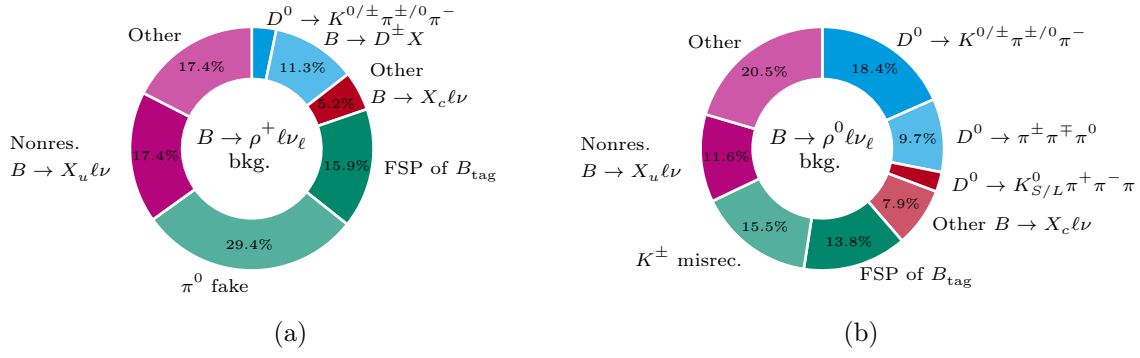


Figure 7.2: Composition of the background templates in both $B \rightarrow \rho \ell \nu_\ell$ reconstruction channels, grouped in categories likely to produce peaking structures in $M(\pi\pi)$. Decays are identified by the mother particle of the dipion candidate in simulation. $B \rightarrow X_c \ell \nu_\ell$ processes modeled with EvtGen’s Dalitz models are shown in blue, $B \rightarrow X_c \ell \nu_\ell$ decays differently are shown in red and combinatorial misreconstruction and misidentification is shown in green. The nonresonant component shown in dark purple originates from the BLNP model. The component shown in light purple contains, among others, resonant $B \rightarrow X_u \ell \nu_\ell$ backgrounds, beam background, and hadronic B decays with secondary leptons. More details on the simulation of these processes can be found in Section 5.3.

constraint term to the negative log-likelihood,

$$\left(\frac{\Delta \mathcal{B}_{\text{Belle 2021}} - \frac{1}{2N_{\text{B}\bar{\text{B}}}} \sum_i^6 \frac{Y_i}{\epsilon_i}}{\sigma_{\mathcal{B}_{\text{Belle 2021}}}} \right)^2, \quad (7.17)$$

with the central value $\Delta \mathcal{B}_{\text{Belle 2021}}$ and uncertainty $\sigma_{\mathcal{B}_{\text{Belle 2021}}}$ of the Belle partial branching fraction result and the number of B -meson pairs in the data set $N_{\text{B}\bar{\text{B}}}$. The signal normalizations Y_i corresponding to the templates described above as well as the reconstruction efficiencies ϵ_i for the processes in each of these templates are both allowed to vary in the fit, however the efficiencies are constrained within multiplicative systematic uncertainties to the values determined from simulation. As only the partial branching fraction in the range 0.46 to 1.34 GeV is used in the constraint, the reconstruction efficiency of each reconstructed template is also only determined in this mass range.

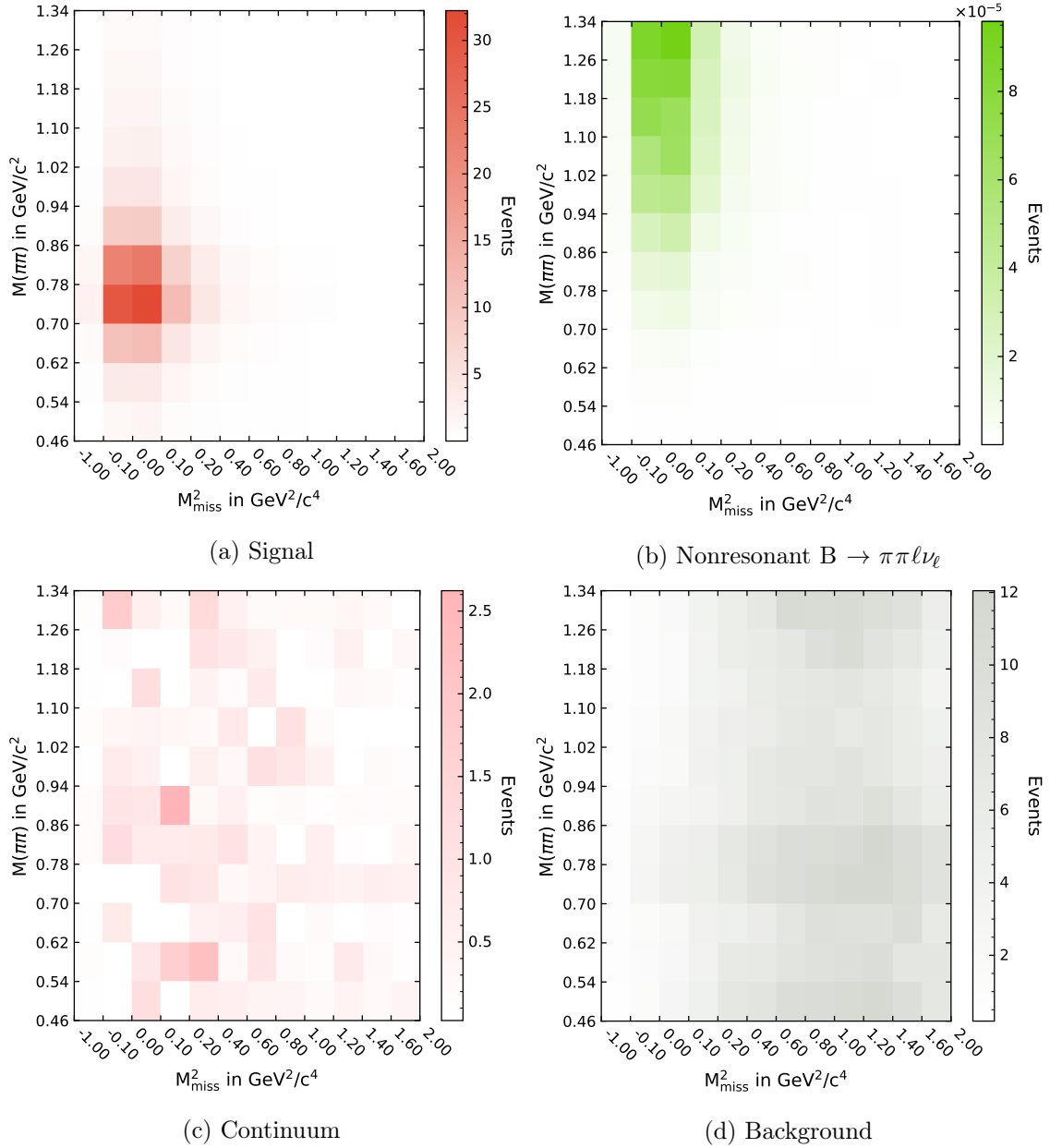


Figure 7.3: Four out of eight fit templates in M_{miss}^2 and $M(\pi\pi)$ in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ reconstruction channel with normalization set to the expectation in simulation. Illustrations of the templates with uncertainties are given in Chapter A.

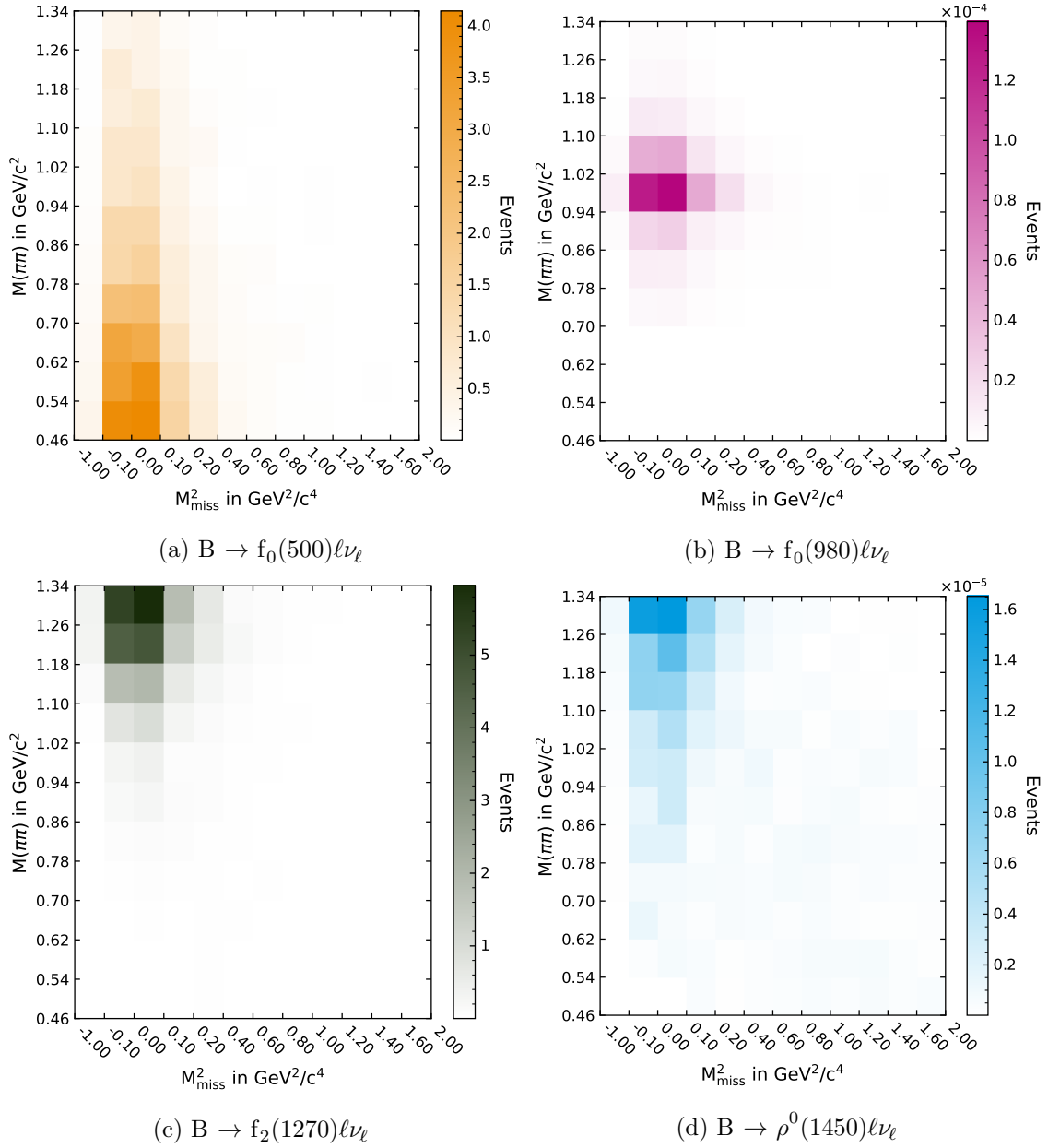


Figure 7.4: Four out of eight fit templates in M_{miss}^2 and $M(\pi\pi)$ in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ reconstruction channel with normalization set to the expectation in simulation. Illustrations of the templates with uncertainties are given in Chapter A.

7.3 Fit Validation with Simulated Data

The signal extraction procedure is validated entirely using simulated data as no proxy sample of recorded data suitable for this task exists. This validation aims to establish the reliability and model-dependence of the signal extraction procedure and is performed blind, i.e. without knowledge of the distributions of the fit variables in recorded data.

A first estimate of the expected sensitivity of the signal extraction procedure is derived using the Asimov data set [124]; the expected template yields and uncertainties are given in Table 7.1. As expected, the fit cannot determine the normalizations of the background templates very well. However, they still contribute to the overall uncertainty so are considered regardless in the background model. Projections of the expected yields in this data set as well as the expected uncertainties are shown in Fig. 7.5.

Due to the low numbers of events in individual bins, bin-wise correlations introduced by the nuisance parameters and the constraints implemented in the fit can introduce an asymmetry in the Likelihood profile. Thus, the Likelihood is scanned by fixing the parameter of interest (here the signal yield) to values 2σ above and below the minimum value found in the fit and minimizing all other parameters of interest. A bifurcated parabola is then fit to the scanned points to determine the parameters of the profile. While the fitted parabola describes the scanned points well in the case of $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$, the likelihood is only parabolic below 1σ in the $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ signal extraction. However, in both fits the Hesse approximation is sufficient to give a conservative estimate of the 1σ interval as illustrated in Fig. 7.6.

7.3.1 Validation using Simplified Statistical Models

To test whether the estimate \hat{y} of the true parameter values y_{true} can be reproduced by the signal extraction procedure, simplified statistical models (“toys”) are used. In a first step, only statistical uncertainties are considered by taking the expected number of events in each bin of the two-dimensional template as the mean λ of a Poisson distribution

Table 7.1: Expected yields in both $B \rightarrow \rho \ell \nu_\ell$ signal extraction channels with uncertainties determined from a fit to Asimov data.

Template	Expected Number of Events	
	$B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$	$B^0 \rightarrow \rho^- \ell^+ \nu_\ell$
$B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$	324.2 ± 32.4	271.7 ± 41.8
Nonresonant. $B \rightarrow \pi \pi \ell \nu_\ell$	27.3 ± 50.9	105.3 ± 46.1
Other Backgrounds	1679.1 ± 50.7	2794.2 ± 59.5
Continuum	94.3 ± 5.3	84.90 ± 6.16
$B \rightarrow f_0(500) \ell \nu_\ell$	17.1 ± 18.0	-
$B \rightarrow f_0(980) \ell \nu_\ell$	10.9 ± 11.8	-
$B \rightarrow f_2(1270) \ell \nu_\ell$	7.6 ± 30.6	-
$B \rightarrow \rho^0(1450) \ell \nu_\ell$	16.4 ± 62.2	-

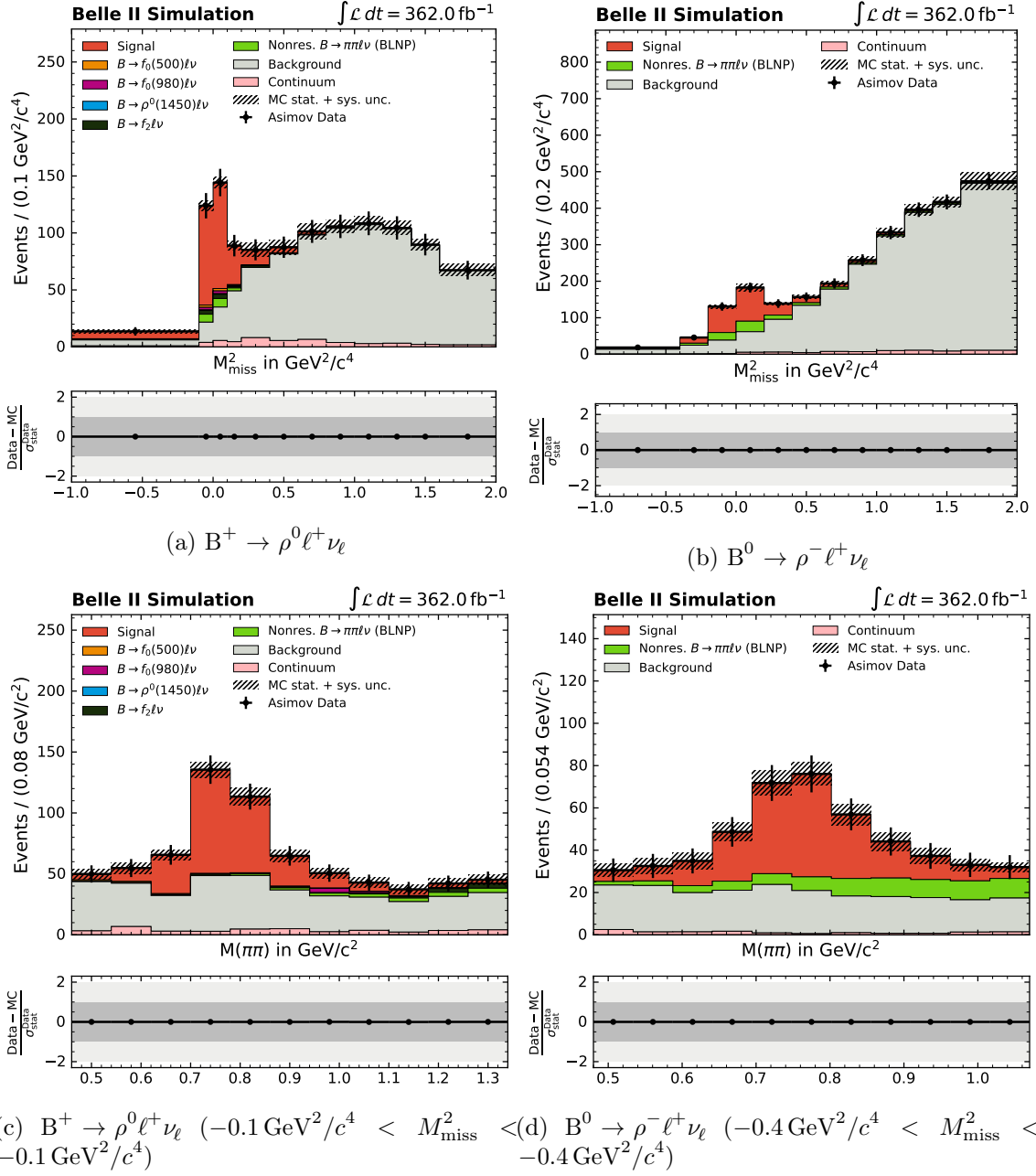


Figure 7.5: Projections of the fit variables (top) M_{miss}^2 in the full $M(\pi\pi)$ range and (bottom) $M(\pi\pi)$ in signal-rich regions of M_{miss}^2 , shown for the Asimov data sets in the (left) $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ and (right) $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ reconstruction channels.

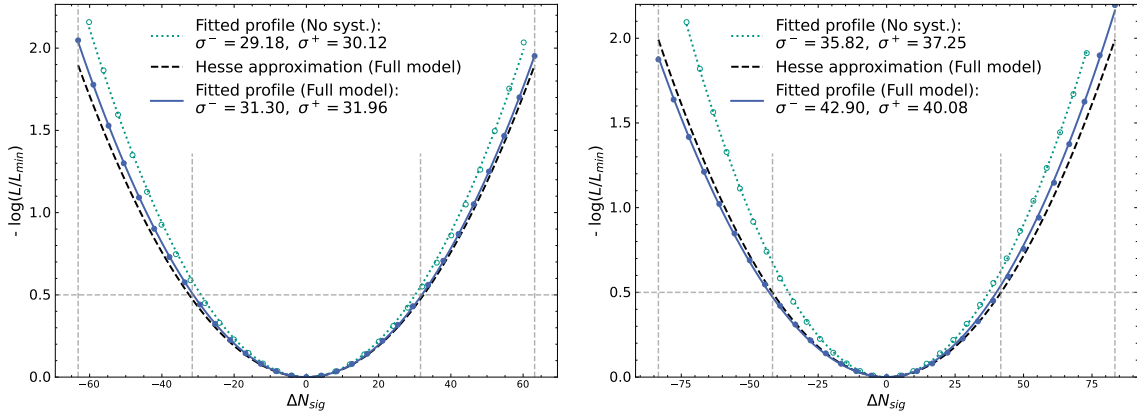


Figure 7.6: Likelihood scans to simulated data in the extraction of the signal yield parameter for (top) $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ and (bottom) $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ decays. Shown are the nominal fit setup and a fit setup in which no systematic uncertainties other than MC statistics are considered. Also shown is the Hesse approximation for the nominal fit setup. The dashed vertical lines indicate the 1σ and 2σ intervals of the nominal model, the dashed horizontal line indicates the point at which the likelihood value increases by a factor of 0.5.

$$\mathcal{P}_\lambda(k) = \frac{\lambda^k}{k!} e^{-\lambda} \quad (7.18)$$

from which the toys are sampled.

These toys are then fit using the nominal signal extraction procedure in which all bin-wise nuisance parameters are fixed to zero (as the sampled uncertainty does not include the effects of the uncertainties described by these parameters).

To also validate that the uncertainty estimate σ_y determined by the fit procedure can reproduce the true uncertainty, the pull distribution p is defined as

$$p = \frac{\hat{y} - y_{\text{true}}}{\sigma_y} \quad (7.19)$$

The standard deviation of this distribution is expected to be consistent with unity if the fit provides adequate coverage. The distributions of the event yields and pulls determined in the toy studies is shown in Fig. 7.7. Both $B \rightarrow \rho \ell \nu_\ell$ channels exhibit some bias on the parameter of interest, which I attribute to deviations from a Gaussian distribution caused by the lower bounds on the normalizations of both resonant and nonresonant $B \rightarrow \pi \pi \ell \nu_\ell$ background templates. In addition, this simplified signal extraction is not completely able to estimate the uncertainty of the sample as the assumption that the events in each bin follow a Gaussian distribution does not hold entirely.

Additional toy studies are conducted to estimate whether the signal extraction can accurately recover the shape-changing effects introduced by systematic uncertainties. To do this, the expectation for each template is modified with values for nuisance parameters drawn from a multi-variate Gaussian distribution for which the correlation structure is given by

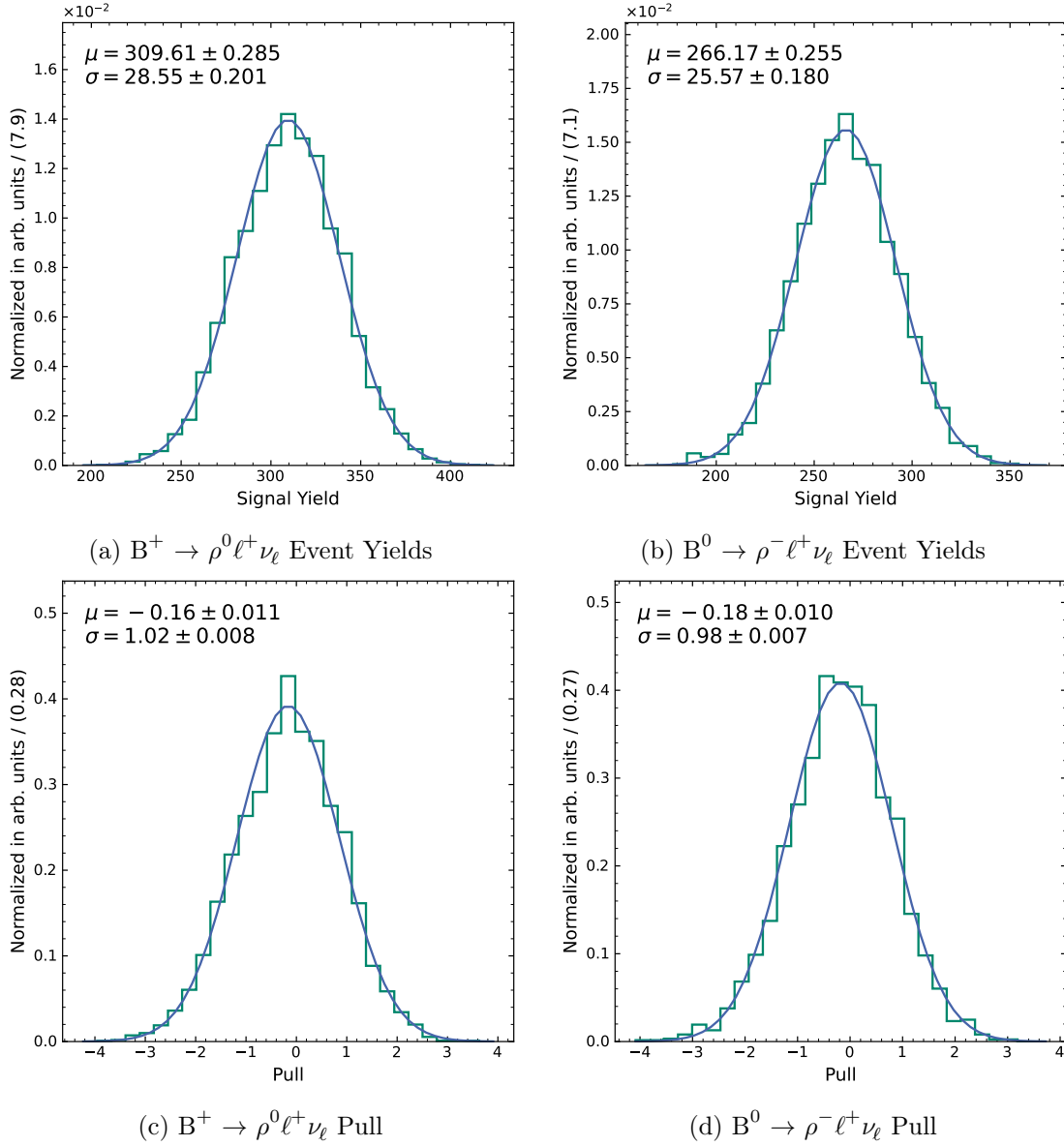


Figure 7.7: Normalized distributions of event yield and pull in the (left) $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ and (right) $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ signal extraction of the signal yield parameter, conducted using Poisson sampling of the nominal templates and template normalizations from MC expectation. Nuisance parameters in the fit are fixed to zero. The fit can reproduce the uncertainties but only reproduces the parameter of interest within 2.5% of the total event yield.

Table 7.2: Toy fit configurations in $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ for different factors $C_B \rightarrow X_u \ell \nu_\ell$ to the template normalizations expected from MC as well as averages determined from 300 toys. The table header identifies the templates by their hadronic component X_u .

ρ (Signal)	$f_0(500)$	$f_0(980)$	$f_2(1270)$	$\rho^0(1450)$	π^\pm π^\mp	$Y_{\text{signal}}/C_{\text{signal}}$
1.0	0.1	0.1	0.1	0.1	1.9	318.5 ± 0.5
0.5	3.3	3.3	3.3	3.3	1.8	320.0 ± 0.5
1.2	0.8	0.8	0.8	0.8	0.0	315.9 ± 0.4
1.0	2.0	0.9	0.9	0.9	0.9	317.8 ± 0.5
1.0	0.8	2.0	0.8	0.8	0.8	317.4 ± 0.5
1.0	1.0	1.0	1.0	2.0	1.0	317.5 ± 0.5
1.0	0.8	0.8	2.0	0.8	0.8	315.8 ± 0.5
0.5	2.0	2.7	2.7	2.0	2.7	322.9 ± 0.5

the bin-wise covariance matrix for each template. This essentially reproduces Eqs. (7.10) and (7.11) with random sampling of the parameters θ_k . The modified templates are then summed and are again used as the means of a Poisson distribution from which the bin contents of each toy are sampled.

This test, illustrated in Fig. 7.8, still indicates a bias in the signal extraction but shows that the method is able to reliably estimate the uncertainty. To incorporate this bias in the result, a corresponding systematic uncertainty is assigned as described in Section 8.2.

7.3.2 Linearity Tests in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ Reconstruction Channel

To test whether the fit in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ reconstruction channel is stable for varying expectations of the signal process and the background processes to the dipion final state, the toy procedure described above is repeated for various configurations given in Table 7.2. Each determination uses 2000 toy fits. In these configurations, the expectation for individual templates is scaled by fixed values $C_B \rightarrow X_u \ell \nu_\ell$ while all other templates are scaled by the same factor determined under the assumption that the sum of the $B \rightarrow \pi \pi \ell \nu_\ell$ branching fractions follows the value determined in Ref. [79]. The largest deviation from the expected signal yield corresponds to a bias of 1.98% which is determined using an enhancement of the $f_2(1270)$ branching fraction by a factor of two while keeping the expected signal yield constant.

7.3.3 Model Dependence of the Fit in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ Channel

To test how strongly the extracted branching fraction depends on the resonant and nonresonant $B \rightarrow \pi \pi \ell \nu_\ell$ background models, I use altered fit models in which the partial $B \rightarrow \pi \pi \ell \nu_\ell$ branching fraction not expected to originate from $B \rightarrow \rho \ell \nu_\ell$ is either completely resonant or completely nonresonant. This is achieved by fixing the normalization of the resonant or nonresonant $B \rightarrow \pi \pi \ell \nu_\ell$ background templates to zero. This model is then used to perform fits to toys drawn from the simulated expectation according to the approach described above

Table 7.3: Ratio of fitted to expected template yields in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ signal extraction in simulation. The fitted values are determined from 1000 toys in which only the resonant (Variant 1) or nonresonant (Variant 2) background $B \rightarrow \pi\pi\ell\nu_\ell$ template normalizations are allowed to float.

Template	$N_{\text{result}}/N_{\text{expected}}$	
	Resonant only	Nonresonant only
$B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$	1.010 ± 0.092	0.980 ± 0.091
Nonresonant. $B \rightarrow \pi\pi\ell\nu_\ell$	0 (fixed)	3.360 ± 0.969
$B \rightarrow f_0(500)\ell\nu_\ell$	1.095 ± 1.223	0 (fixed)
$B \rightarrow f_0(980)\ell\nu_\ell$	1.260 ± 0.764	0 (fixed)
$B \rightarrow f_2(1270)\ell\nu_\ell$	0.963 ± 0.817	0 (fixed)
$B \rightarrow \rho^0(1450)\ell\nu_\ell$	6.866 ± 7.594	0 (fixed)
Other Backgrounds	0.996 ± 0.029	1.004 ± 0.027
Continuum	1.006 ± 0.006	1.007 ± 0.006

in which the templates are modified using the nuisance parameters. The results of this fit are given in Table 7.3, the larger deviation of 2.0% is observed in a fit using a completely nonresonant model for the $B \rightarrow \pi\pi\ell\nu_\ell$ backgrounds.

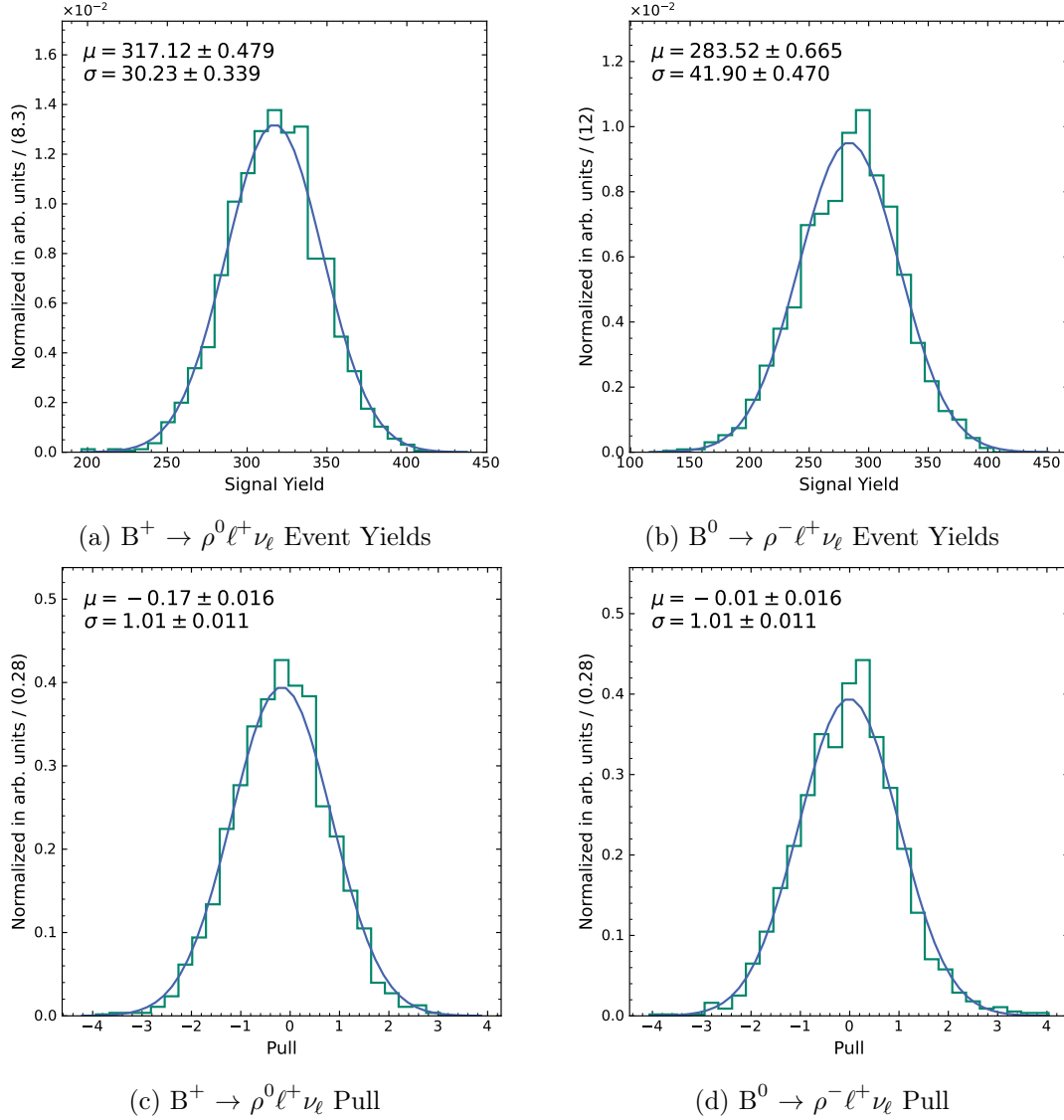


Figure 7.8: Normalized distributions of event yield and pull in the (left) $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ and (right) $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ signal extraction of the signal yield parameter, conducted using Poisson sampling of templates modified by sampling the nuisance parameters and template normalizations from MC expectation. Nuisance parameters in the fit are floating. The fit can reproduce the uncertainties for both $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ channels and can completely reproduce the $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ signal yield. For $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$, the toys reproduce the signal yield within 1.6%.

Chapter 8

Systematic Uncertainties

In this chapter, various sources of uncertainty and their estimation are discussed. Here, systematic uncertainties are defined as unknowns which do not scale directly with the size of the available data set. They can either originate from approximations in the underlying physical models, reconstruction effects, or limitations of the signal extraction procedure. The sources of systematic uncertainty are split into two categories: Additive and multiplicative uncertainties. For the former ones, I determine the correlation between all bins in each template and propagate the uncertainty into the fit via nuisance parameter as described in the previous chapter. For the latter ones, I estimate the effect using the fractional change in efficiency when varying each uncertain parameter. The relative sizes of the estimated systematic uncertainties are given in Table 8.3 with the shape-changing effects of all systematic sources combined to a single entry labeled “template shape.”

8.1 Correlation Matrices Estimation

To propagate uncertainties into each template t , candidate-wise weights w determined for each source of uncertainty are used. Uncertainties originating from a single parameter uncorrelated to all other sources are treated by varying the w up and down within its 1σ interval and recording the varied templates t_{up} and t_{down} . The covariance matrix C_{tk} for each template t and each source k is then estimated from the difference $t_{\text{diff}} = (t_{\text{up}} - t_{\text{down}})/2$ as

$$C_{tk} = t_{\text{diff}} \otimes t_{\text{diff}}. \quad (8.1)$$

This approach can be extended to asymmetric uncertainties σ_{up} and σ_{down} for which the covariance matrix is estimated as

$$C_{tk} = \frac{(t_{\text{up}} - t) \otimes (t_{\text{up}} - t) + (t_{\text{down}} - t) \otimes (t_{\text{down}} - t)}{2}. \quad (8.2)$$

For multiple correlated parameters, two approaches are followed depending on the degree of correlation and the number of parameters. In the first approach, referred to as “eigenvariations,”

the covariance matrix giving the correlation between the parameters is decomposed into its eigenvectors and eigenvalues which are then used to determine variations on the nominal parameters. The resulting uncorrelated variations are each treated the same way as above. In the second approach, referred to as “toy variations,” the correlation between parameters such as calibration factors is used to sample weights from a multivariate Gaussian distribution. The varied templates resulting from each of these toy variations are then used to estimate the bin-wise sample covariance.

8.2 Sources of Systematic Uncertainty

I first describe approximations from the theoretical models or the simulation, followed by systematic uncertainties from reconstruction effects and uncertainty originating from the signal extraction procedure.

8.2.1 Uncertainties from Theoretical Models and Simulation

$\rho - \omega - \pi\pi$ interference Comparison between simulated and recorded data in a signal-rich region of M_{miss}^2 in Fig. 8.1a shows an overestimation of the ρ signal at approximately the ω mass $m_\omega = 782$ MeV. This effect is attributed to $\rho - \omega$ interference. While the branching fraction of $B \rightarrow \omega(\rightarrow \pi\pi)\ell\nu_\ell$ is two magnitudes smaller than that of $B^+ \rightarrow \rho^0\ell^+\nu_\ell$, this interference can still distort the line shape. This has previously been observed in other processes in which two pions are produced such as e.g. $e^+e^- \rightarrow \pi^+\pi^-(\gamma)$ [125, 126]. Ideally, the size of this distortion, which depends on the strong phase difference between the decay via a ρ meson and the decay via a ω meson, would be estimated directly in the signal extraction procedure. The lack of a sample with high enough purity and large enough sample size, however, render this infeasible. Instead, an estimate of the size of this effect is determined in a fit to the $B \rightarrow \pi\pi\ell\nu_\ell$ spectrum measured in [79] is performed by Ref. [74]. This fit, referred to as an “amplitude fit” to distinguish it from the nominal signal extraction procedure in this analysis, has floating parameters for the $B \rightarrow \rho\ell\nu_\ell$ and nonresonant $B \rightarrow \pi\pi\ell\nu_\ell$ branching fractions as well as the strong phase difference between the decay via ρ and via ω . The covariance between the three floating parameters, which is determined by the fit, is used to create independent versions of the parameter sets. The amplitude fit and its 1σ eigenvariations are illustrated in Fig. 8.1, however, as the nominal model does not include any interference the eigenvariations with respect to the best-fit values are not independent with respect to this nominal model. This means incorporating all of them in the nominal fit overestimates the uncertainty significantly. Instead, eigenvariation 3, which has the largest effect on the $B^+ \rightarrow \rho^0\ell^+\nu_\ell$ signal extraction and which is almost identical to the best-fit values, is used as a downward systematic uncertainty. The corresponding upward systematic uncertainty is obtained by taking the difference between the nominal line shape and the one corresponding to eigenvariation 3 and adding it to the nominal line shape.

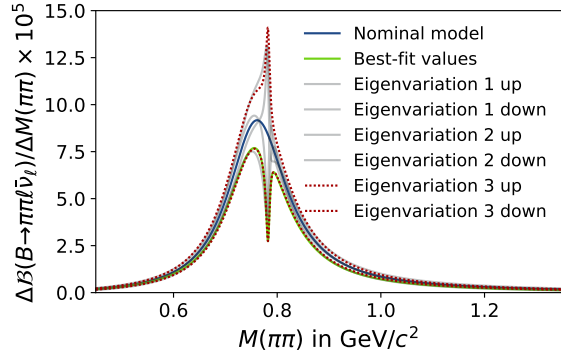
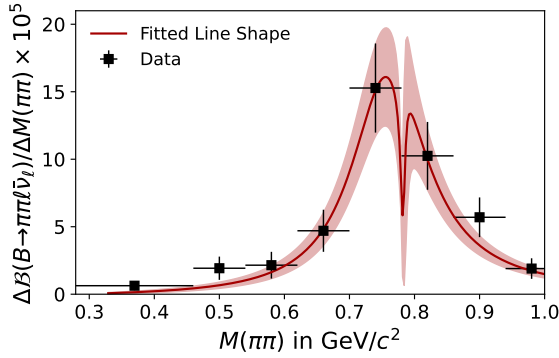
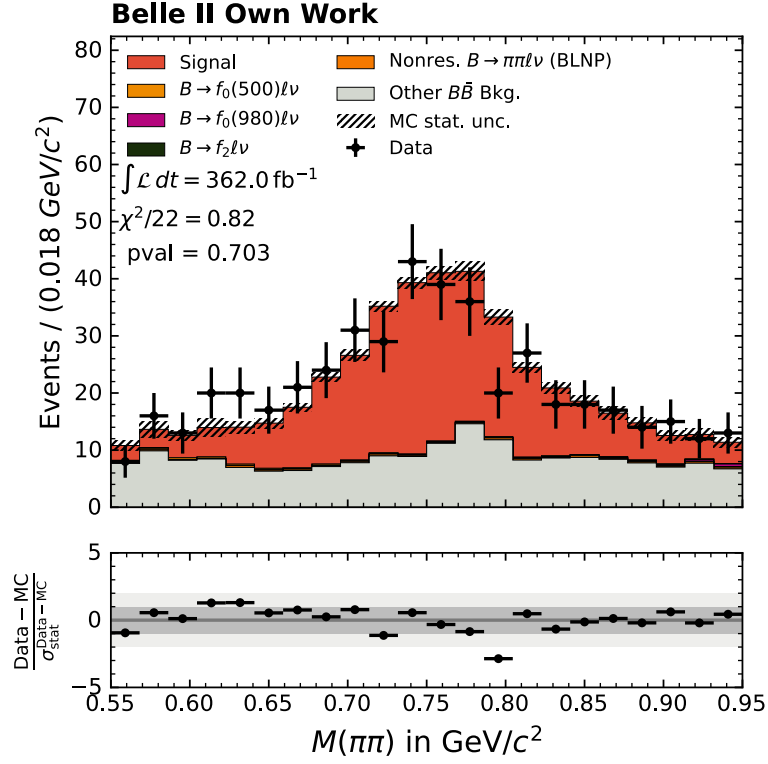


Figure 8.1: Overestimation of the number of events in several bins of simulated data in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ reconstruction, amplitude fit to Ref. [79] and varied versions of the fit parameters [74]. Eigenvariation 3, which has the largest effect, is used as a systematic uncertainty estimate in the fit.

The resulting systematic uncertainty is only propagated to the signal template. It is validated by applying the nominal signal extraction procedure to a simulated data set in which the ρ line shape is distorted according to the best-fit values obtained in the amplitude fit.

BLNP model parameters As described in Section 5.2.2, the inclusive BLNP model is parameterized by two HQE parameters m_b and μ_π^2 and three scales μ_h , μ_i and $\bar{\mu}$. The HQE parameters, which are determined from $B \rightarrow X_c \ell \nu_\ell$ moments in Ref. [16], have correlated uncertainties from this determination itself and uncorrelated uncertainties from the conversion of the fit result to a different scheme. Consequently, after converting m_b and μ_π^2 into two uncorrelated parameters, in total four (two up, two down) variations of the HQE parameters are obtained for the fit uncertainties and four additional variations are required for the uncorrelated scheme uncertainties. A third set of four variations arises from the scale parameters, the parameters μ_h and $\bar{\mu}$ are each varied between $1/\sqrt{2}$ and $\sqrt{2}$ to obtain uncertainties. I use EvtGen to produce simulated samples corresponding to these 12 variations, illustrated in Fig. 8.2, and reweight the reconstructed simulated samples to match their description. Large uncertainties are expected here as the model cannot provide very accurate descriptions in the region $m_X < 1.5 \text{ GeV } c^{-1}$. Events simulated with the BLNP model completely define the nonresonant $B \rightarrow \pi\pi\ell\nu_\ell$ templates and contribute heavily to the combined background templates. To maintain the consistency of the model within the combined background templates, the hybrid weights introduced in Section 5.2.3 are recalculated for each variation of the BLNP parameters and considered in the covariance between the bins. For the nonresonant $B \rightarrow \pi\pi\ell\nu_\ell$ templates, only the variations of the BLNP parameters themselves are considered as uncertainties. The varied hybrid weights are not considered here since the floating normalization already incorporates possible deviations from the inclusive $B \rightarrow X_u \ell \nu_\ell$ branching fraction constraint.

PYTHIA fragmentation To account for the fragmentation and hadronization in PYTHIA and the resulting effects on the nonresonant $B \rightarrow \pi\pi\ell\nu_\ell$ background template, an uncertainty is assigned based on the correction procedure described in Section 5.2.2. The distribution as fragmented by PYTHIA is considered the down-variation and the total m_X distribution (which is equivalent to only allowing the X_u system from BLNP to decay to two charged pions) is considered the up-variation. Full correlation is assumed between bins of this uncertainty, the total uncertainty is illustrated in Fig. 8.3.

Uncertainties introduced by the hybrid modelling technique Additional systematic uncertainty is assigned by replacing the ISGW2 model used to simulate the four resonant $B \rightarrow X_u \ell \nu_\ell$ modes with

$X_u \in \{f_0(500), f_0(980), f_2(1270), \rho^0(1450)\}$ with EvtGen's PHSP model in which the particles are distributed uniformly in the allowed phase space. While the $M(\pi\pi)$ and M_{miss}^2 distributions

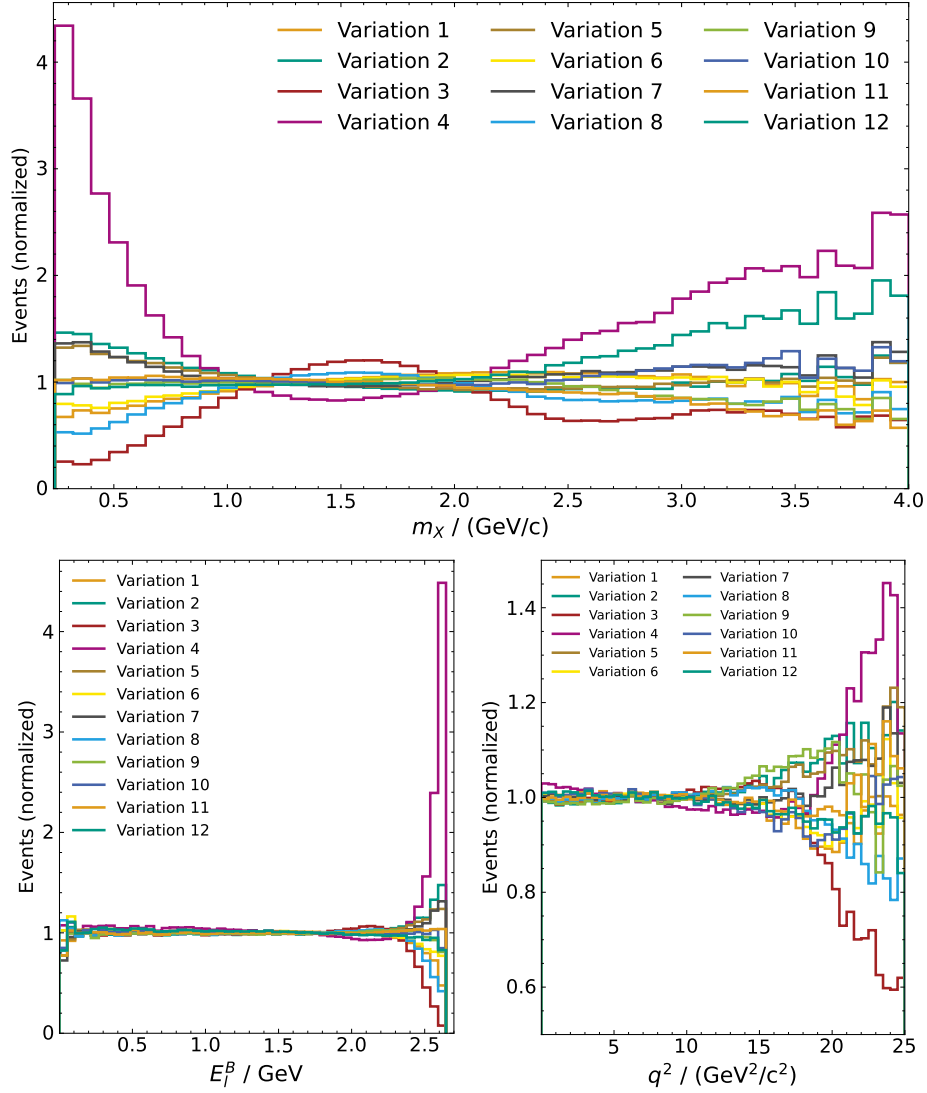


Figure 8.2: m_X , E_l^B and q^2 distributions in simulated data without reconstruction effects for 12 variations of the parameters used in the BLNP model, normalized to the distributions obtained with the nominal parameters. The region $m_X < 1.5 \text{ GeV } c^{-1}$ exhibits large systematic uncertainty which is reflected in the $B \rightarrow \pi\pi\ell\nu_\ell$ templates in the signal extraction (see Figs. A.2 and A.6).

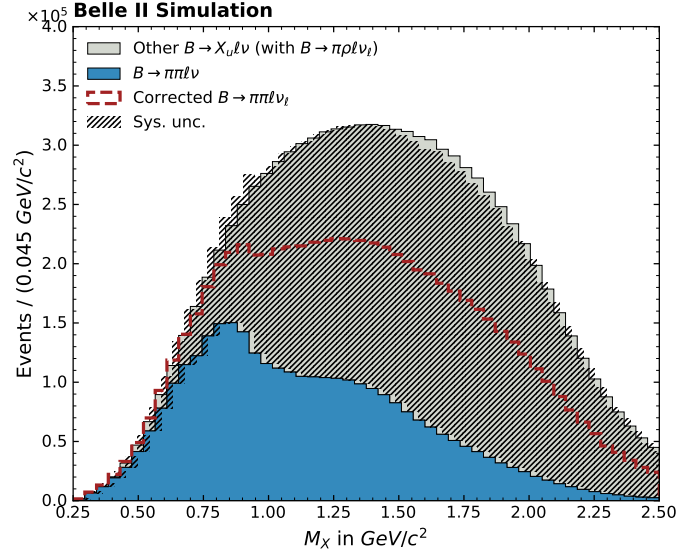


Figure 8.3: Systematic uncertainty assigned to the shape of the $B \rightarrow \pi\pi l\nu_\ell$ template partially originating from fragmentation of the X_u system generated in the BLNP model. The template is allowed to fluctuate between the nominal shape and the overall m_X distribution. Full correlation is assumed between the bins. The figure shows a slight discrepancy between the assigned uncertainty and the sum of the stacked histograms which originates from binning effects.

of the samples after selection do not differ between those simulated with ISGW2 and those simulated with PHSP, the E_ℓ^B and q^2 distributions do differ. This in turn changes the hybrid model composition which is determined in bins of the latter two variables. To incorporate this uncertainty into the signal extraction, I calculate hybrid weights using both models and incorporate the difference between them into the fit as a fully correlated 1σ systematic uncertainty.

$B \rightarrow X_u l\nu_\ell$ and $B \rightarrow X_c l\nu_\ell$ branching fractions Uncertainties can be assigned to each the $B \rightarrow X_c l\nu_\ell$ branching fractions given in Table 5.2 and the $B \rightarrow X_u l\nu_\ell$ background branching fractions in Table 5.1. These originate mostly from measurements or isospin relations to measured decays, the only exceptions are $B \rightarrow D^* \eta l\nu_\ell$ and $B \rightarrow D \eta l\nu_\ell$ decays for which no measurements exist and which are used to fill the difference between the sum of exclusive modes and the inclusive $B \rightarrow X_c l\nu_\ell$ branching fraction. For these unmeasured decays, an uncertainty of 100% is assigned. Events corresponding to these decay modes in the simulated sample are weighted up and down within their uncertainties. As the uncertainties between different channels are uncorrelated, the individual uncertainties are then propagated separately into the combined background template using the covariance matrix approach given above.

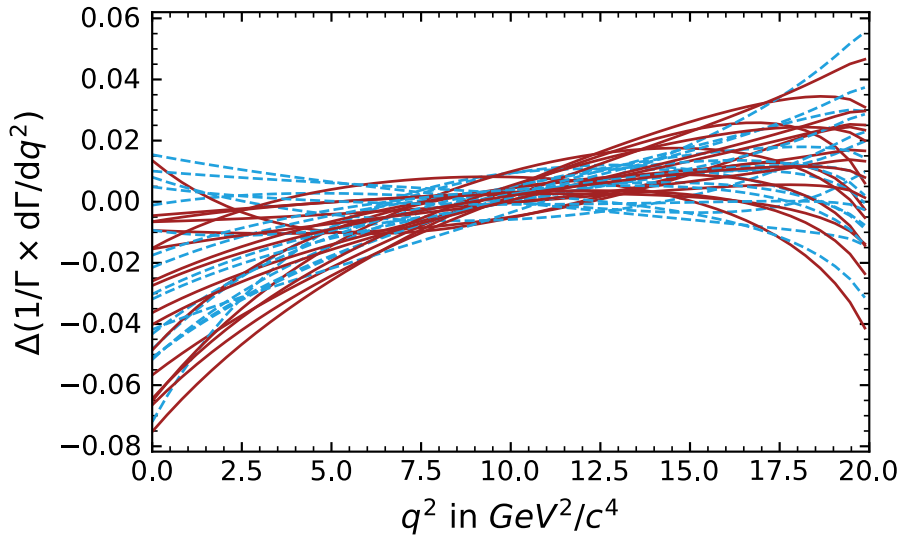


Figure 8.4: 19 independent (dashed) up and (solid) down variations of the parameters used in the $B \rightarrow \rho$ form factor, normalized to the form factor with best-fit parameters.

Uncertainties from the branching fractions of D^0 - and D^\pm -meson decays To accurately incorporate the uncertainties of the branching fractions of D^0 and D^\pm mesons, the 20 decay modes contributing more than 80% to each of the two background templates in both $B \rightarrow \rho l \nu_\ell$ reconstruction channels are identified and uncertainties are assigned to the template using the method described above. No correlation is assumed between measurements of the individual decay modes. The values used as uncertainties are given in Table 8.1 for D^0 decays and Table 8.2 for D^\pm decays. When consistent with the values used to simulate the MC sample, uncertainties from Ref. [5] (preferred) and Ref. [127] are used directly, otherwise an uncertainty which covers the central value in Ref. [5] is used.

Uncertainties from D decays are not only considered for D mesons originating from the B_{sig} decays as identified via the lepton mother but also for D decays on the tag-side. This is necessary as the FEI calibration factors do not incorporate this source of uncertainty into the provided total uncertainty.

$B \rightarrow X_u l \nu_\ell$ and $B \rightarrow X_c l \nu_\ell$ form factors As described in Chapter 5, the form factor parameters of the $B \rightarrow \rho l \nu_\ell$ signal process, other $B \rightarrow X_u l \nu_\ell$ decays, and $B \rightarrow X_c l \nu_\ell$ processes are determined from measurements. Consequently, they are given with uncertainties which I propagate into the analysis using correlation matrices for each set of parameters and the eigenvariation approach outlined above. The varied candidate weights from this approach are then considered for both the signal and combined background templates. For uncertainties from $B \rightarrow \rho l \nu_\ell$ form factors, which in contrast to other semileptonic processes must not only be considered for their effect on the template shapes, the form factor variations are illustrated in Fig. 8.4.

Table 8.1: Branching fractions in the simulated sample and uncertainties for D^0 decays. The column “MC Model” corresponds to the decay model in EvtGen (also see Ref. [71]). If this column is “D_DALITZ” the given decay mode and branching fraction should be interpreted to include intermediate resonances, otherwise the branching fraction should be interpreted as exclusive.

D^0 decay mode	MC Model	\mathcal{B}	in	%
$K^- \pi^+ \pi^0$	D_DALITZ	14.4	\pm	0.5
$K^- a_1(1260)$	SVS	14.02	\pm	0.99
$K^*(892) \rho^\pm$	SVV_HELAMP	6.79	\pm	2.60
$K^- \pi^+$	PHSP	3.947	\pm	0.03
$K^- e^+ \nu_e$	SLBKPOLE	3.549	\pm	0.026
$K^- \mu^+ \nu_\mu$	SLBKPOLE	3.41	\pm	0.04
$K_S^0 \pi^+ \pi^-$	D_DALITZ	2.8	\pm	0.18
$K_L^0 \pi^+ \pi^-$	D_DALITZ	2.8	\pm	0.18
$\bar{K}^*(892)^0 \pi^+ \pi^-$	PHSP	2.4	\pm	0.05
$K^- \pi^+ \omega$	PHSP	1.9	\pm	0.78
$\bar{K}^*(892)^0 \pi^+ \pi^- \pi^0$	PHSP	1.9	\pm	0.9
$K^*(892) \mu^+ \nu_\mu$	SLPOLE	1.89	\pm	0.24
$K^- \pi^+ \pi^+ \pi^-$	PHSP	1.81	\pm	0.07
$K_1(1270) \pi^+$	SVS	1.6	\pm	0.45
$\bar{K}^*(892)^0 \rho^0$	SVV_HELAMP	1.58	\pm	0.08
$\pi^+ \pi^- \pi^0$	D_DALITZ	1.49	\pm	0.06
$K_S^0 \pi^0$	PHSP	1.24	\pm	0.022
ωK_L^0	SVS	1.1	\pm	0.4
$\pi^+ \pi^- \pi^0 \pi^0$	PHSP	0.936	\pm	0.09
$K_L^0 \pi^+ \pi^- \pi^0$	PHSP	0.6	\pm	0.6
$\pi^+ \pi^+ \pi^- \pi^-$	PHSP	0.562	\pm	0.23
$K^- \pi^+ \rho^0$	PHSP	0.5	\pm	0.23
$K^- K^+$	PHSP	0.408	\pm	0.006
$\pi^- \mu^+ \nu_\mu$	SLBKPOLE	0.27	\pm	0.012
$\pi^+ \pi^-$	PHSP	0.15	\pm	0.02
Sum		71.78	\pm	3.22

Table 8.2: Branching fraction in the simulated sample and uncertainties for D^+ decays. The column “MC Model” corresponds to the decay model in EvtGen. If this is “D_DALITZ” the given decay mode and branching fraction should be interpreted to include intermediate resonances. If the MC model is not a Dalitz model, the branching fraction should be interpreted as exclusive.

D^+ decay mode	MC Model	\mathcal{B}	in	%
$K^- \pi^+ \pi^+$	D_DALITZ	9.38	\pm	0.16
$\bar{K}^0 \mu^+ \nu_\mu$	SLBKPOLE	8.76	\pm	0.19
$\bar{K}^0 e^+ \nu_e$	SLBKPOLE	8.72	\pm	0.09
$K_S^0 \pi^+ \pi^0$	D_DALITZ	7.36	\pm	0.21
$K_L^0 \pi^+ \pi^0$	D_DALITZ	7.36	\pm	0.21
$\bar{K}^*(892)^0 e^+ \nu_e$	SLPOLE	5.40	\pm	0.1
$\bar{K}^*(892)^0 \mu^+ \nu_\mu$	SLPOLE	5.27	\pm	0.15
$\bar{K}^*(892)^0 \pi^0 \pi^+$	PHSP	4.71	\pm	1.2
$K_0^{*1} \pi^+$	SVS	2.71	\pm	0.91
$K_S^0 a_1(1260)$	SVS	2.59	\pm	0.79
$K_L^0 a_1(1260)$	SVS	2.59	\pm	0.79
$K_S^0 \pi^+$	PHSP	1.56	\pm	0.03
$K_L^0 \pi^+$	PHSP	1.46	\pm	0.05
$\bar{K}^*(892)^0 \rho^+$	SVV_HELAMP	1.34	\pm	0.8
$\pi^+ \pi^+ \pi^- \pi^0$	PHSP	1.16	\pm	0.08
$K^+ K^- \pi^+$	D_DALITZ	0.968	\pm	0.018
$K_{*-}^0 \pi^+ \pi^+$	PHSP	0.92	\pm	0.9
$K^- \rho^+ \pi^+$	PHSP	0.80	\pm	0.40
$\eta' \pi^+$	PHSP	0.497	\pm	0.019
$\eta \pi^+ \pi^+ \pi^-$	PHSP	0.34	\pm	0.02
$\pi^+ \pi^+ \pi^-$	D_DALITZ	0.327	\pm	0.018
Sum		74.22	\pm	2.31

8.2.2 Uncertainties from Reconstruction Effects

Particle identification and π^0 reconstruction efficiency For uncertainties on the lepton and pion identification efficiency and misidentification rate as well as the π^0 reconstruction efficiency, the toy variation approach is used. The correction factors are given in bins of the particle momentum and, for the particle identification correction, also in bins of the polar angle. This introduces complex correlations between the correction factors as no correlation between the factors must be assumed for the statistical component of the uncertainty and full correlation must be assumed for the systematic part. To correctly model these correlations, 400 variations are determined for each source and the propagated into the correlation matrix using the sample covariance. For the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ reconstruction channel, this is done twice for the pion identification calibration factors to account for the two pions used in the reconstruction. The uncertainty on the π^0 reconstruction efficiency, which is only considered in the $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ reconstruction channel, is given with two independent sources of systematic uncertainties for each factor. These are each treated as fully correlated between the factors but uncorrelated between the two sources.

FEI efficiency Systematic uncertainties are assigned for the FEI calibration factors obtained with the χ^2 fit procedure described in Section 6.2.2. The calibration factors are provided together with covariance matrices describing the correlation structure between factors determined for each FEI channel subset which are illustrated in Figs. D.1 and D.2.

Using the toy variation approach described above, 500 variations of the calibration factors are determined and used to propagate the correlated uncertainties into each template. The multiplicative uncertainty from the FEI calibration is the dominant systematic effect on the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ branching fraction determination.

Tracking efficiency Systematic uncertainty arises from the efficiency of the tracking algorithms which has been studied in Ref. [128]. Based on this study, a systematic uncertainty of 0.24% is assigned for each track used to reconstruct a B_{sig} candidate. Due to its comparatively small value, this uncertainty is considered to have a purely multiplicative effect on the result.

8.2.3 Uncertainties from the Signal Extraction Procedure

Multiplicative uncertainties In addition to the sources described above which must be considered in the fit, two additional quantities enter the branching fraction determinations in Eq. (9.1): The number of B mesons in the data set $N_{B\bar{B}}$ determined in analogy with the method described in Ref. [17, Chapter 3] and the ratio of $\Upsilon(4S)$ decays to charged and neutral B mesons f_{+0} determined in Ref. [129].

$N_{B\bar{B}}$ is determined independently for data taking periods in which the conditions at the accelerator are assumed to be unchanged. The determination relies on the comparison of data taken at the $\Upsilon(4S)$ resonance and data taken 60 MeV below it. Unfortunately, these

Table 8.3: Sources of systematic uncertainty, quoted as a percentage of the $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ measured branching fractions. The template shape uncertainty is determined by fixing all shape nuisance parameters in the fit to data to zero and subtracting in quadrature the resulting uncertainty from the uncertainty of the nominal fit to data. The value given in the row labeled “Total (without shape)” is used as an uncertainty on the reconstruction efficiency in Table 9.3.

Source	$B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ uncert. / %	$B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ uncert. / %
Template shape	2.63	3.10
MC statistics	0.44	0.38
D^0/D^+ decay BFs	1.68	1.90
$B \rightarrow \rho \ell \nu_\ell$ Form Factor	0.27	0.25
FEI calibration	3.46	3.57
Pion ID eff.	< 0.1	< 0.1
Lepton ID eff.	0.36	0.39
$\rho - \omega - \pi\pi$ interference	–	0.18
π^0 efficiency	3.61	–
$N_{B\bar{B}}$	1.55	1.55
f_{+0}	2.52	2.68
Tracking eff.	0.48	0.72
Fit model dependence	3.11	2.00
Total	7.33	6.35
Total (without shape)	6.85	5.55

off-resonance data sets are not available for each data taking period. In this case, the ratio must be extrapolated from regions where they are. The dominant uncertainty in the determination of $N_{B\bar{B}}$ originates from this extrapolation. The uncertainty on the ratio f_{+0} is mostly theoretical in nature and originates from the assumption that isospin symmetry holds between charged and neutral $B \rightarrow J/\psi K$ decays. These two uncertainties have purely multiplicative effects on the result and are therefore not considered directly in the signal extraction.

Fit bias As described in Section 7.3, the signal extraction procedure cannot completely recover the expected branching fraction due to non-linearities in the likelihood function introduced by the systematic uncertainties and parameter limits. The size of this effect is estimated by repeating the toy studies described in Section 7.3 but using the best-fit template yields as the expectation. This determination yields uncertainties of 3.11% for the $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ signal yield and 1.42% for the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ signal yield. However, for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ I also consider the tests in Sections 7.3.2 and 7.3.3 which give a maximum deviation of 2.0% from the nominal result. This is used as a more conservative estimate on the ability of the fit to recover the branching fraction of the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ process. These effects are included as as “Fit model dependence” in Table 8.3.

Chapter 9

Results

To measure the $B \rightarrow \rho\ell\nu_\ell$ branching fractions, first an estimate on the number of signal events in the data set is obtained as described in Chapter 7. This signal yield is then combined with the efficiency determined from simulation, which incorporates both detector acceptance and selection efficiency, and an estimate of the number of B mesons in the data set.

9.1 Fit to Data

In a first step, the fit procedure introduced in Section 7.2 is applied to data and the yields given in Table 9.1 are obtained. These are then used to re-determine the hybrid weights which depend on the $B \rightarrow \pi\pi\ell\nu_\ell$ background normalizations as described Section 5.2.3 and which thus indirectly influence the distribution of the combined background template in each $B \rightarrow \rho\ell\nu_\ell$ channel. Backgrounds for which the normalization parameter is at the limit are set to zero.

In a second step, the fit to data is repeated with these modified templates and the values in Table 9.2, which decrease by 1.54% for $B^+ \rightarrow \rho^0\ell^+\nu_\ell$ and by 3.52% for $B^0 \rightarrow \rho^-\ell^+\nu_\ell$ from the first determination, are obtained. Projections of the fit results to both signal extraction variables in both $B \rightarrow \rho\ell\nu_\ell$ channels are shown in Fig. 9.2. Bin-by-bin figures for the $B^0 \rightarrow \rho^-\ell^+\nu_\ell$ fit can be found in Section B.0.1, the equivalent figures for the $B^+ \rightarrow \rho^0\ell^+\nu_\ell$ fit can be found in Section B.0.2. The likelihood profile is illustrated in Fig. 9.1 and exhibits good agreement between the uncertainties approximated with the Hesse matrix and those determined with likelihood profiling in the 1σ interval. Beyond the 1σ interval, the point-wise evaluation of the likelihood profile determined for the $B^0 \rightarrow \rho^-\ell^+\nu_\ell$ decay channel shows some disagreement with the Hesse parabola. As the likelihood profile evaluation performed with the MINOS algorithm, however, indicates no large downward uncertainty and the agreement is good in the 1σ interval, the Hesse matrix approximation is still considered a valid estimate of the uncertainty. The asymmetric uncertainties determined with MINOS are given in Table 9.2. Again, several background yields have best-fit points below the limit at zero which means that their uncertainties cannot be estimated via likelihood profiling.

Table 9.1: Ratio of obtained to expected template normalizations in a fit to data. As expected, the normalizations of several templates used for unmeasured background processes deviate from the expected normalization or are at the limit. As parameter values are not the final signal yields and are only used to re-determine the background template, they are given without uncertainties.

Process	$N_{\text{fit}}/N_{\text{expected}}$	
	$B^0 \rightarrow \rho^- \ell^+ \nu_\ell$	$B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$
Signal	0.87	0.95
Nonres. $B \rightarrow \pi\pi\ell\nu_\ell$	0.00	0.00
Background	0.94	0.91
Continuum	1.00	1.00
$B \rightarrow f_0(500)\ell\nu_\ell$		7.73
$B \rightarrow f_0(980)\ell\nu_\ell$		0.00
$B \rightarrow f_2(1270)\ell\nu_\ell$		3.88
$B \rightarrow \rho^0(1450)\ell\nu_\ell$		1.85

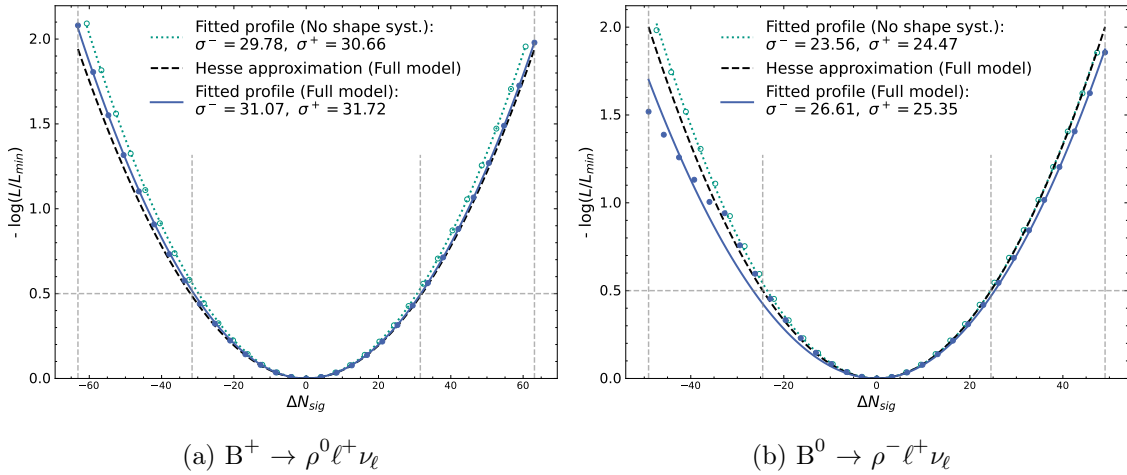


Figure 9.1: Likelihood profiles in the fit to data, determined by pointwise evaluation of the likelihood and fit with a bifurcated parabola. Good agreement is observed between pointwise evaluation and the Hesse matrix approximation within the 1σ interval in both $B \rightarrow \rho\ell\nu_\ell$ channels. In the $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ signal extraction, slight disagreement between the pointwise determination and the Hesse matrix approximation is observed in the interval -2σ to -1.5σ which is attributed to the parameter limits.

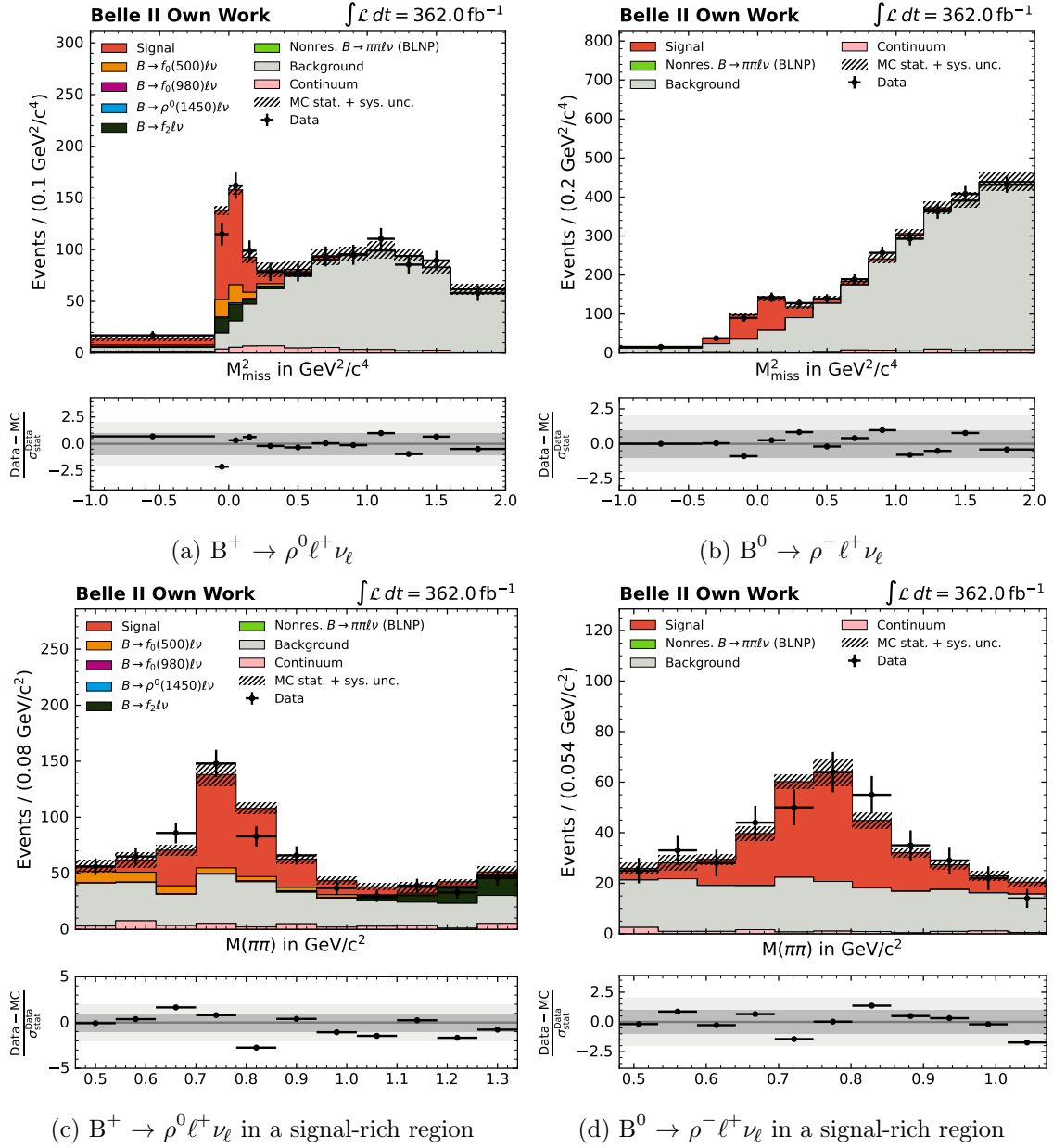
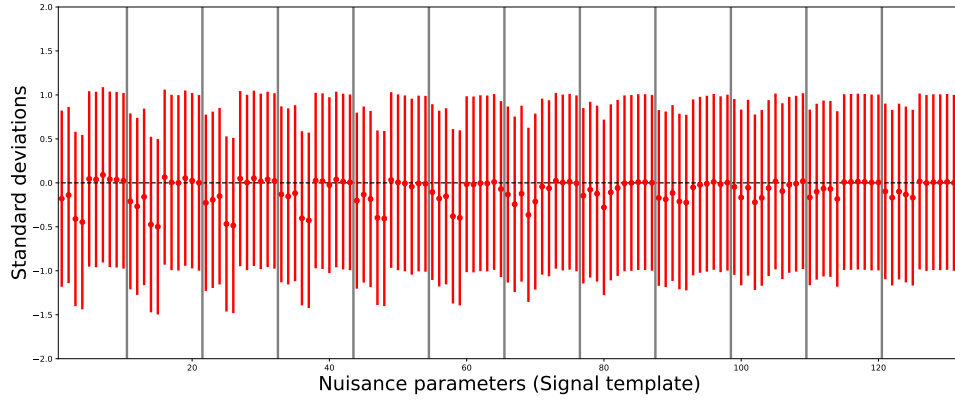


Figure 9.2: Projections of the fit variables (top) M_{miss}^2 in the full $M(\pi\pi)$ range and (bottom) $M(\pi\pi)$ in signal-rich regions of M_{miss}^2 , shown for the fit to data in the (left) $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ and (right) $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ reconstruction channels. The signal-rich regions correspond to -0.1 to $0.6 \text{ GeV}^2/c^4$ for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ and -0.4 to $0.4 \text{ GeV}^2/c^4$ for $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$.

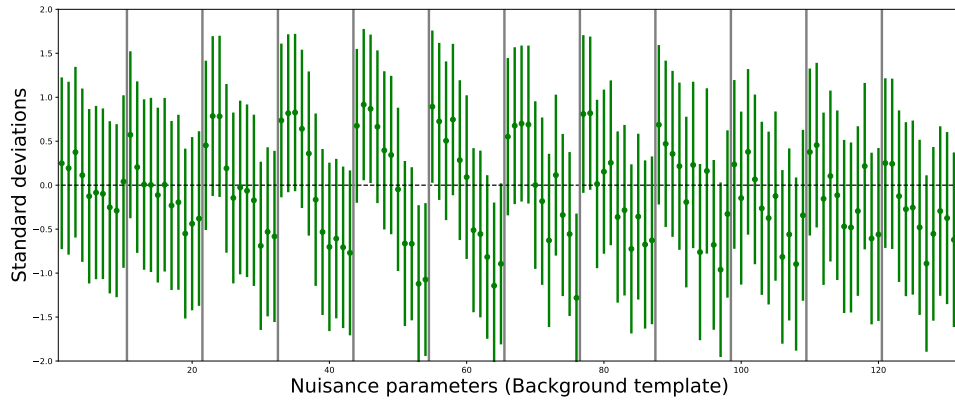
Table 9.2: Yields from reconstructed $B \rightarrow \rho \ell \nu_\ell$ decays obtained from maximum likelihood fits to 362 fb^{-1} of data, with uncertainties including template shape uncertainties determined via likelihood profiling with MINOS. The normalizations of the nonresonant templates as well as that of the resonant $B \rightarrow f_0(980) \ell \nu_\ell$ processes are at their limits so likelihood profiling cannot be used to estimate the uncertainties.

Process	$B^0 \rightarrow \rho^- \ell^+ \nu_\ell$	$B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$
Signal	$243.4^{+25.1}_{-24.0}$	$325.2^{+31.8}_{-31.4}$
Nonres. $B \rightarrow \pi \pi \ell \nu_\ell$	0.0	0.0
Background	$2688.0^{+56.8}_{-55.3}$	$1520.2^{+46.1}_{-48.0}$
Continuum	$73.9^{+5.3}_{-5.3}$	$94.6^{+5.2}_{-5.2}$
$B \rightarrow f_0(500) \ell \nu_\ell$		$61.8^{+25.8}_{-24.1}$
$B \rightarrow f_0(980) \ell \nu_\ell$		0.0
$B \rightarrow f_2(1270) \ell \nu_\ell$		$52.9^{+16.7}_{-41.6}$
$B \rightarrow \rho^0(1450) \ell \nu_\ell$		$7.0^{+70.3}_{-7.0}$

The bin-wise nuisance parameters of the signal and background templates are shown for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ in Fig. 9.3 and for $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ in Fig. 9.4. These nuisance parameter pulls on all other templates are small, illustrations of them can be found in Chapter C.

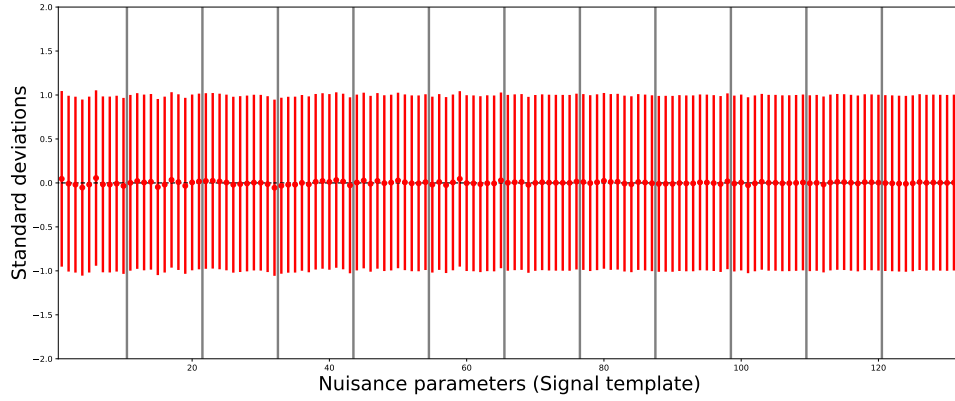


(a) Signal template

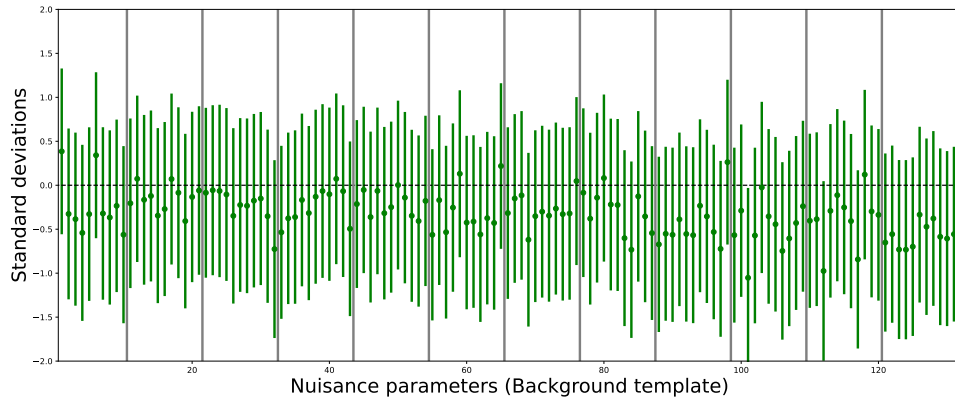


(b) Background template

Figure 9.3: Pulls on the 132 bin-wise, strongly correlated nuisance parameters for the signal and background templates in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ fit to data. The error bars correspond to the uncertainty determined in the fit. The solid vertical lines from top to bottom indicate bins in M_{miss}^2 . The signal template parameters exhibit pulls of up to 0.55σ in the third and fourth $M(\pi\pi)$ bin in each M_{miss}^2 bin due to the interference with ω mesons. The background template exhibits pulls with up to 1.28σ due to the uncertainty on the D meson decays and the large uncertainty from the MC sample size.



(a) Signal template



(b) Background template

Figure 9.4: Pulls on the 121 bin-wise, strongly correlated nuisance parameters for the signal and background templates in the $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ fit to data. The error bars correspond to the uncertainty determined in the fit. The solid vertical lines from top to bottom indicate bins in M_{miss}^2 . The signal template parameters exhibit no strong pulls. The background template parameters exhibit pulls with up to 1.05σ due to the uncertainty on the D meson decays and the large uncertainty from the MC sample size.

9.2 Branching Fraction Determination

The branching fractions are determined from the signal yields and a number of inputs using the relations

$$\mathcal{B}(B^0 \rightarrow \rho^- \ell^+ \nu_\ell) = \frac{N_{\text{sig}}^{\text{data}}(1 + f_{+0})}{4 \times N_{B\bar{B}} \times \epsilon} \text{ and} \quad (9.1)$$

$$\mathcal{B}(B^+ \rightarrow \rho^0 \ell^+ \nu_\ell) = \frac{N_{\text{sig}}^{\text{data}}(1 + f_{+0})}{4 \times N_{B\bar{B}} \times f_{+0} \times \epsilon}, \quad (9.2)$$

where $N_{\text{sig},i}^{\text{data}}$ is the signal yield obtained in the fit to data, f_{+0} is the ratio between the branching fractions of the decays of the $\Upsilon(4S)$ meson to pairs of charged and neutral B -mesons and $N_{B\bar{B}}$ is the number of B -meson pairs counted in the current data set, and ϵ is the reconstruction efficiency obtained from simulation. The factor of four present in the denominator accounts for the two B -mesons in the $\Upsilon(4S)$ decay and the reconstruction of both light lepton flavors. $N_{B\bar{B}}$ is determined in analogy to the method described in Ref. [17, Chapter 3] and f_{+0} is measured in Ref [129].

The measured branching fractions as well as the values of the input parameters are given in Table 9.3 where they are compared to the average performed by the PDG in Ref. [5] as well as the evaluation performed by the HFLAV in Ref. [16]. The former is a weighted average in which the uncertainty is increased to obtain reasonable agreement between the results while the latter is determined using fits to the differential measurements published by Belle and BaBar as well as isospin relations between ρ^0 and ρ^+ . Figure 9.5 illustrates the results compared to previous tagged and untagged determinations of the branching fractions. Good agreement is observed between the results and the PDG and HFLAV averages in both analysis channels.

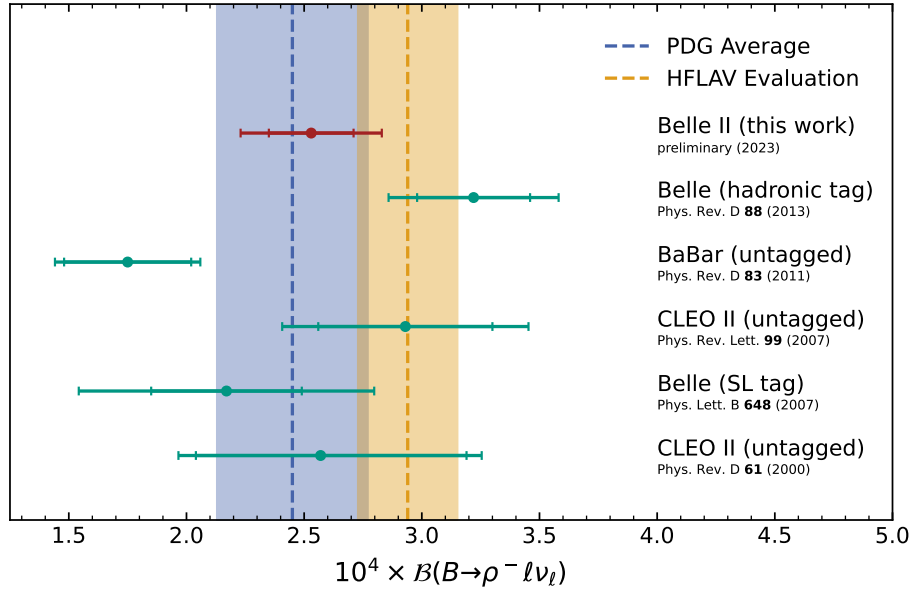
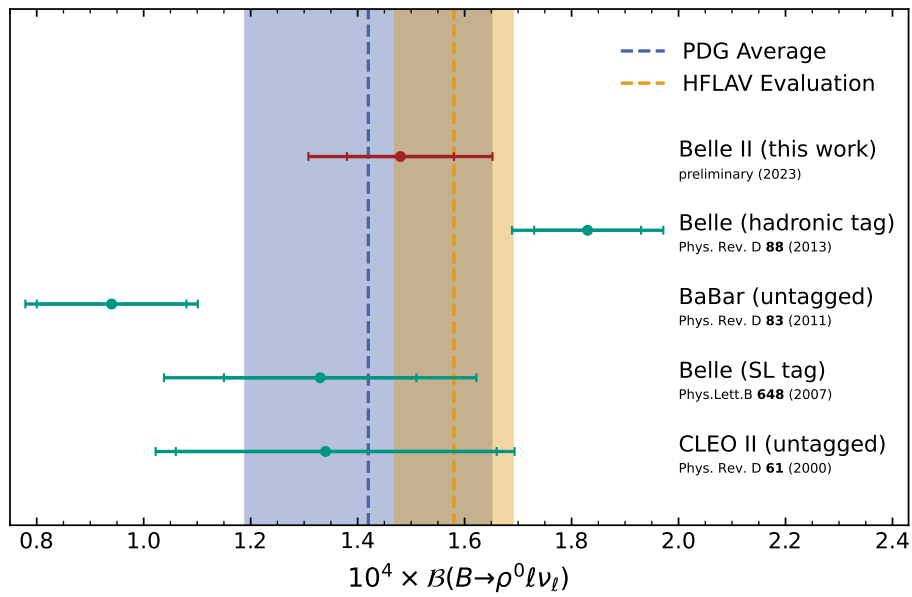
(a) $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ (b) $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$

Figure 9.5: The measured branching fractions of (top) $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ and (bottom) $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ compared to previous determinations, the PDG average and the HFLAV evaluation.

Table 9.3: Measured total branching fractions of $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays with statistical and systematic uncertainties, compared with world averages from the PDG [5] and the HFLAV [16]. For $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ the first uncertainty in the HFLAV evaluation is statistical and the second one is systematic, for the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ evaluation and the PDG average the given uncertainty combines both. The values of the parameters used in the measurement are also given with the uncertainties on $N_{\text{sig}}^{\text{data}}$ corresponding to the statistical and systematic uncertainty in the fit. The former is determined in a fit with fixed nuisance parameters, the latter is determined by subtracting in quadrature the statistical uncertainty from the value obtained with floating nuisance parameters. The uncertainty given for the efficiency incorporates all systematic effects without influence on the template shape.

	$B^0 \rightarrow \rho^- \ell^+ \nu_\ell$	$B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$
$N_{\text{sig}}^{\text{data}}$	$243.39 \pm 23.43(\text{stat}) \pm 6.39(\text{sys})$	$325.19 \pm 30.41(\text{stat}) \pm 10.08(\text{sys})$
f_{+0}		1.065 ± 0.052
$N_{B\bar{B}}$		$(387 \pm 6) \times 10^6$
ϵ	$(13.2 \pm 0.9) \times 10^4$	$(27.5 \pm 1.5) \times 10^4$
\mathcal{B}	$(2.46 \pm 0.24(\text{stat}) \pm 0.18(\text{syst})) \times 10^{-4}$	$(1.48 \pm 0.14(\text{stat}) \pm 0.10(\text{syst})) \times 10^{-4}$
$\mathcal{B}_{\text{HFLAV}}$	$(2.94 \pm 0.11 \pm 0.18) \times 10^{-4}$	$(1.58 \pm 0.11) \times 10^{-4}$
\mathcal{B}_{PDG}	$(2.45 \pm 0.32) \times 10^{-4}$	$(1.42 \pm 0.23) \times 10^{-4}$

Chapter 10

Conclusion

In this thesis I present branching fraction measurements of the processes $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ with a new tagging method. This is the first measurement of this process at Belle II.

To improve simulation and analysis workflows at Belle II, I also describe a caching workflow which points a way towards a more efficient use of computing resources by incorporating data transfer with XRootD and caching with XCache. This achieves a better distribution of computing workloads in the Belle II collaboration and reduces the amount of data transferred between computing sites.

The presented $B \rightarrow \rho \ell \nu_\ell$ analysis makes use of data collected between 2019 and 2022 at the $\Upsilon(4S)$ resonance and corresponding to an integrated luminosity of 362 fb^{-1} . In the analysis, events are fully reconstructed on the hadronic tag side and reconstructed from two pions and a lepton on the signal side.

A two-dimensional fit in the squared missing mass and the invariant dipion mass is conducted to distinguish both misreconstructed events and other processes with the $B \rightarrow \pi \pi \ell \nu_\ell$ final state from the signal processes. Shape-changing systematic uncertainties are incorporated in the signal extraction to maximally exploit the available information.

The measurements are the first determinations of these branching fractions at Belle II which find

$$\mathcal{B}(B^0 \rightarrow \rho^- \ell^+ \nu_\ell) = (2.46 \pm 0.24 \pm 0.18) \times 10^{-4}$$

and

$$\mathcal{B}(B^+ \rightarrow \rho^0 \ell^+ \nu_\ell) = (1.48 \pm 0.14 \pm 0.10) \times 10^{-4}$$

where the first uncertainty is statistical and the second uncertainty is systematic. These values are in good agreement with the combined evaluation of all relevant previous determinations performed by HFLAV [16].

Main areas of improvement of the simulated model and uncertainty treatment compared to previous determinations are improved descriptions of the $B \rightarrow X_c \ell \nu_\ell$ background and an inclusive model of the $B \rightarrow X_u \ell \nu_\ell$ background, updated signal modeling in both the hadronic

form factor and the resonant line shape. Additionally, an external constraint from Ref. [79] is used in the determination of $\mathcal{B}(B^+ \rightarrow \rho^0 \ell^+ \nu_\ell)$. This external input adds more control of as of yet unmeasured background processes with the same final state as the signal process. Sources of systematic uncertainty previously not considered or treated differently are the $B \rightarrow \rho \ell \nu_\ell$ form factor description, possible $\rho - \omega$ interference distorting the ρ line shape, D meson decay branching fractions in the background, and contributions from a possible nonresonant $B \rightarrow \pi \pi \ell \nu_\ell$ background.

In the coming years, Belle II will accumulate a large data set of decays in the $B \rightarrow \pi \pi \ell \nu_\ell$ final state. As the measurement is currently statistically dominated, this will significantly reduce the overall uncertainties seen today. In addition, systematic uncertainties are expected to shrink with larger data sets as subsidiary measurements at Belle II will become more precise. This will enable differential measurements of the $B \rightarrow \rho \ell \nu_\ell$ decay, allow further insight into the CKM matrix element $|V_{ub}|$, and fully determine the resonant structure of $B \rightarrow \pi \pi \ell \nu_\ell$ decays.

Danksagung

Als Erstes möchte ich mich bei Prof. Dr. Torben Ferber für die Übernahme des Referats und bei Prof. Dr. Günter Quast für die Übernahme des Korreferats bedanken. Beiden möchte ich sehr für die angenehme Atmosphäre am Institut, die fachliche Unterstützung und die stets offenen Türen (in denen ich häufig ratsuchend stand) danken. Ich habe mich, trotz der thematischen Verschiebung unserer Arbeitsgruppe, immer gut betreut und beraten gefühlt.

Als Nächstes möchte ich mich bei Dr. Pablo Goldenzweig bedanken, der mich in die Arbeitsgruppe aufgenommen hat, die gesamte Promotion betreut hat und bei allen Problemen jederzeit ein offenes Ohr hatte. Ohne ihn wäre diese Arbeit unmöglich gewesen und er hat durch viele Kommentare maßgeblich zum Gelingen der Arbeit beigetragen. Er ist es auch, der mir geholfen hat, einen Platz in der Belle-II-Kollaboration zu finden, und wusste wirklich immer, wie Licht ins Dunkel zu bringen ist, wenn sich neue, unverstandene Merkwürdigkeiten in der Analyse ergeben haben.

Des Weiteren möchte ich mich bei meinen Kolleg:innen aller Kollaborationen am Institut für Experimentelle Teilchenphysik (ETP) und in Belle II bedanken, die immer bereit waren, mit mir verschiedenste Ideen und, gerade zu Ende der Promotion, auch viele Probleme zu diskutieren. Hier möchte ich besonders meine Bürokollegen im Raum 9-11 des Physikhochhauses, zu Beginn Dr. Felix Metzner und Patrick Ecker sowie später Dr. Giacomo De Pietro, nennen, von denen ich viel über Physik, Programmieren und Belle II gelernt habe. Trotz oder vielleicht aufgrund unserer sehr unterschiedlichen Arbeitsweisen hatten wir zahlreiche interessante Gespräche, nicht immer fachlicher Art aber stets unterhaltsam und lehrreich. Alle Kolleg:innen am Institut haben dazu beigetragen, dass ich mich gefreut hatte nach der pandemiebedingten Zeit im Homeoffice wieder in das Büro zurückzukehren.

Besonders hervorheben möchte ich außerdem alle noch nicht bereits Erwähnten, die mit inhaltlichen und textlichen Vorschlägen zum Entstehen dieser Dissertation beigetragen haben. Zu nennen sind hier Alexander Heidelberg, Isabel Haide, Jonas Eppelt, Dr. Marco Link, Dr. Max Neukum, Maximilian Stemmer-Grabow, Noreen Rauls und Dr. Slavomira Stefkova. Vielen Dank für die Vielzahl an hilfreichen Vorschlägen, die den Text erheblich klarer und lesbarer gemacht haben.

Weiterhin möchte ich mich bei Dr. Markus Prim und Prof. Dr. Florian Bernlochner bedanken, die sowohl das Thema an mich herangetragen haben als auch wertvolle Vorschläge und Hinweise gegeben haben wenn die Analyse ins Stocken geriet.

Schlussendlich möchte ich mich bei meinen Eltern, meiner Schwester und insbesondere meiner Partnerin Anna bedanken. Sie haben mich während des Studiums und der Promotion immer unterstützt und (hoffentlich) gerne zugehört, wenn ich mit der Arbeit oder dem Thema kämpfte. Gerade in den letzten Monaten der Promotion war diese Unterstützung unverzichtbar und hat es mir erst ermöglicht die Doktorarbeit zu schreiben.

Bibliography

- [1] E. Kou et al. “The Belle II physics book”. In: *Prog. Theor. Exp. Phys.* 2019 (2019), p. 123C01. DOI: 10.1093/ptep/ptz106.
- [2] M. Kobayashi and T. Maskawa. “CP-violation in the renormalizable theory of weak interaction”. In: *Prog. Theor. Phys* 49 (1973), pp. 652–657. DOI: 10.1143/PTP.49.652.
- [3] N. Cabibbo. “Unitary symmetry and leptonic decays”. In: *Phys. Rev. Lett.* 10 (1963), pp. 531–533. DOI: 10.1103/PhysRevLett.10.531.
- [4] L.-L. Chau and W.-Y. Keung. “Comments on the parametrization of the Kobayashi-Maskawa matrix”. In: *Phys. Rev. Lett.* 53 (1984), pp. 1802–1805. DOI: 10.1103/PhysRevLett.53.1802.
- [5] **Particle Data Group**, R. L. Workman et al. “Review of particle physics”. In: *Prog. Theor. Exp. Phys.* 2022 (2022), p. 083C01. DOI: 10.1093/ptep/ptac097.
- [6] J. Charles et al. “CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories”. In: *Eur. Phys. J. C* 41 (2005), pp. 1–131. DOI: 10.1140/epjc/s2005-02169-1.
- [7] J. D. Richman and P. R. Burchat. “Leptonic and semileptonic decays of charm and bottom hadrons”. In: *Rev. Mod. Phys.* 67 (1995), pp. 893–976. DOI: 10.1103/RevModPhys.67.893.
- [8] I. I. Bigi, N. G. Uraltsev, and A. I. Vainshtein. “Nonperturbative corrections to inclusive beauty and charm decays. QCD versus phenomenological models”. In: *Phys. Lett. B* 293 (1992), pp. 430–436. DOI: 10.1016/0370-2693(92)90908-M.
- [9] I. I. Y. Bigi et al. “A QCD ‘manifesto’ on inclusive decays of beauty and charm”. In: *7th Meeting of the APS Division of Particles Fields.* 1992, pp. 610–613.
- [10] I. I. Bigi, M. Shifman, N. G. Uraltsev, and A. Vainshtein. “QCD predictions for lepton spectra in inclusive heavy flavor decays”. In: *Phys. Rev. Lett.* 71 (1993), pp. 496–499. DOI: 10.1103/PhysRevLett.71.496.
- [11] I. I. Bigi and N. G. Uraltsev. “Weak annihilation and the end-point spectrum in semileptonic B decays”. In: *Nucl. Phys. B* 423 (1994), pp. 33–55. DOI: 10.1016/0550-3213(94)90564-9.

- [12] M. Neubert. “B decays and the heavy-quark expansion”. In: *Heavy Flavours II*. Vol. Volume 15. Advanced Series on Directions in High Energy Physics Volume 15. WORLD SCIENTIFIC, 1998, pp. 239–293. ISBN: 978-981-02-2215-4. DOI: 10.1142/9789812812667_0003.
- [13] B. O. Lange, M. Neubert, and G. Paz. “Theory of charmless inclusive B decays and the extraction of V_{ub} ”. In: *Phys. Rev. D* 72 (2005), p. 073006. DOI: 10.1103/PhysRevD.72.073006.
- [14] F. D. Fazio and M. Neubert. “ $B \rightarrow X_u \ell \nu_\ell$ decay distributions to order α_s ”. In: *J. High Energy Phys.* 1999 (1999), p. 017. DOI: 10.1088/1126-6708/1999/06/017.
- [15] P. Gambino, P. Giordano, G. Ossola, and N. Uraltsev. “Inclusive semileptonic B decays and the determination of $|V_{ub}|$ ”. In: *J. High Energy Phys.* 2007 (2007), p. 058. DOI: 10.1088/1126-6708/2007/10/058.
- [16] **Heavy Flavor Averaging Group**, Y. Amhis et al. “Averages of b -hadron, c -hadron, and τ -lepton properties as of 2021”. In: *Phys. Rev. D* 107 (2023), p. 052008. DOI: 10.1103/PhysRevD.107.052008.
- [17] A. J. Bevan et al. “The physics of the B factories”. In: *Eur. Phys. J. C* 74 (2014), p. 3026. DOI: 10.1140/epjc/s10052-014-3026-9.
- [18] A. Bharucha, D. M. Straub, and R. Zwicky. “ $B \rightarrow V \ell^+ \ell^-$ in the Standard Model from light-cone sum rules”. In: *J. High Energy Phys.* 2016 (2016), p. 98. DOI: 10.1007/JHEP08(2016)098.
- [19] F. U. Bernlochner, M. T. Prim, and D. J. Robinson. “ $B \rightarrow \rho \ell \bar{\nu}$ and $\omega \ell \bar{\nu}$ in and beyond the Standard Model: improved predictions and $|V_{ub}|$ ”. In: *Phys. Rev. D* 104 (2021), p. 034032. DOI: 10.1103/PhysRevD.104.034032.
- [20] J. Dingfelder and T. Mannel. “Leptonic and semileptonic decays of B mesons”. In: *Rev. Mod. Phys.* 88 (2016), p. 035008. DOI: 10.1103/RevModPhys.88.035008.
- [21] **CLEO Collaboration**, J. P. Alexander et al. “First measurement of the $B \rightarrow \pi \ell \nu$ and $B \rightarrow \rho(\omega) \ell \nu$ branching fractions”. In: *Phys. Rev. Lett.* 77 (1996), pp. 5000–5004. DOI: 10.1103/PhysRevLett.77.5000.
- [22] **Belle II Collaboration**, T. Abe et al. *Belle II technical design report*. BELLE2-REPORT-2016-001. 2010. URL: <https://docs.belle2.org/record/329>.
- [23] D. M. J. Lovelock et al. “Masses, widths, and leptonic widths of the higher upsilon resonances”. In: *Phys. Rev. Lett.* 54 (1985), pp. 377–380. DOI: 10.1103/PhysRevLett.54.377.
- [24] D. Besson et al. “Observation of new structure in the e^+e^- cross section above the $\Upsilon(4S)$ ”. In: *Phys. Rev. Lett.* 54 (1985), pp. 381–384. DOI: 10.1103/PhysRevLett.54.381.

- [25] D. Besson and T. Skwarnicki. “Upsilon spectroscopy: Transitions in the bottomonium system”. In: *Annu. Rev. Nucl. Part. Sci.* 43 (1993), pp. 333–378. DOI: 10.1146/annurev.ns.43.120193.002001.
- [26] D. Cassel. “CLEO B Physics”. In: *ARGUS Symposium, DESY Auditorium, 09 November 2007*. ARGUS Fest 2007. URL: https://argus-fest.desy.de/e21/index_eng.html.
- [27] J. Iizuka. “A systematics and phenomenology of meson family”. In: *Progress of Theoretical Physics Supplement* 37–38 (1966), pp. 21–34. DOI: 10.1143/PTPS.37.21.
- [28] S. Okubo. “ φ -meson and unitary symmetry model”. In: *Phys. Lett.* 5 (1963), pp. 165–168. DOI: 10.1016/S0375-9601(63)92548-9.
- [29] G. Zweig. *An SU_3 model for strong interaction symmetry and its breaking; Version 2*. 1964. DOI: 10.17181/CERN-TH-412. preprint.
- [30] **LHCb Collaboration**, A. Augusto Alves Jr. et al. “The LHCb detector at the LHC”. In: *J. Inst.* 3 (2008), S08005. DOI: 10.1088/1748-0221/3/08/S08005.
- [31] L. Evans and P. Bryant. “LHC Machine”. In: *J. Inst.* 3 (2008), S08001. DOI: 10.1088/1748-0221/3/08/S08001.
- [32] **Belle Collaboration**, J. Brodzicka et al. “Physics achievements from the Belle experiment”. In: *Prog. Theor. Exp. Phys.* 2012 (2012), p. 04D001. DOI: 10.1093/ptep/pts072.
- [33] **SuperKEKB accelerator team**, K. Akai, K. Furukawa, and H. Koiso. “SuperKEKB collider”. In: *Nucl. Instrum. Methods Phys. Res. A* 907 (2018), pp. 188–199. DOI: 10.1016/j.nima.2018.08.017.
- [34] **SuperB Collaboration**, M. Baszczyk et al. *SuperB technical design report*. 2013. URL: <https://cds.cern.ch/record/1557673>.
- [35] Y. Ohnishi et al. “Accelerator design at SuperKEKB”. In: *Prog. Theor. Exp. Phys.* 2013 (2013), 03A011. DOI: 10.1093/ptep/pts083.
- [36] Y. Funakoshi et al. “The SuperKEKB has broken the world record of the luminosity”. In: 13th International Particle Accelerator Conference (IPAC’22), Bangkok, Thailand, 12-17 June 2022. JACOW Publishing, Geneva, Switzerland, 2022, pp. 1–5. ISBN: 978-3-95450-227-1. DOI: 10.18429/JACoW-IPAC2022-MOPLXGD1.
- [37] *Belle II Website / Archives*. URL: <https://www.belle2.org/archives/> Access date: 2023-09-03.
- [38] H. Ye et al. “Commissioning and performance of the Belle II pixel detector”. In: *Nucl. Instrum. Methods Phys. Res. A* 987 (2021), p. 164875. DOI: 10.1016/j.nima.2020.164875.
- [39] **Belle II SVD Collaboration**, K. Adamczyk et al. “The design, construction, operation and performance of the Belle II Silicon Vertex Detector”. In: *J. Inst.* 17 (2022), P11042. DOI: 10.1088/1748-0221/17/11/P11042.

- [40] N. Taniguchi. “Central Drift Chamber for Belle II”. In: *J. Inst.* 12 (2017), p. C06014. DOI: 10.1088/1748-0221/12/06/C06014.
- [41] P. Križan. “Particle identification at Belle II”. In: *J. Inst.* 9 (2014), p. C07018. DOI: 10.1088/1748-0221/9/07/C07018.
- [42] **Belle II Collaboration**, F. Abudinén et al. “Measurement of the integrated luminosity of the Phase 2 data of the Belle II experiment”. In: *Chin. Phys. C* 44 (2020), p. 021001. DOI: 10.1088/1674-1137/44/2/021001.
- [43] Belle II Collaboration. *Belle II Luminosity | Page Version 21*. URL: <https://confluence.desy.de/display/BI/Belle+II+Luminosity> Access date: 2023-08-14.
- [44] A. Tsaregorodtsev et al. *DIRACGrid/DIRAC: v7.2.50*. Zenodo, 2022. DOI: 10.5281/zenodo.7071472.
- [45] M. Barisits et al. “Rucio: Scientific Data Management”. In: *Comput. Softw. Big Sci.* 3 (2019), p. 11. DOI: 10.1007/s41781-019-0026-3.
- [46] M. Böhler et al. “Transparent integration of opportunistic resources into the WLCG compute infrastructure”. In: *EPJ Web Conf.* 251 (2021), p. 02039. DOI: 10.1051/epjconf/202125102039.
- [47] A. Dorigo, P. Elmer, F. Furano, and A. Hanushevsky. “XROOTD/TXNetFile: a highly scalable architecture for data access in the ROOT environment”. In: *Proceedings of the 4th WSEAS International Conference on Telecommunications and Informatics. TELE-INFO’05*. Stevens Point, Wisconsin, USA: World Scientific and Engineering Academy and Society (WSEAS), 2005, pp. 1–6. ISBN: 978-960-8457-11-9.
- [48] A. Hanushevsky, G. Ganis, and G. Amadio. *Home Page | XRootD*. Home Page | XRootD. 2023. URL: <https://xrootd.slac.stanford.edu/index.html> Access date: 2023-02-27.
- [49] **CMS Collaboration**, L. A. T. Bauerdick et al. “XRootd, disk-based, caching proxy for optimization of data access, data placement and data replication”. In: *J. Phys.: Conf. Ser.* 513 (2014), p. 042044. DOI: 10.1088/1742-6596/513/4/042044.
- [50] T. Li, R. Currie, and A. Washbrook. “A data caching model for Tier 2 WLCG computing centres using XCache”. In: *EPJ Web Conf.* 214 (2019), p. 04047. DOI: 10.1051/epjconf/201921404047.
- [51] S. A. Weil et al. “Ceph: a scalable, high-performance distributed file system”. In: *Proceedings of the 7th Symposium on Operating Systems Design and Implementation. OSDI ’06*. USA: USENIX Association, 2006, pp. 307–320. ISBN: 978-1-931971-47-8.
- [52] M. Fischer. *Xrootdlib - Tools for working with the XRootD middleware*. 2022. URL: <https://github.com/maxfischer2781/xrootdlib> Access date: 2023-02-23.

- [53] Elastic. *Elasticsearch*. Elastic NV, 2023. URL: <https://github.com/elastic/elasticsearch> Access date: 2023-02-23.
- [54] H. Voss. “Classification”. In: *Data Analysis in High Energy Physics*. John Wiley & Sons, Ltd, 2013, pp. 153–186. ISBN: 978-3-527-65341-6. DOI: 10.1002/9783527653416.ch5.
- [55] L. Breiman. “Bagging predictors”. In: *Mach. Learn.* 24 (1996), pp. 123–140. DOI: 10.1007/BF00058655.
- [56] Y. Freund and R. E. Schapire. “A decision-theoretic generalization of on-line learning and an application to boosting”. In: *J. Comput. Syst. Sci.* 55 (1997), pp. 119–139. DOI: 10.1006/jcss.1997.1504.
- [57] J. H. Friedman. “Greedy function approximation: A gradient boosting machine.” In: *Ann. Stat.* 29 (2001), pp. 1189–1232. DOI: 10.1214/aos/1013203451.
- [58] J. H. Friedman. “Stochastic gradient boosting”. In: *Comput. Stat. Data Anal.* 38 (2002), pp. 367–378. DOI: 10.1016/S0167-9473(01)00065-2.
- [59] T. Keck. “FastBDT: a speed-optimized multivariate classification algorithm for the Belle II experiment”. In: *Comput. Softw. Big Sci.* 1 (2017), p. 2. DOI: 10.1007/s41781-017-0002-8.
- [60] Belle II Collaboration. *Belle II Analysis Software Framework (basf2)*. Zenodo, 2022. DOI: 10.5281/zenodo.6949513.
- [61] **Belle II Framework Software Group**, T. Kuhr et al. “The Belle II core software”. In: *Comput. Softw. Big Sci.* 3 (2018), p. 1. DOI: 10.1007/s41781-018-0017-9.
- [62] M. Milesi, J. Tan, and P. Urquijo. “Lepton identification in Belle II using observables from the electromagnetic calorimeter and precision trackers”. In: *EPJ Web Conf.* 245 (2020), p. 06023. DOI: 10.1051/epjconf/202024506023.
- [63] J.-F. Krohn et al. “Global decay chain vertex fitting at Belle II”. In: *Nucl. Instrum. Methods Phys. Res. A* 976 (2020), p. 164269. DOI: 10.1016/j.nima.2020.164269.
- [64] W. D. Hulsbergen. “Decay chain fitting with a Kalman filter”. In: *Nucl. Instrum. Methods Phys. Res. A* 552 (2005), pp. 566–575. DOI: 10.1016/j.nima.2005.06.078.
- [65] M. Feindt et al. “A hierarchical NeuroBayes-based algorithm for full reconstruction of B mesons at B factories”. In: *Nucl. Instrum. Methods Phys. Res. A* 654 (2011), pp. 432–440. DOI: 10.1016/j.nima.2011.06.008.
- [66] **Belle Collaboration**, A. Sibidanov et al. “Study of exclusive $B \rightarrow X_u \ell \nu$ decays and extraction of $|V_{ub}|$ using full reconstruction tagging at the Belle experiment”. In: *Phys. Rev. D* 88 (2013), p. 032005. DOI: 10.1103/PhysRevD.88.032005.
- [67] T. Keck et al. “The Full Event Interpretation”. In: *Comput. Softw. Big Sci.* 3 (2019), p. 6. DOI: 10.1007/s41781-019-0021-8.

- [68] J. Tanaka. “Precise measurements of charm meson lifetimes and search for D^0 - \bar{D}^0 mixing”. PhD thesis. University of Tokyo, 2002. URL: <https://inspirehep.net/literature/565918> Access date: 2023-09-06.
- [69] **Geant4 Collaboration**, S. Agostinelli et al. “Geant4—a simulation toolkit”. In: *Nucl. Instrum. Methods Phys. Res. A* 506 (2003), pp. 250–303. DOI: 10.1016/S0168-9002(03)01368-8.
- [70] D. J. Lange. “The EvtGen particle decay simulation package”. In: *Nucl. Instrum. Methods Phys. Res. A* 462 (2001), pp. 152–155. DOI: 10.1016/S0168-9002(01)00089-4.
- [71] A. Ryd, D. Lange, and N. Kuznetsova. *EvtGen User Guide*. URL: <https://evtgen.hepforge.org/doc/EvtGenGuide.pdf> Access date: 2023-09-18.
- [72] J. M. Blatt and V. F. Weisskopf. *Theoretical nuclear physics*. New York: Springer, 1952. ISBN: 978-0-471-08019-0. DOI: 10.1007/978-1-4612-9959-2.
- [73] F. von Hippel and C. Quigg. “Centrifugal-barrier effects in resonance partial decay widths, shapes, and production amplitudes”. In: *Phys. Rev. D* 5 (1972), pp. 624–638. DOI: 10.1103/PhysRevD.5.624.
- [74] F. Bernlochner. *Private communication on $\rho - \omega - \pi\pi$ interference and the ρ line shape*. E-mail. 2023.
- [75] H. B. O’Connell, B. C. Pearce, A. W. Thomas, and A. G. Williams. “Rho-omega mixing, vector meson dominance and the pion form-factor”. In: *Prog. Part. Nucl. Phys* 39 (1997), pp. 201–252. DOI: 10.1016/S0146-6410(97)00044-6.
- [76] **Fermilab Lattice and MILC collaborations**, J. A. Bailey et al. “ $|V_{ub}|$ from $B \rightarrow \pi\ell\nu$ decays and (2+1)-flavor lattice QCD”. In: *Phys. Rev. D* 92 (2015), p. 014024. DOI: 10.1103/PhysRevD.92.014024.
- [77] G. Duplančić and B. Melić. “Form factors of $B, B_s \rightarrow \eta^{(\prime)}$ and $D, D_s \rightarrow \eta^{(\prime)}$ transitions from QCD light-cone sum rules”. In: *J. High Energy Phys.* 2015 (2015), p. 138. DOI: 10.1007/JHEP11(2015)138.
- [78] D. Scora and N. Isgur. “Semileptonic meson decays in the quark model: An update”. In: *Phys. Rev. D* 52 (1995), pp. 2783–2812. DOI: 10.1103/PhysRevD.52.2783.
- [79] **Belle Collaboration**, C. Beleño et al. “Measurement of the branching fraction of the decay $B^+ \rightarrow \pi^+\pi^-\ell^+\nu_\ell$ in fully reconstructed events at Belle”. In: *Phys. Rev. D* 103 (2021), p. 112001. DOI: 10.1103/PhysRevD.103.112001.
- [80] C. Bierlich et al. *A comprehensive guide to the physics and usage of PYTHIA 8.3*. 2022. DOI: 10.48550/arXiv.2203.11601. preprint.
- [81] **Belle Collaboration**, L. Cao et al. “Measurements of partial branching fractions of inclusive $B \rightarrow X_u\ell^+\nu_\ell$ decays with hadronic tagging”. In: *Phys. Rev. D* 104 (2021), p. 012008. DOI: 10.1103/PhysRevD.104.012008.

- [82] C. Ramirez, J. F. Donoghue, and G. Burdman. “Semileptonic $b \rightarrow u$ decay”. In: *Phys. Rev. D* 41 (1990), pp. 1496–1503. DOI: 10.1103/PhysRevD.41.1496.
- [83] **Belle Collaboration**, M. T. Prim et al. “Search for $B^+ \rightarrow \mu^+ \nu_\mu$ and $B^+ \rightarrow \mu^+ N$ with inclusive tagging”. In: *Phys. Rev. D* 101 (2020), p. 032007. DOI: 10.1103/PhysRevD.101.032007.
- [84] F. G. K. Metzner. “Preparation of a measurement of $R(D^{(*)})$ with leptonic τ and hadronic FEI tag at the Belle experiment”. PhD thesis. Karlsruhe Institut für Technologie (KIT) / Karlsruhe Institut für Technologie (KIT), 2022. 188 pp. DOI: 10.5445/IR/1000148812.
- [85] Belle II Collaboration. *Correction of $BF(B \rightarrow X\ell\nu)$ in generic MC | page version 142*. URL: <https://confluence.desy.de/pages/viewpage.action?pageId=202394372> Access date: 2023-08-14.
- [86] D. Ferlewicz, P. Urquijo, and E. Waheed. “Revisiting fits to $B^0 \rightarrow D^{*-} \ell^+ \nu_\ell$ to measure $|V_{cb}|$ with novel methods and preliminary LQCD data at nonzero recoil”. In: *Phys. Rev. D* 103 (2021), p. 073005. DOI: 10.1103/PhysRevD.103.073005.
- [87] **Belle Collaboration**, R. Glattauer et al. “Measurement of the decay $B \rightarrow D\ell\nu_\ell$ in fully reconstructed events and determination of the Cabibbo-Kobayashi-Maskawa matrix element $|V_{cb}|$ ”. In: *Phys. Rev. D* 93 (2016), p. 032006. DOI: 10.1103/PhysRevD.93.032006.
- [88] R. A. Briere and E. R. Oxford. *Update of charm decay tables*. Belle II Note BELLE2-NOTE-PH-2020-008. Belle II, 2020.
- [89] **Mark III Collaboration**, J. Adler et al. “Resonant substructure in $K\pi\pi$ decays of charmed D mesons”. In: *Phys. Lett. B* 196 (1987), pp. 107–112. DOI: 10.1016/0370-2693(87)91685-6.
- [90] J. C. Anjos et al. “Dalitz plot analysis of $D \rightarrow K\pi\pi$ decays”. In: *Phys. Rev. D* 48 (1993), pp. 56–62. DOI: 10.1103/PhysRevD.48.56.
- [91] **CLEO Collaboration**, P. del Amo Sanchez et al. “Dalitz plot analysis of $D_s^+ \rightarrow K^+ K^- \pi^+$ ”. In: *Phys. Rev. D* 83 (2011), p. 052001. DOI: 10.1103/PhysRevD.83.052001.
- [92] **BABAR Collaboration**, B. Aubert et al. “Measurement of CP violation parameters with a Dalitz plot analysis of $B^\pm \rightarrow D_{\pi^+\pi^-\pi^0} K^\pm$ ”. In: *Phys. Rev. Lett.* 99 (2007), p. 251801. DOI: 10.1103/PhysRevLett.99.251801.
- [93] **CLEO Collaboration**, G. Bonvicini et al. “Dalitz plot analysis of the $D^+ \rightarrow K^- \pi^+ \pi^+$ decay”. In: *Phys. Rev. D* 78 (2008), p. 052001. DOI: 10.1103/PhysRevD.78.052001.
- [94] **CLEO Collaboration**, P. Rubin et al. “Search for CP violation in the Dalitz-plot analysis of $D^\pm \rightarrow K^+ K^- \pi^\pm$ ”. In: *Phys. Rev. D* 78 (2008), p. 072003. DOI: 10.1103/PhysRevD.78.072003.

- [95] **Fermilab E791 Collaboration**, E. M. Aitala et al. “Experimental evidence for a light and broad scalar resonance in $D^+ \rightarrow \pi^- \pi^+ \pi^+$ decay”. In: *Phys. Rev. Lett.* 86 (2001), pp. 770–774. DOI: 10.1103/PhysRevLett.86.770.
- [96] **Fermilab E791 Collaboration**, E. M. Aitala et al. “Study of the $D_s^+ \rightarrow \pi^- \pi^+ \pi^+$ decay and measurement of f_0 masses and widths”. In: *Phys. Rev. Lett.* 86 (2001), pp. 765–769. DOI: 10.1103/PhysRevLett.86.765.
- [97] **FOCUS Collaboration**, J. M. Link et al. “Study of the doubly and singly Cabibbo suppressed decays $D^+ \rightarrow K^+ \pi^+ \pi^-$ and $D_s^+ \rightarrow K^+ \pi^+ \pi^-$ ”. In: *Phys. Lett. B* 601 (2004), pp. 10–19. DOI: 10.1016/j.physletb.2004.09.022.
- [98] **BABAR Collaboration**, B. Aubert. *Dalitz plot analysis of D^0 hadronic decays $D^0 \rightarrow K^0 K^- \pi^+$, $D^0 \rightarrow \bar{K}^0 K^+ \pi^-$ and $D^0 \rightarrow \bar{K}^0 K^+ K^-$* . 2002. DOI: 10.48550/arXiv.hep-ex/0207089. preprint.
- [99] **CLEO Collaboration**, D. M. Asner et al. “Hadronic structure in the decay $\tau^- \rightarrow \nu_\tau \pi^- \pi^0 \pi^0$ and the sign of the tau neutrino helicity”. In: *Phys. Rev. D* 61 (1999), p. 012002. DOI: 10.1103/PhysRevD.61.012002.
- [100] S. Jadach, B. F. L. Ward, and Z. Wąs. “The precision Monte Carlo event generator KK for two-fermion final states in e^+e^- collisions”. In: *Comput. Phys. Commun.* 130 (2000), pp. 260–325. DOI: 10.1016/S0010-4655(00)00048-5.
- [101] S. Jadach, J. H. Kühn, and Z. Wąs. “TAUOLA - a library of Monte Carlo programs to simulate decays of polarized τ leptons”. In: *Comput. Phys. Commun.* 64 (1991), pp. 275–299. DOI: 10.1016/0010-4655(91)90038-M.
- [102] T. Ferber and P. Urquijo. *Overview of the Belle II physics generators*. Belle II Note BELLE2-NOTE-PH-2015-006. Belle II, 2015.
- [103] V. Bertacchi et al. “Track finding at Belle II”. In: *Comput. Phys. Commun.* 259 (2021), p. 107610. DOI: 10.1016/j.cpc.2020.107610.
- [104] M. Röhrken. *Time-Dependent CP Violation Measurements: Analyses of Neutral B Meson to Double-Charm Decays at the Japanese Belle Experiment*. Springer Theses. Cham: Springer International Publishing, 2014. ISBN: 978-3-319-00725-0 978-3-319-00726-7. DOI: 10.1007/978-3-319-00726-7.
- [105] **CLEO Collaboration**, D. M. Asner et al. “Search for exclusive charmless hadronic B decays”. In: *Phys. Rev. D* 53 (1996), pp. 1039–1050. DOI: 10.1103/PhysRevD.53.1039.
- [106] G. C. Fox and S. Wolfram. “Observables for the analysis of event shapes in e^+e^- annihilation and other processes”. In: *Phys. Rev. Lett.* 41 (1978), pp. 1581–1585. DOI: 10.1103/PhysRevLett.41.1581.
- [107] R. A. Fisher. “The use of multiple measurements in taxonomic problems”. In: *Ann. Eugen.* 7 (1936), pp. 179–188. DOI: 10.1111/j.1469-1809.1936.tb02137.x.

- [108] **Belle II Lepton ID Group**, M. Milesi, D. Ferlewicz, P. Urquijo, and M. Hohmann. *Muon and electron identification performance with 189 fb^{-1} of Belle II data*. Belle II Note BELLE2-NOTE-TE-2021-011. 2021.
- [109] F. Abudinen et al. *Impact of photon-energy biases on candidate-reconstruction variables*. Belle II Note BELLE2-NOTE-TE-2020-011. Belle II, 2020.
- [110] W. Sutcliffe and F. Bernlochner. *Calibration of hadronic tagging with $B \rightarrow X\ell\nu$ decays with the LS1 dataset*. Belle II Note BELLE2-NOTE-PH-2023-008. Belle II, 2023.
- [111] M. Liu, N. Rout, K. Trabelsi, and V. S. Vobbiliseti. *Hadronic FEI calibration with $D^{(*)}\pi$ samples*. Belle II Note BELLE2-NOTE-PH-2023-004. Belle II, 2023.
- [112] T. Skwarnicki. “A study of the radiative CASCADE transitions between the Upsilon-Prime and Upsilon resonances”. PhD thesis. Cracow, INP, 1986.
- [113] W. Sutcliffe et al. *Combination of calibration factors for hadronic tagging*. Belle II Note BELLE2-NOTE-PH-2023-029. Belle II, 2023.
- [114] C. Hainje, A. Albert, C. Hadjivasiliou, and J. Strube. *A comprehensive study of Belle II particle identification performance, efficiencies, and detector effects*. Belle II Note BELLE2-NOTE-TE-2021-022. 2021.
- [115] G. Cowan. *Statistical data analysis*. Oxford, New York: Oxford University Press, 1998. 212 pp. ISBN: 978-0-19-850155-8.
- [116] O. Behnke and L. Moneta. “Parameter estimation”. In: *Data Analysis in High Energy Physics*. John Wiley & Sons, Ltd, 2013, pp. 27–73. ISBN: 978-3-527-65341-6. DOI: 10.1002/9783527653416.ch2.
- [117] J. F. Kenney and E. S. Keeping. *Mathematics of statistics*. Vol. 2. 1961. ISBN: 978-1-133-09235-3.
- [118] S. S. Wilks. “The large-sample distribution of the likelihood ratio for testing composite hypotheses”. In: *Ann. Math. Stat.* 9 (1938), pp. 60–62. DOI: 10.1214/aoms/1177732360.
- [119] F. James and M. Roos. “Minuit - a system for function minimization and analysis of the parameter errors and correlations”. In: *Comput. Phys. Commun.* 10 (1975), pp. 343–367. DOI: 10.1016/0010-4655(75)90039-9.
- [120] H. Dembinski et al. *Scikit-hep/iminuit*. Version v2.24.0. Zenodo, 2023. DOI: 10.5281/ZENODO.3949207.
- [121] F. Metzner. *TemplateFitter*. 2022. URL: <https://github.com/FelixMetzner/TemplateFitter> Access date: 2023-08-30.
- [122] **Belle II Collaboration**, F. Abudinén et al. *Exclusive $B \rightarrow X_u\ell\nu_\ell$ decays with hadronic full-event-interpretation tagging in 62.8 fb^{-1} of Belle II data*. 2021. DOI: 10.48550/arXiv.2111.00710. preprint.

- [123] **Belle II Collaboration**, F. Abudinén et al. *Reconstruction of $B \rightarrow \rho \ell \nu_\ell$ decays identified using hadronic decays of the recoil B meson in 2019 – 2021 Belle II data*. 2022. DOI: 10.48550/arXiv.2211.15270. preprint.
- [124] G. Cowan, K. Cranmer, E. Gross, and O. Vitells. “Asymptotic formulae for likelihood-based tests of new physics”. In: *Eur. Phys. J. C* 71 (2011), p. 1554. DOI: 10.1140/epjc/s10052-011-1554-0.
- [125] **SND Collaboration**, M. N. Achasov et al. “Measurement of the $e^+e^- \rightarrow \pi^+\pi^-$ process cross section with the SND detector at the VEPP-2000 collider in the energy region $0.525 < \sqrt{s} < 0.883$ GeV”. In: *J. High Energy Phys.* 2021 (2021), p. 113. DOI: 10.1007/JHEP01(2021)113.
- [126] **BABAR Collaboration**, J. P. Lees et al. “Precise measurement of the $e^+e^- \rightarrow \pi^+\pi^-(\gamma)$ cross section with the initial-state radiation method at BABAR”. In: *Phys. Rev. D* 86 (2012), p. 032013. DOI: 10.1103/PhysRevD.86.032013.
- [127] **Particle Data Group**, C. Amsler et al. “Review of particle physics”. In: *Phys. Lett. B* 667 (2008), pp. 1–6. DOI: 10.1016/j.physletb.2008.07.018.
- [128] A. Glazov et al. *Measurement of the tracking efficiency in Phase 3 data using tau-pair events*. Belle II Note BELLE2-NOTE-PH-2020-006. Belle II, 2020.
- [129] **Belle Collaboration**, S. Choudhury et al. “Measurement of the B^+/B^0 production ratio in e^+e^- collisions at the $\Upsilon(4S)$ resonance using $B \rightarrow J/\psi(\ell\ell)K$ decays at Belle”. In: *Phys. Rev. D* 107 (2023), p. L031102. DOI: 10.1103/PhysRevD.107.L031102.
- [130] K. Berkelman. *A personal history of CESR and CLEO: The cornell electron storage ring and its main particle detector facility*. WORLD SCIENTIFIC, 2004. ISBN: 978-981-238-697-7 978-981-256-724-6. DOI: 10.1142/5426.

Appendices

Appendix A

Two-dimensional Projections of the Fit Templates with Bin-wise Uncertainties

A.0.1 $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ Templates

Detailed visualizations of the four templates in the $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ signal extraction are given in Figs. A.1 to A.4 with the normalization set to the expectation from simulated data. Relative uncertainties are given in each bin, the top value indicates the uncertainty from the limited size of the simulated sample, the bottom uncertainty the uncertainty from other systematic effects. The text `nan` indicates bins in which no events are expected for which no *relative* uncertainty can be given.

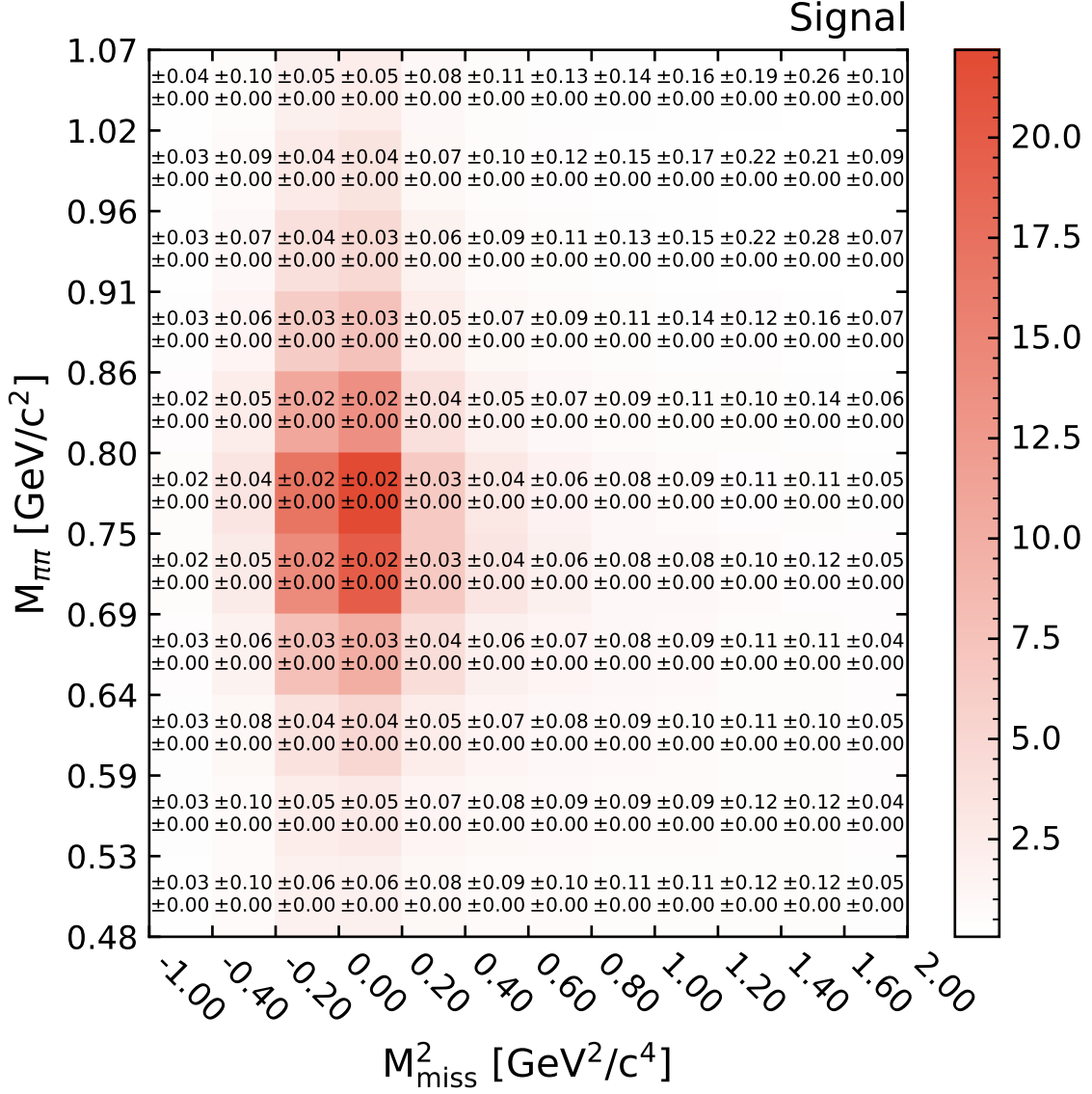


Figure A.1: Two-dimensional projection of the signal template in the $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ signal extraction with relative systematic uncertainties in each bin. Further description is given in the text.

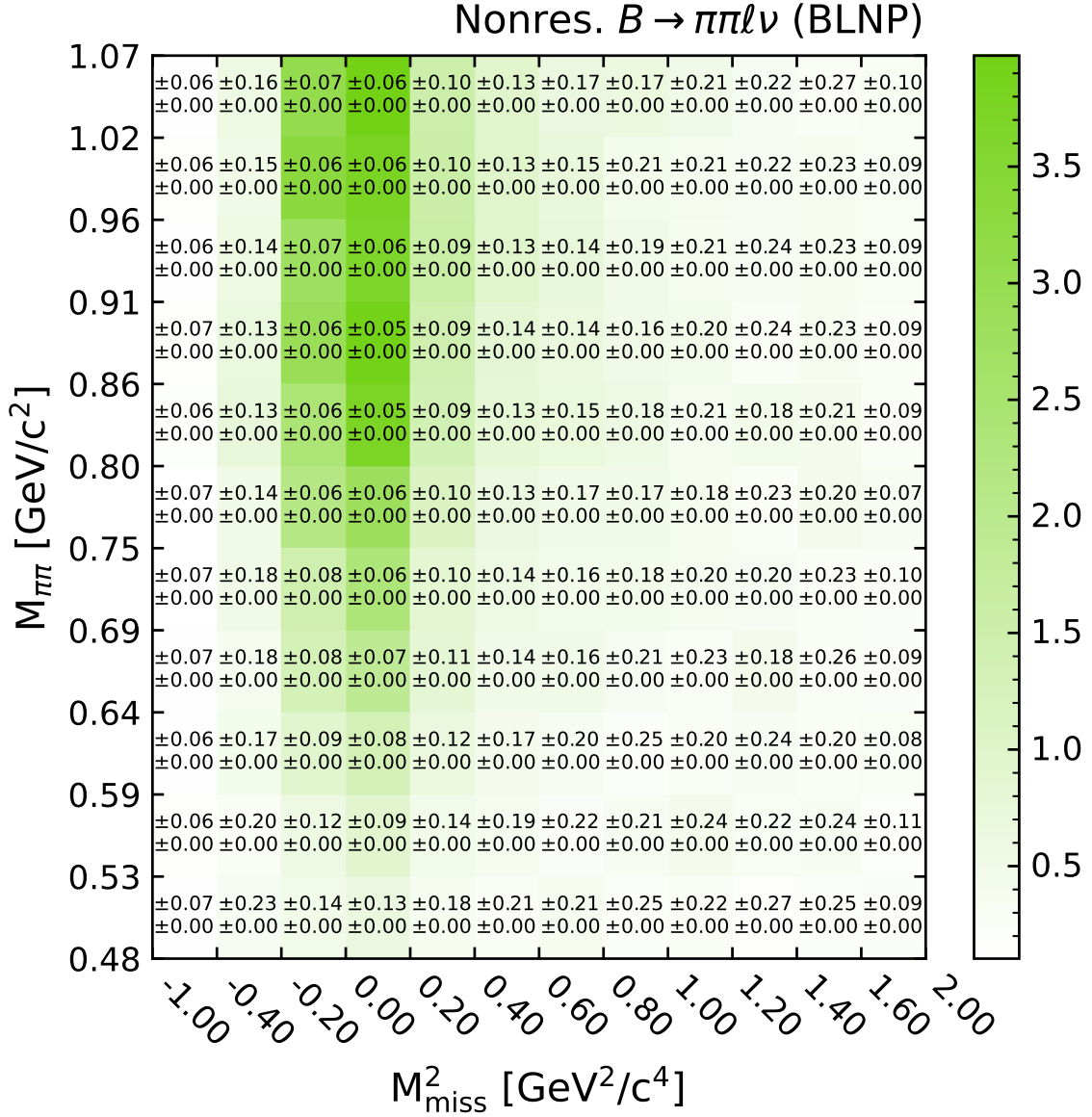


Figure A.2: Two-dimensional projection of the nonresonant $B \rightarrow \pi\pi\ell\nu$ template in the $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ signal extraction with relative systematic uncertainties in each bin. Further description is given in the text.

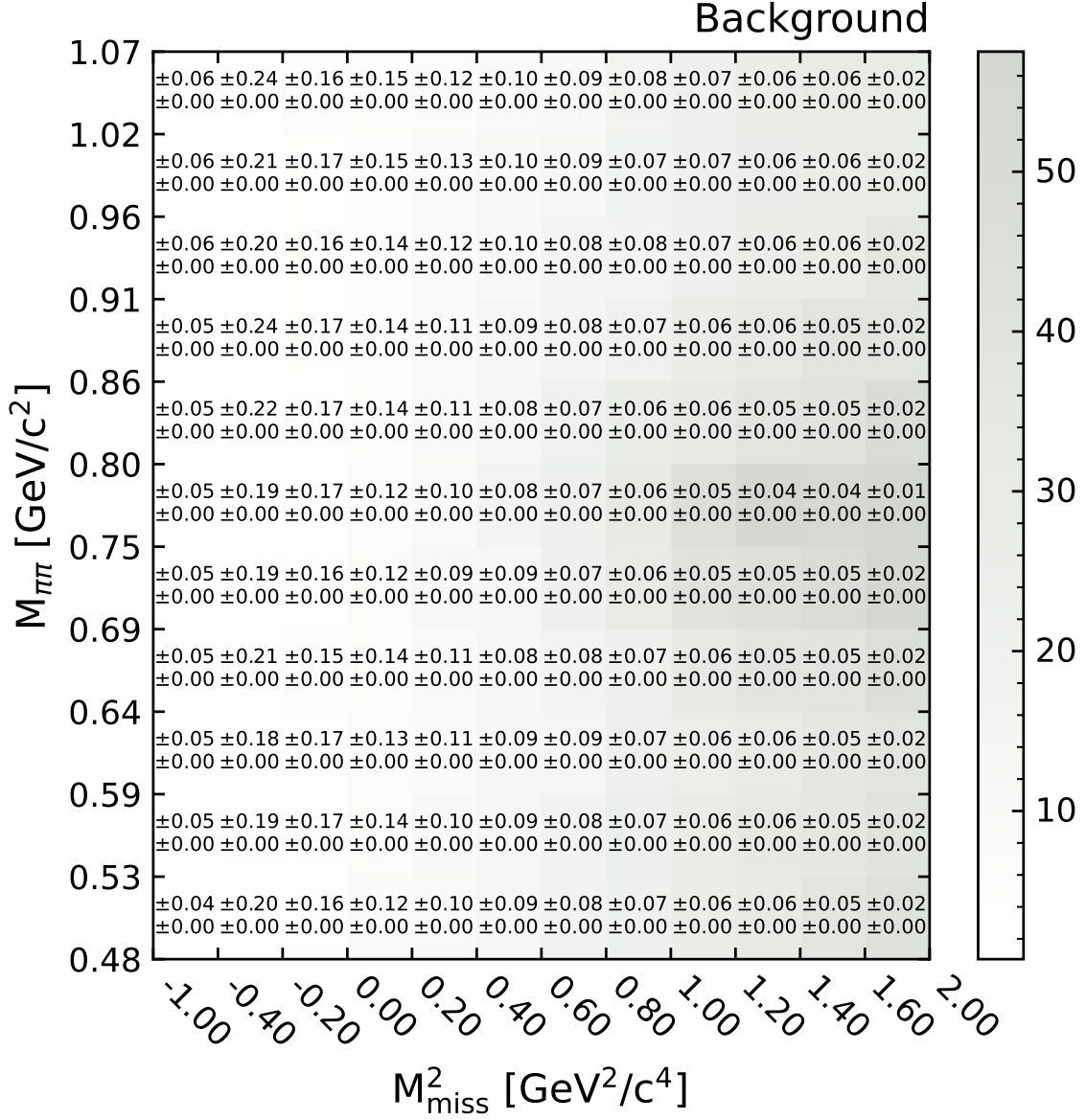


Figure A.3: Two-dimensional projection of the combined background template in the $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ signal extraction with relative systematic uncertainties in each bin. Further description is given in the text.

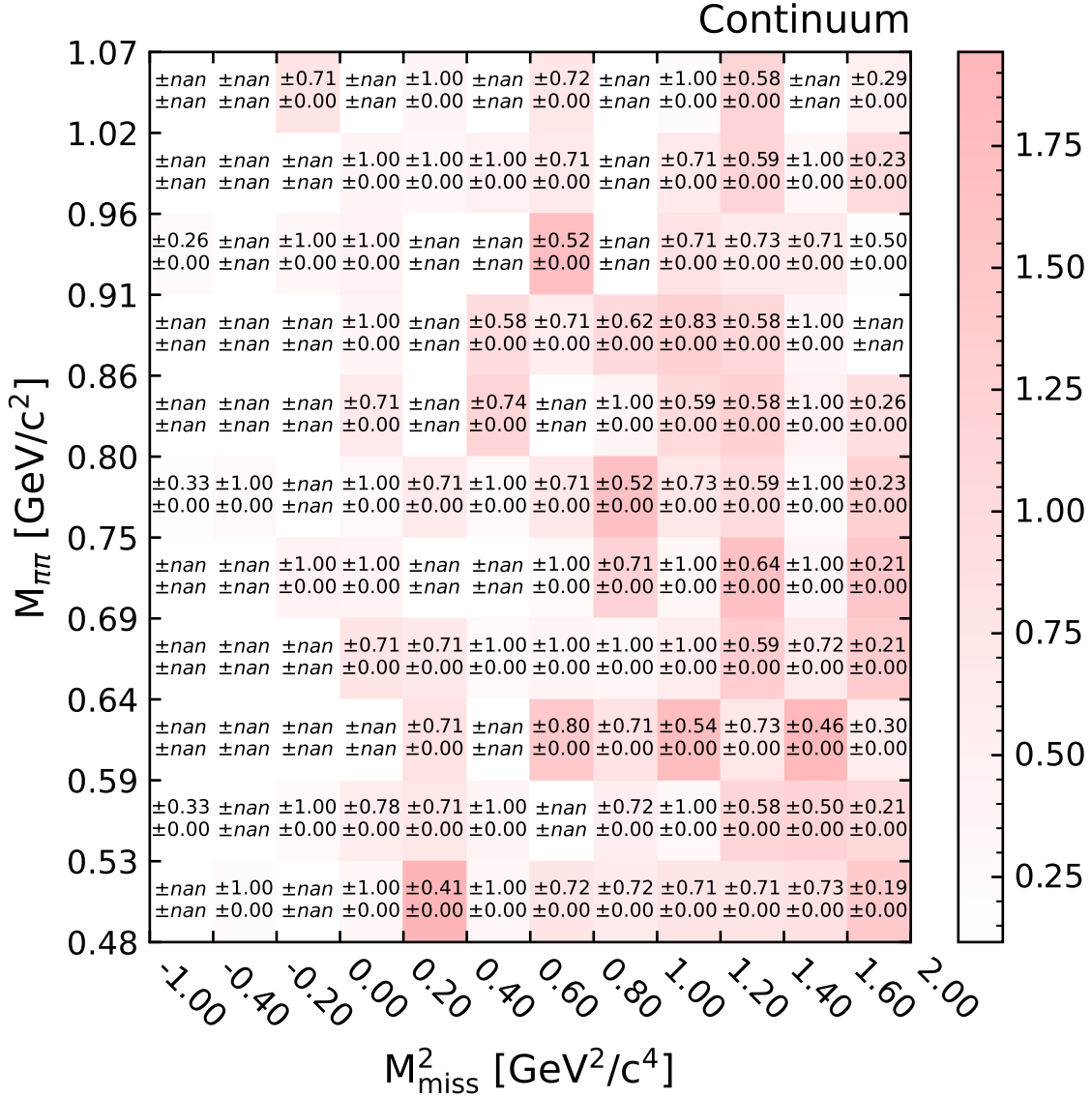


Figure A.4: Two-dimensional projection of the continuum template in the $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ signal extraction with relative systematic uncertainties in each bin. Further description is given in the text.

A.0.2 $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ Templates

Detailed visualizations of the four templates in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ signal extraction are given in Figs. A.5 to A.8 with the normalization set to the expectation from simulated data. Relative uncertainties are given in each bin, the top value indicates the uncertainty from the limited size of the simulated sample, the bottom uncertainty the uncertainty from other systematic effects. The text `nan` indicates bins in which no events are expected for which no *relative* uncertainty can be given.

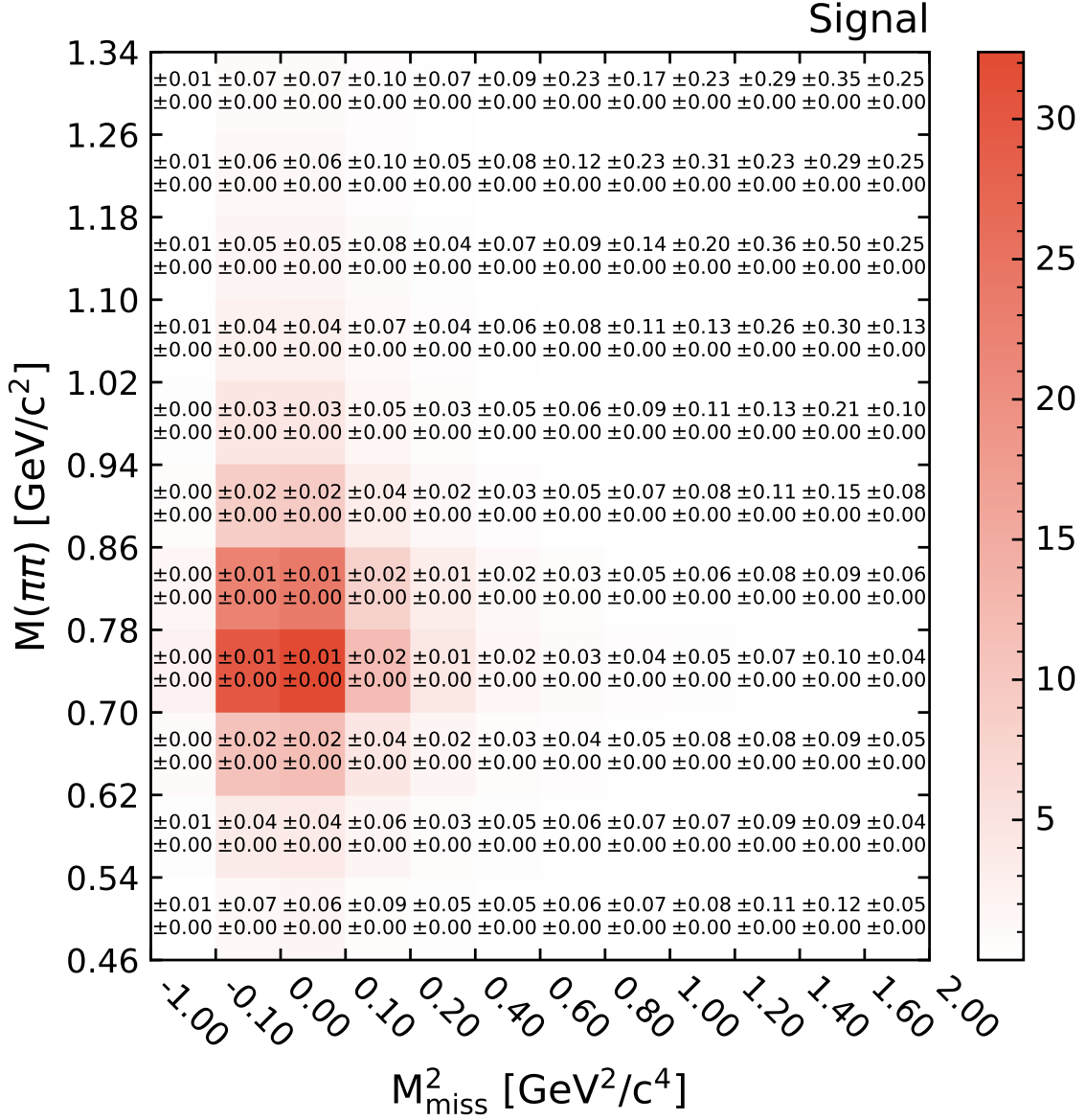


Figure A.5: Two-dimensional projection of the signal template in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ signal extraction with relative systematic uncertainties in each bin. Further description is given in the text.

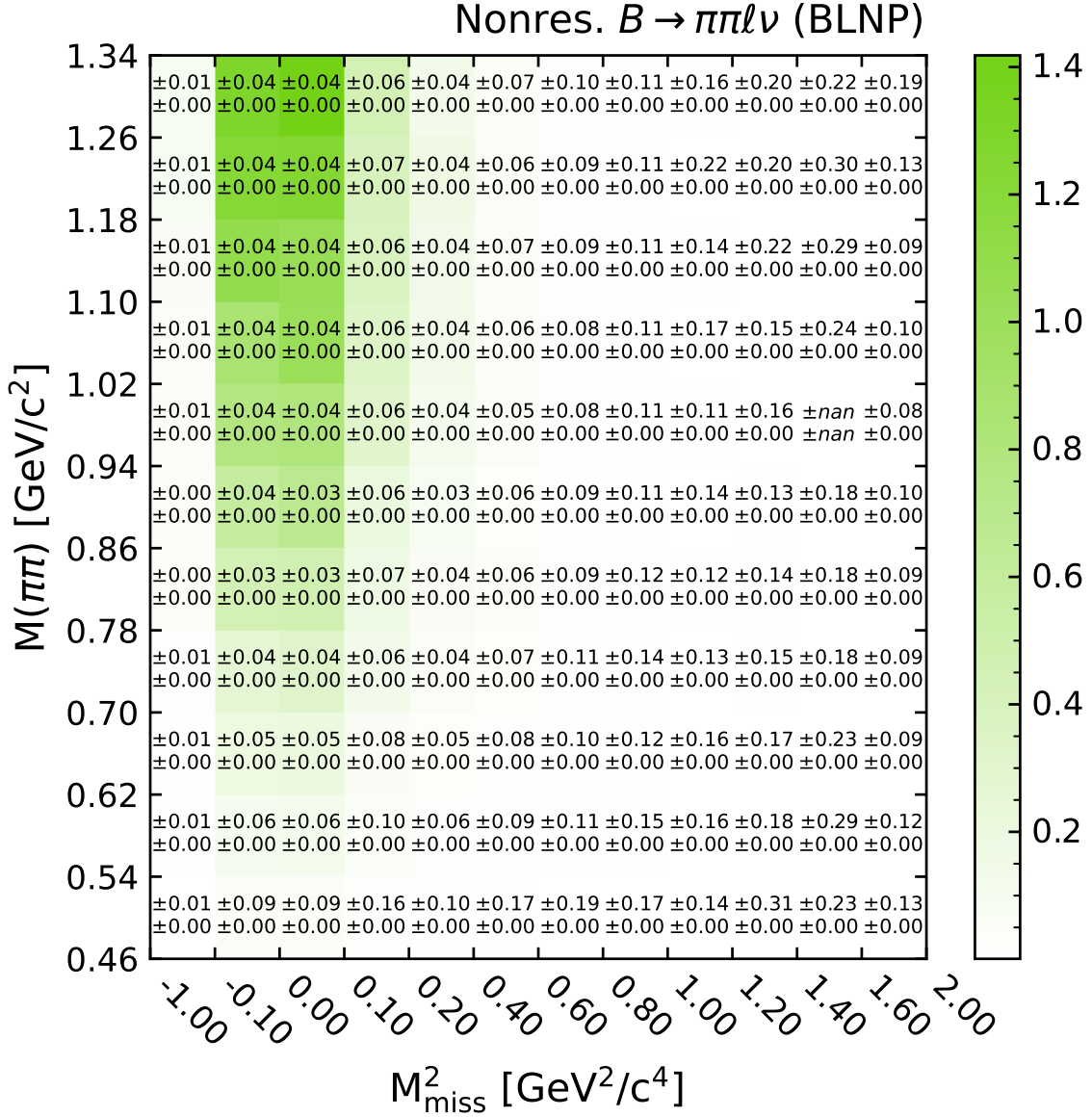


Figure A.6: Two-dimensional projection of the nonresonant $B \rightarrow \pi\pi\ell\nu_\ell$ template in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ signal extraction with relative systematic uncertainties in each bin. Further description is given in the text.

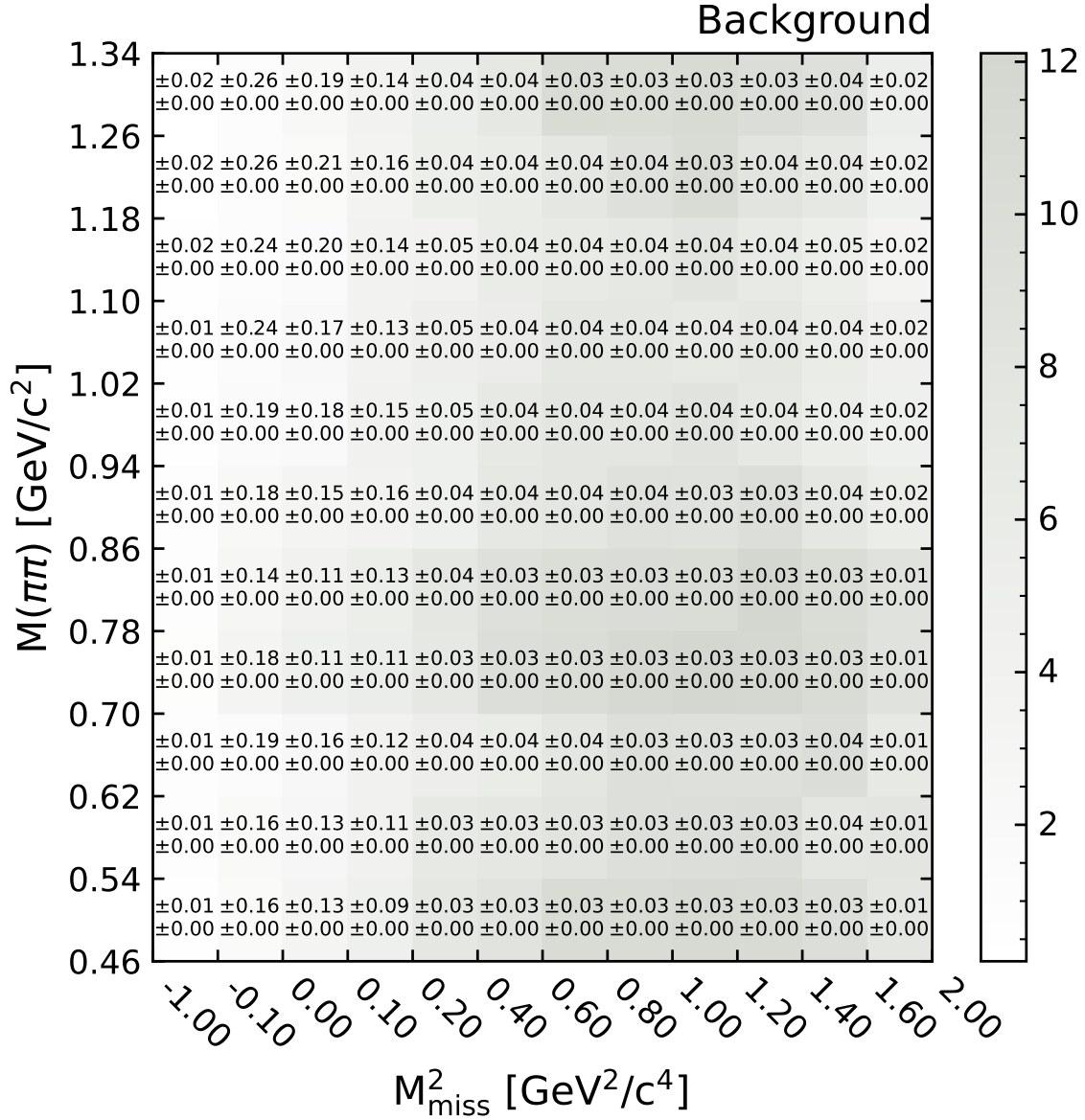


Figure A.7: Two-dimensional projection of the combined background template in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ signal extraction with relative systematic uncertainties in each bin. Further description is given in the text.

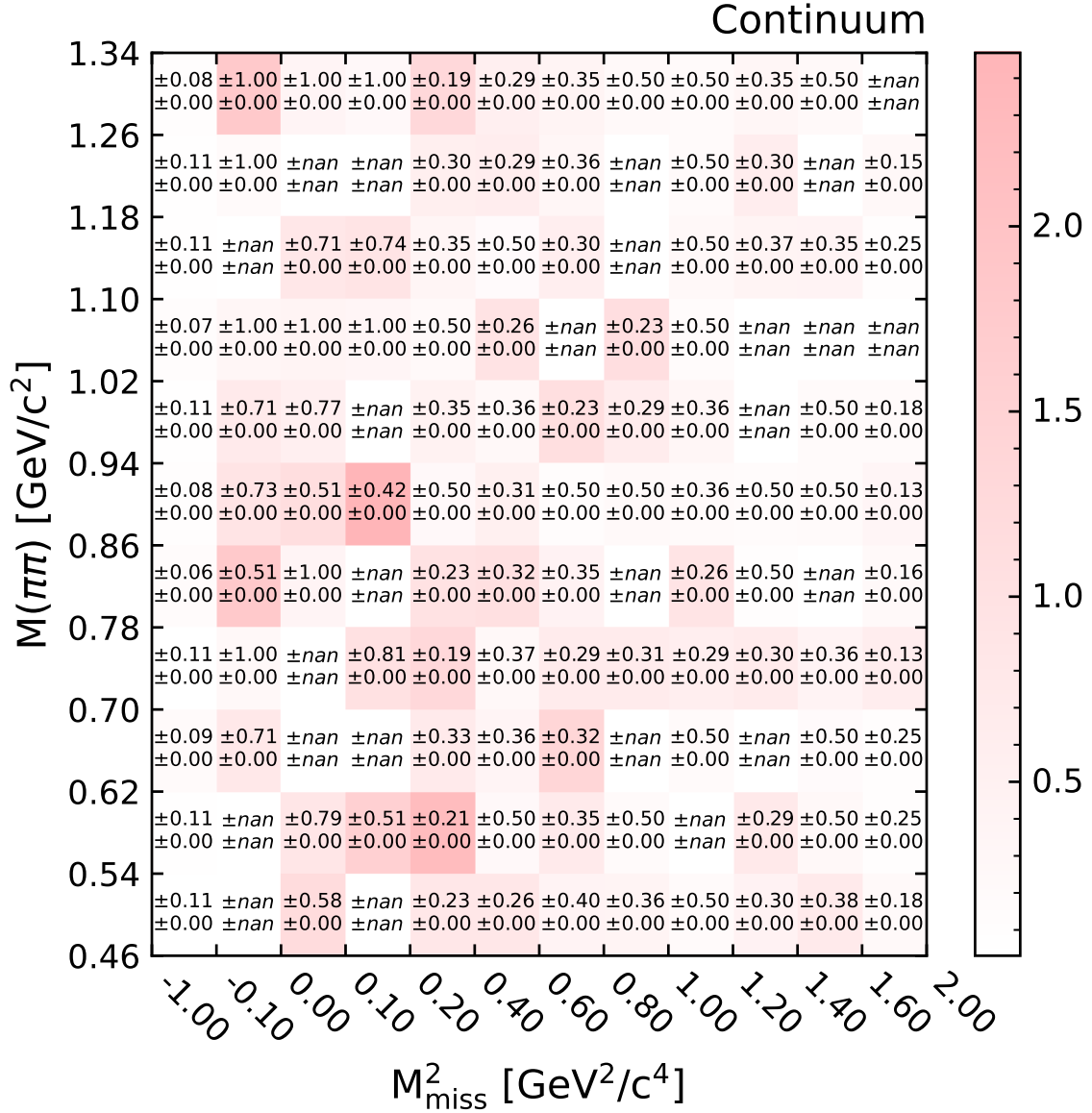


Figure A.8: Two-dimensional projection of the continuum template in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ signal extraction with relative systematic uncertainties in each bin. Further description is given in the text.

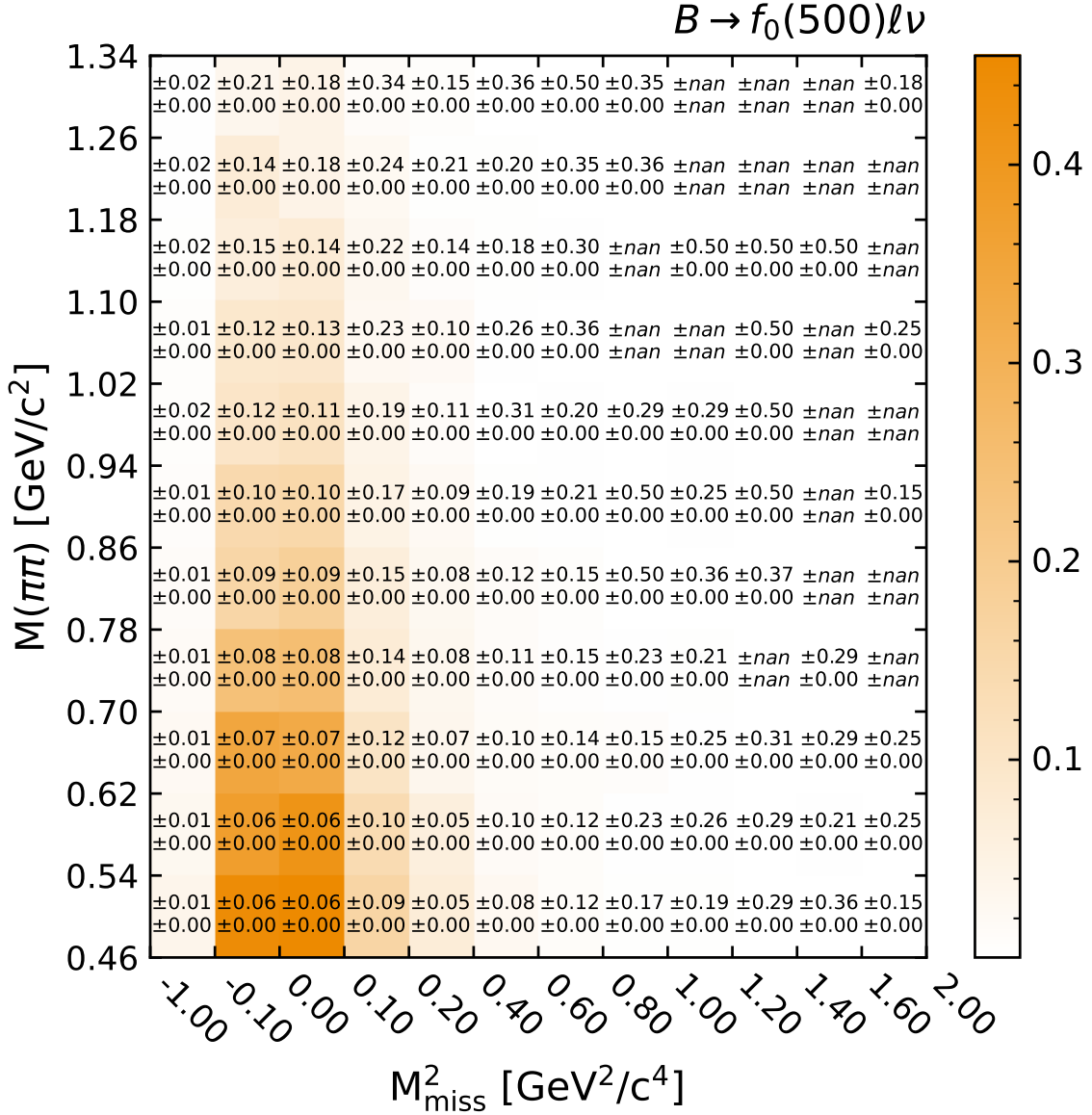


Figure A.9: Two-dimensional projection of the $B \rightarrow f_0(500)\ell\nu$ template in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ signal extraction with relative systematic uncertainties in each bin. Further description is given in the text.

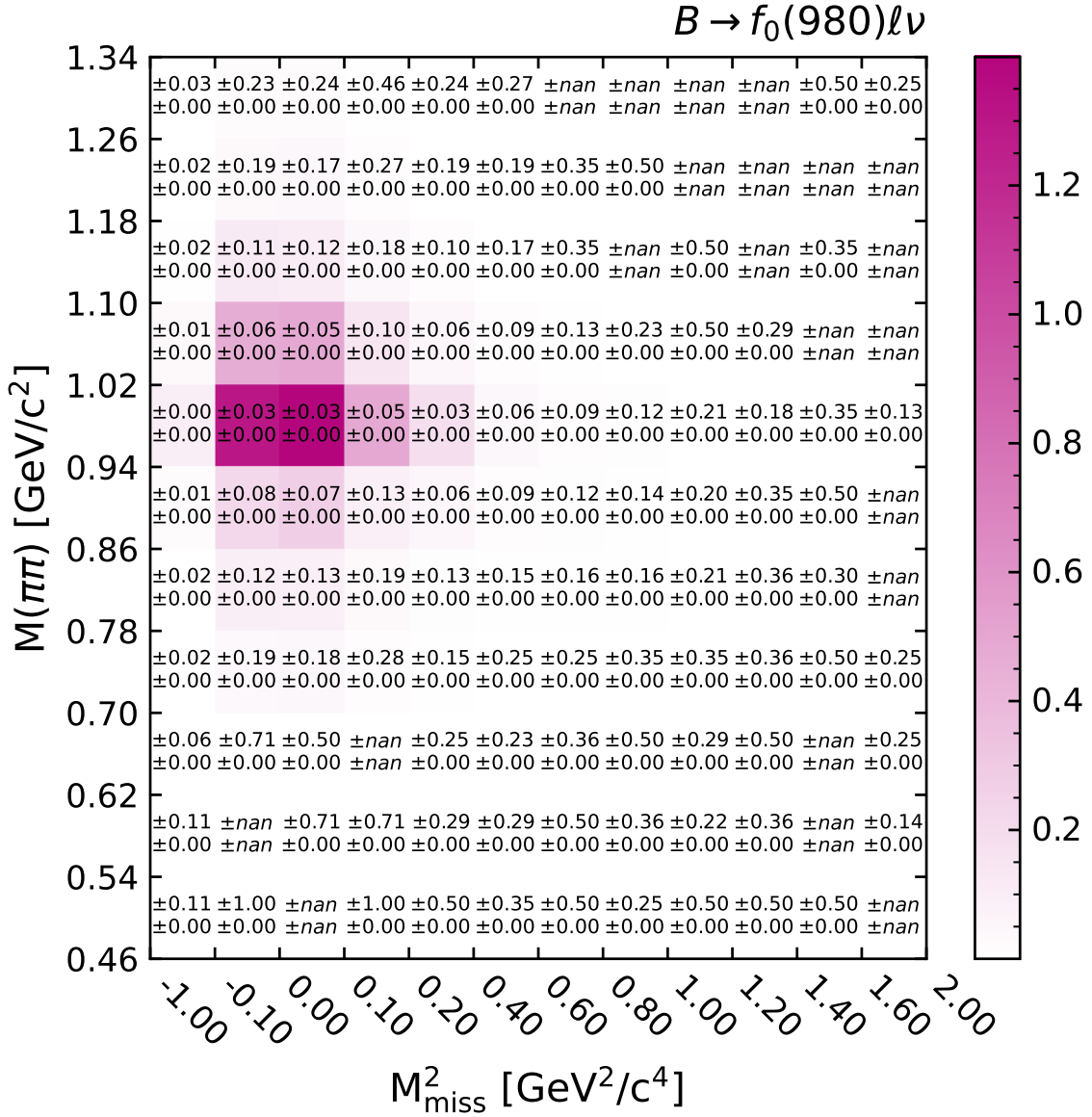


Figure A.10: Two-dimensional projection of the $B \rightarrow f_0(980)\ell\nu_\ell$ template in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ signal extraction with relative systematic uncertainties in each bin. Further description is given in the text.

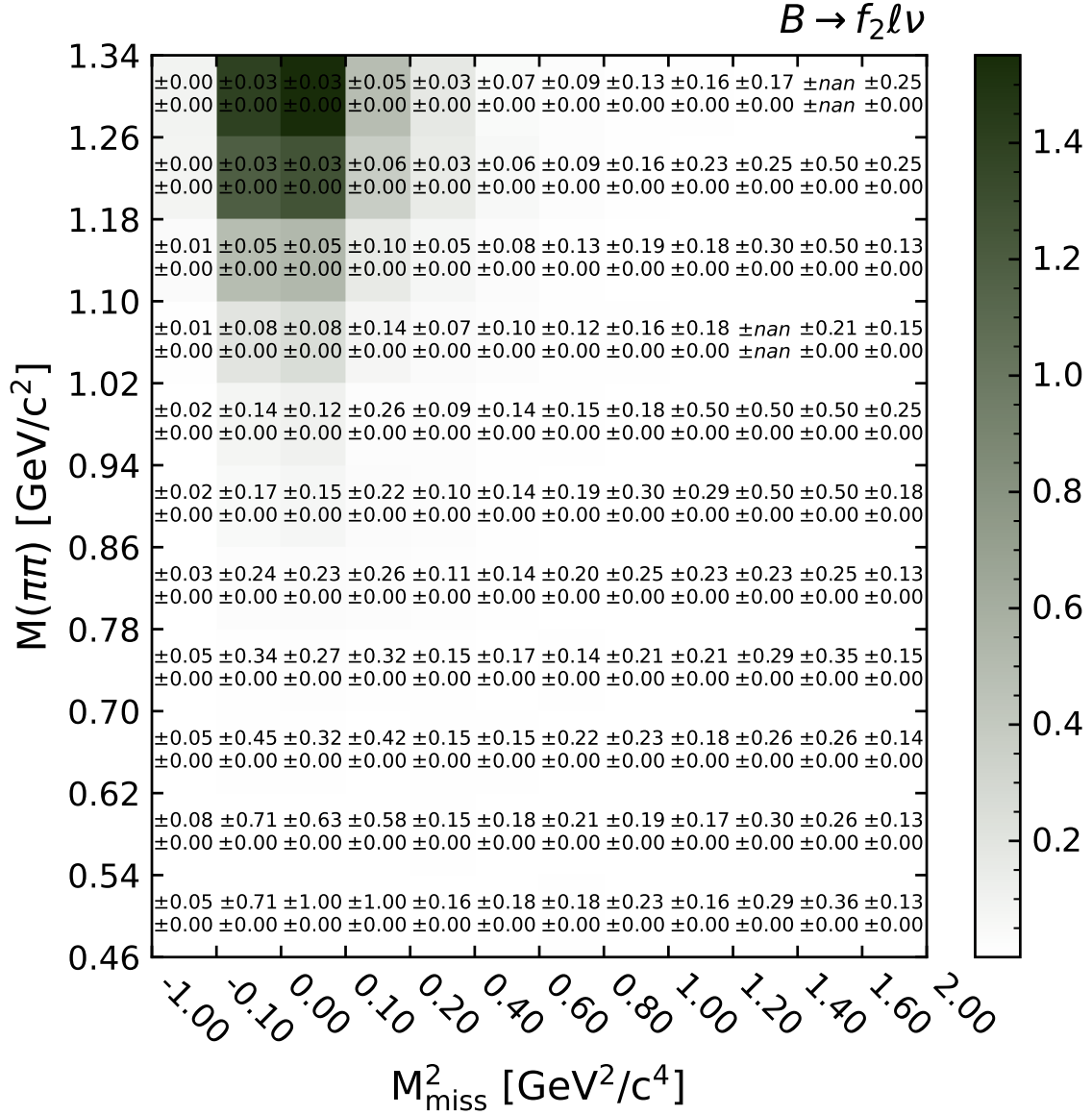


Figure A.11: Two-dimensional projection of the $B \rightarrow f_2(1270)\ell\nu_\ell$ template in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ signal extraction with relative systematic uncertainties in each bin. Further description is given in the text.

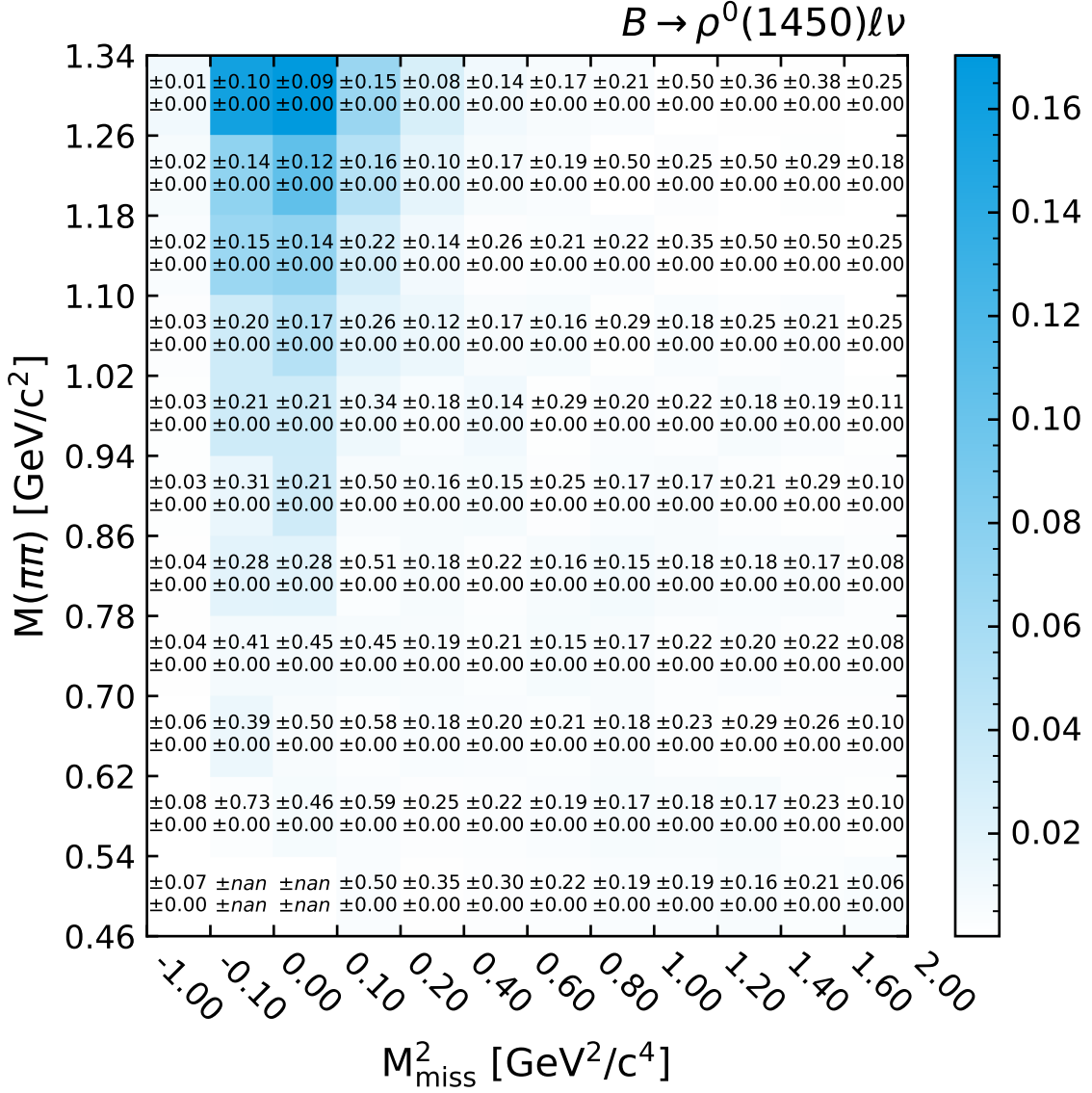


Figure A.12: Two-dimensional projection of the $B \rightarrow \rho^0(1450)\ell\nu_\ell$ template in the $B^+ \rightarrow \rho^0\ell^+\nu_\ell$ signal extraction with relative systematic uncertainties in each bin. Further description is given in the text.

Appendix B

Bin-by-bin Fit Result Figures and Comparison to Simulation

B.0.1 $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ Results

Supplemental to the projections of the fit result to M_{miss}^2 and $M(\pi\pi)$ found in Section 9.1, Figs. B.1 and B.2 give the fit to $M(\pi\pi)$ in bins of M_{miss}^2 and Figs. B.3 and B.4 give the fit to M_{miss}^2 in bins of $M(\pi\pi)$.

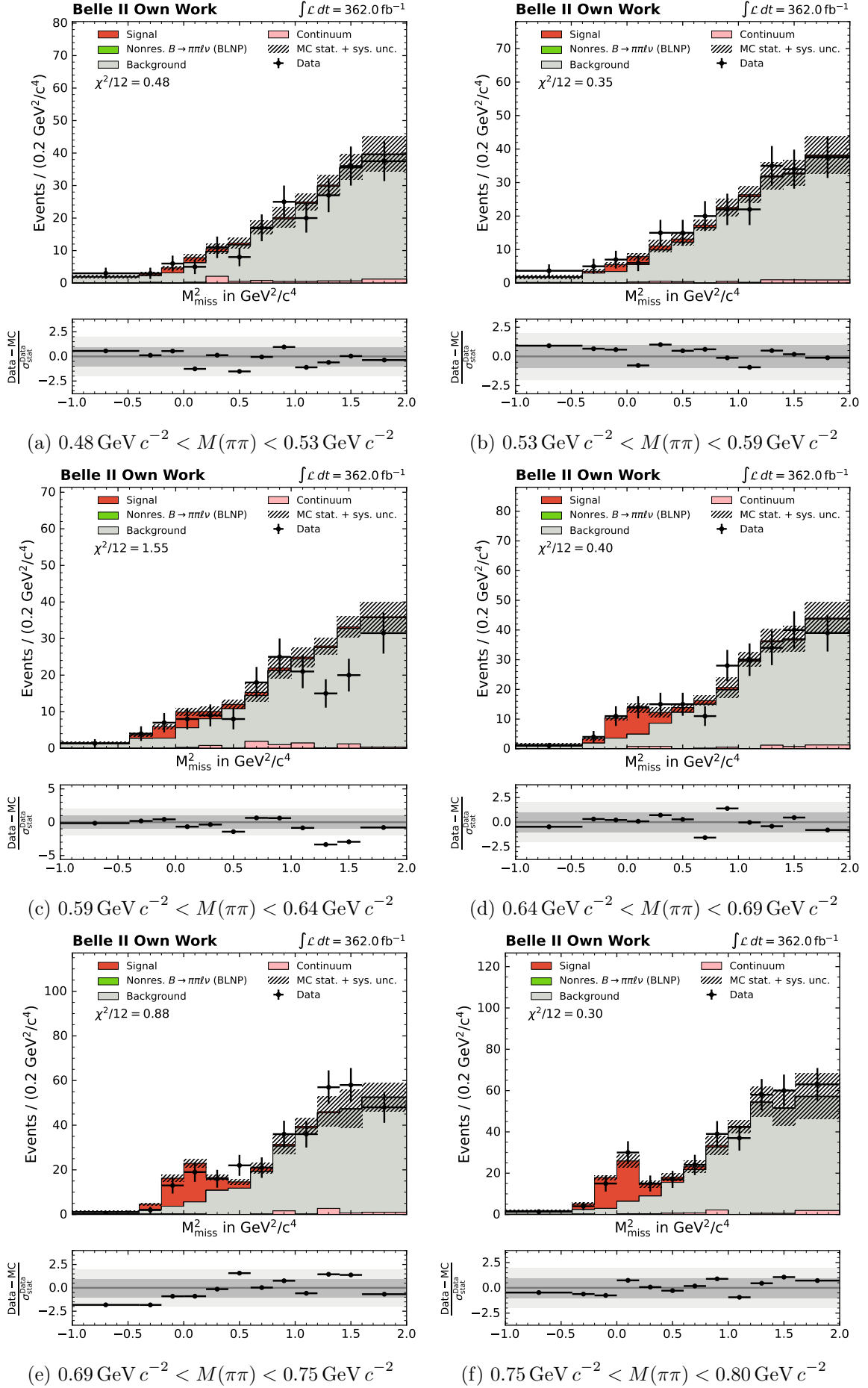


Figure B.1: Fit result to data in M_{miss}^2 shown in bins of $M(\pi\pi)$ in the range 0.48 to $0.80 \text{ GeV } c^{-2}$.

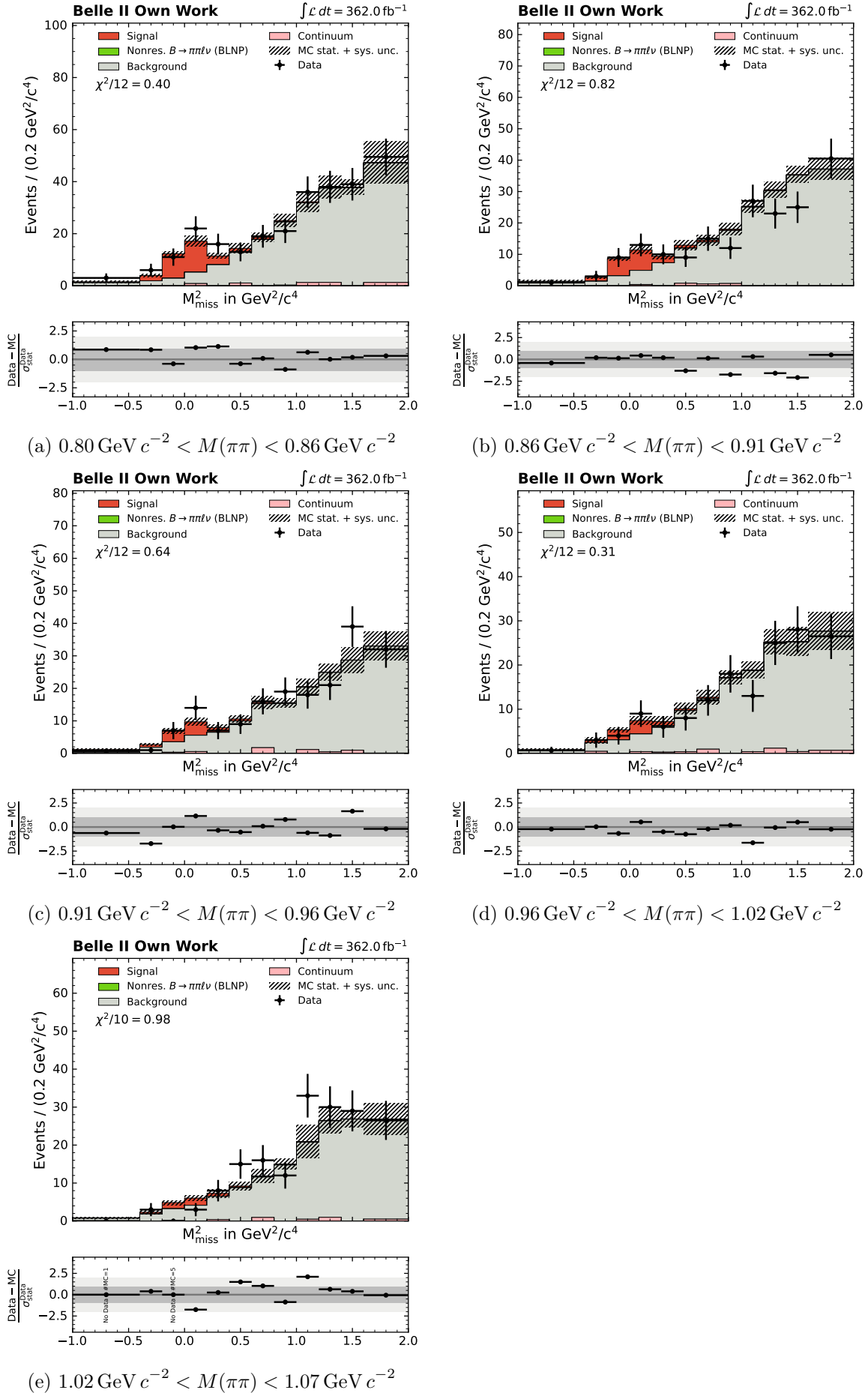


Figure B.2: Fit result to data in M_{miss}^2 shown in bins of $M(\pi\pi)$ in the range 0.80 to $1.07 \text{ GeV } c^{-2}$.

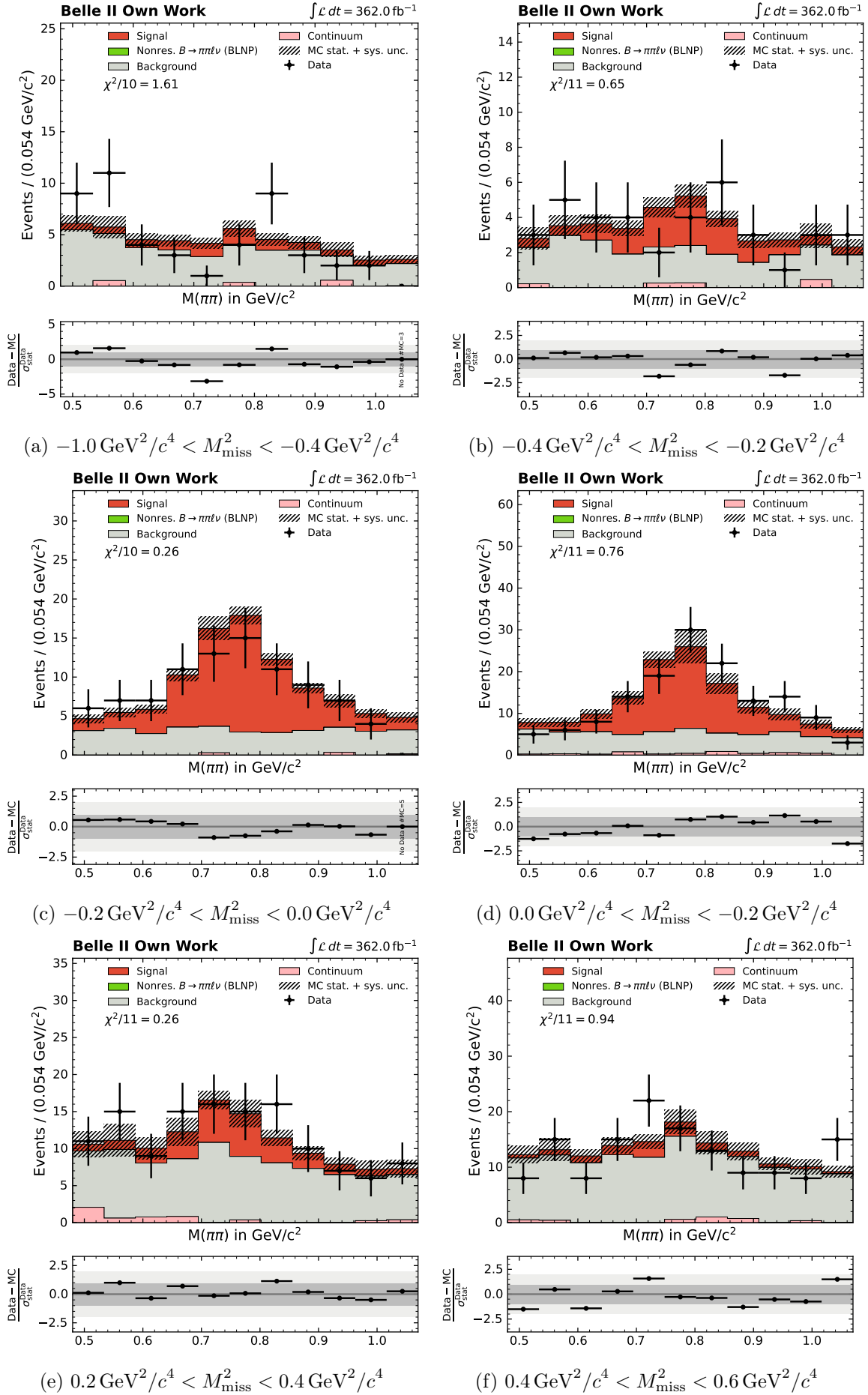


Figure B.3: Fit result to data in $M(\pi\pi)$ shown in bins of M_{miss}^2 in the range -1.0 to $0.6 \text{ GeV}^2/c^4$.

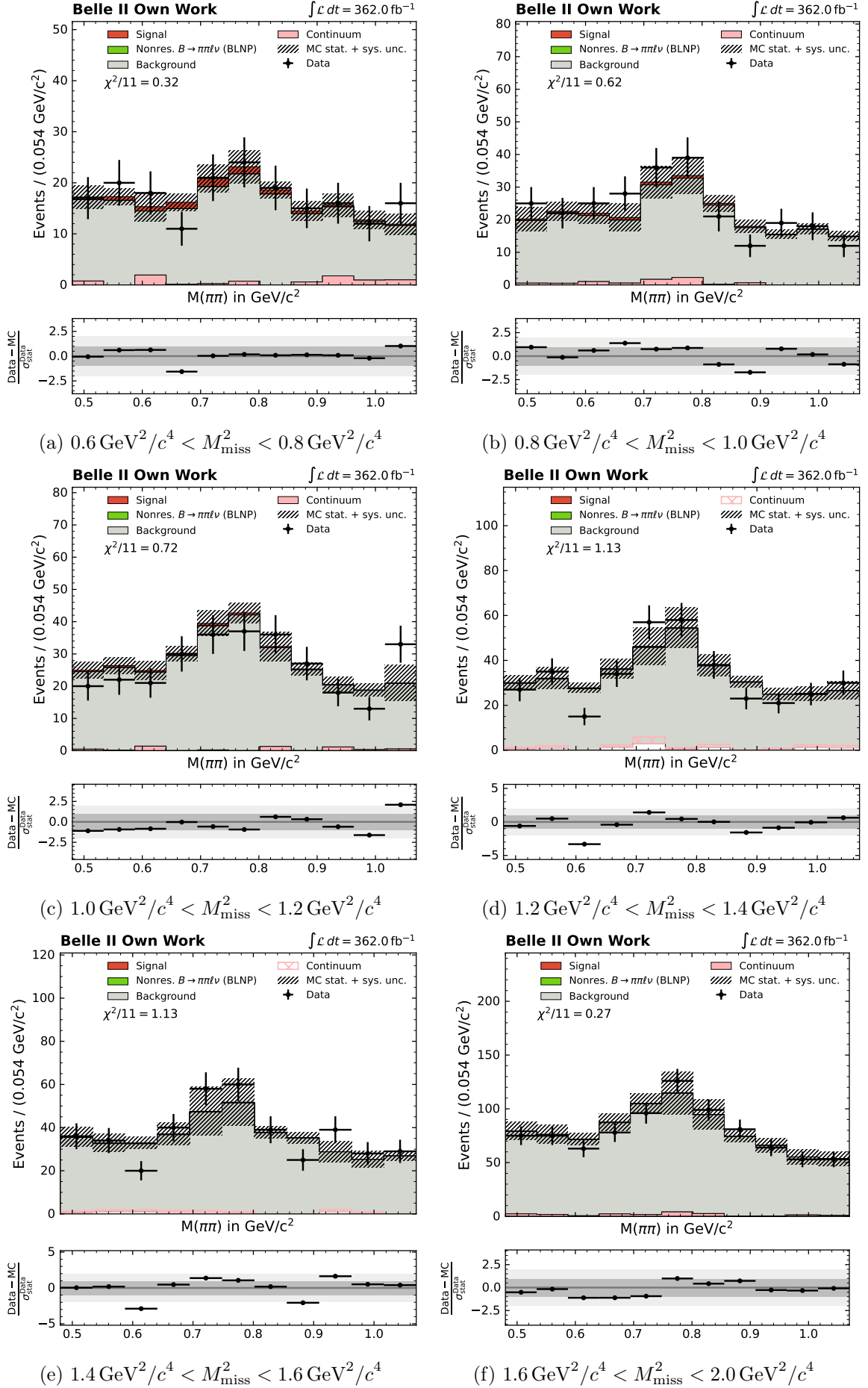


Figure B.4: Fit result to data in $M(\pi\pi)$ shown in bins of M_{miss}^2 in the range -1.0 to $0.6 \text{ GeV}^2/c^4$.

B.0.2 $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ Results

Supplemental to the projections of the fit result to M_{miss}^2 and $M(\pi\pi)$ found in Section 9.1, Figs. B.5 and B.6 give the fit to $M(\pi\pi)$ in bins of M_{miss}^2 and Figs. B.7 and B.8 give the fit to M_{miss}^2 in bins of $M(\pi\pi)$.

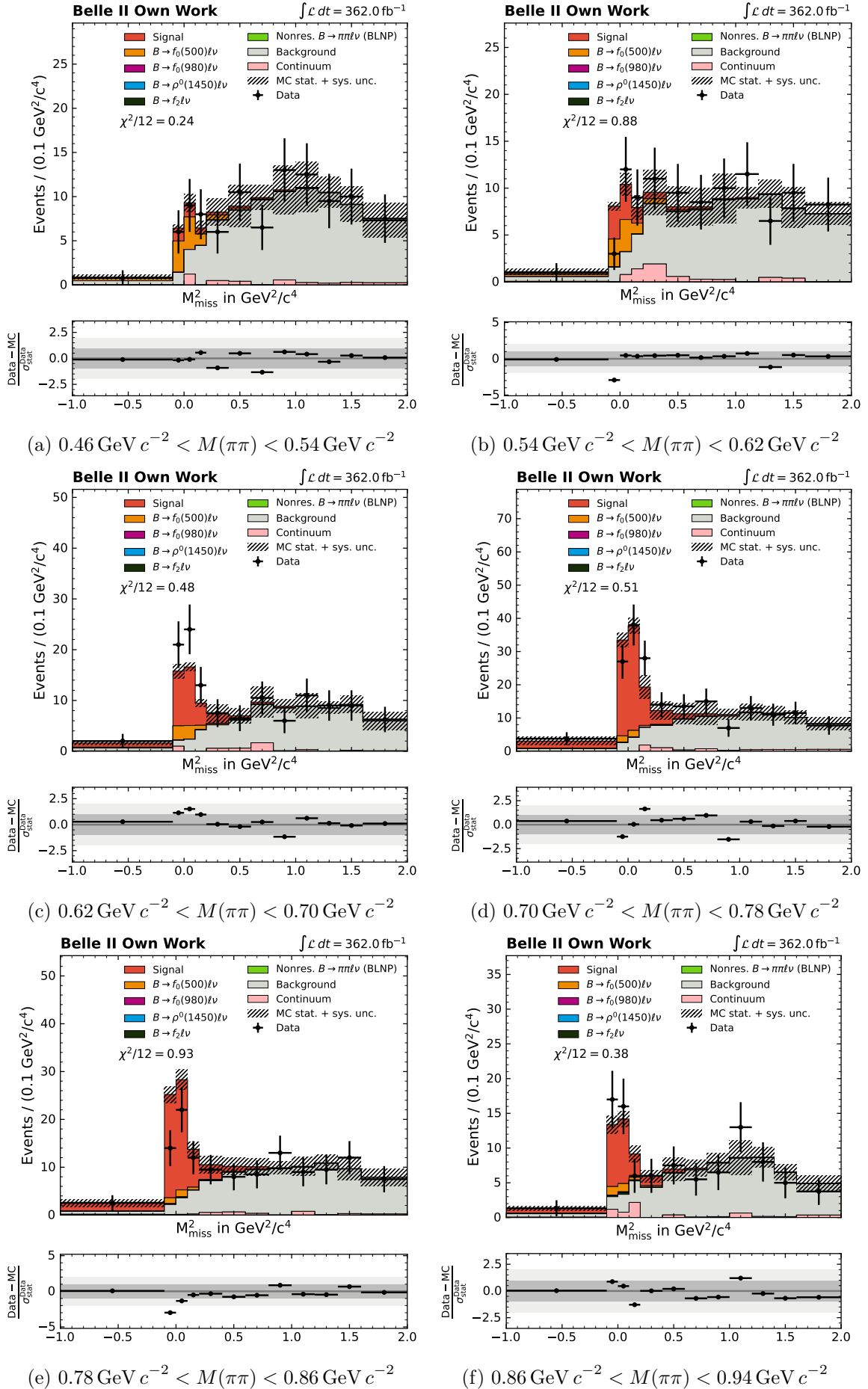


Figure B.5: Fit result to data in M_{miss}^2 shown in bins of $M(\pi\pi)$ in the range 0.48 to $0.80 \text{ GeV } c^{-2}$.

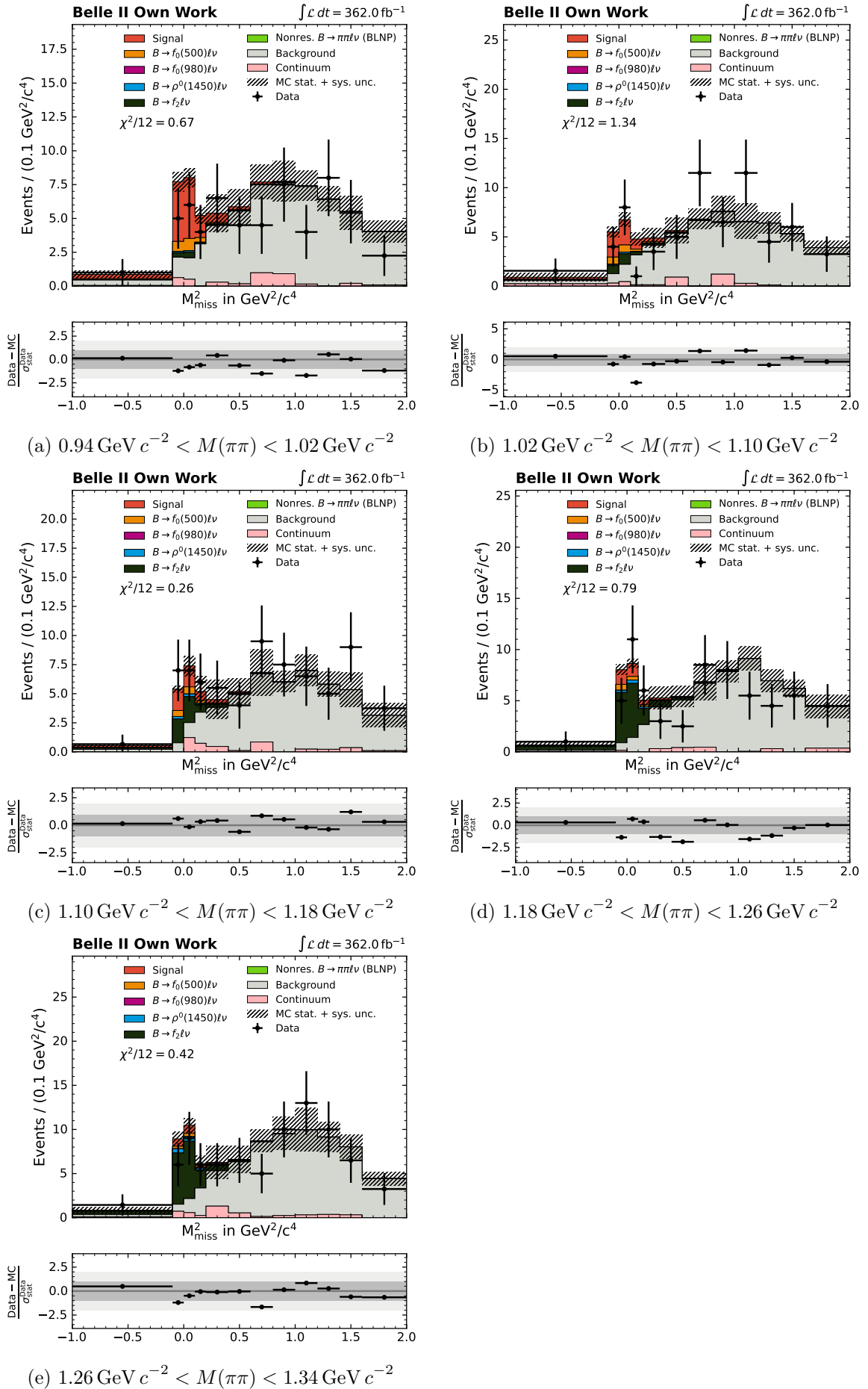


Figure B.6: Fit result to data in M_{miss}^2 shown in bins of $M(\pi\pi)$ in the range 0.48 to $0.80 \text{ GeV } c^{-2}$.

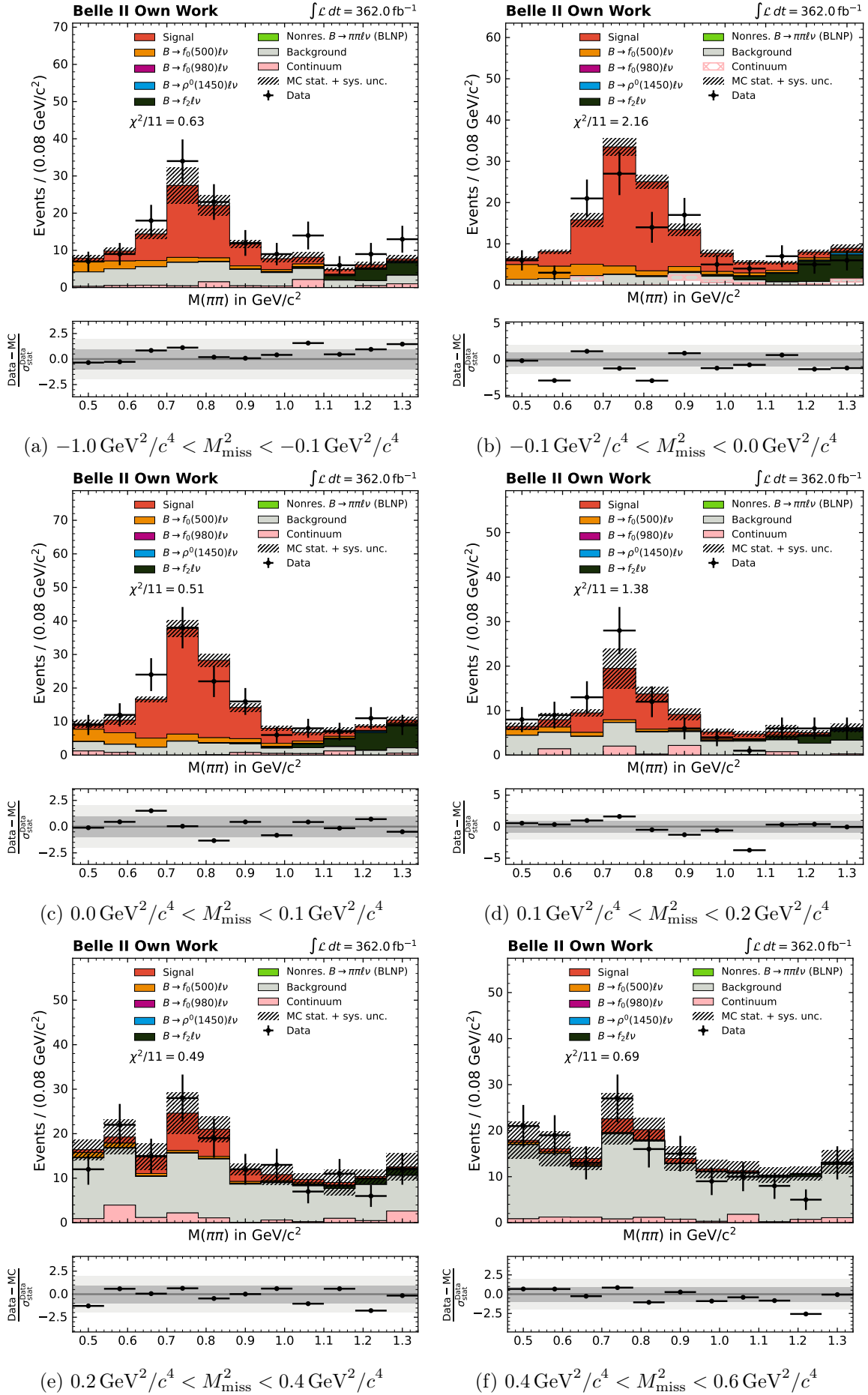


Figure B.7: Fit result to data in $M(\pi\pi)$ shown in bins of M_{miss}^2 in the range -1.0 to $0.2 \text{ GeV}^2/c^4$.

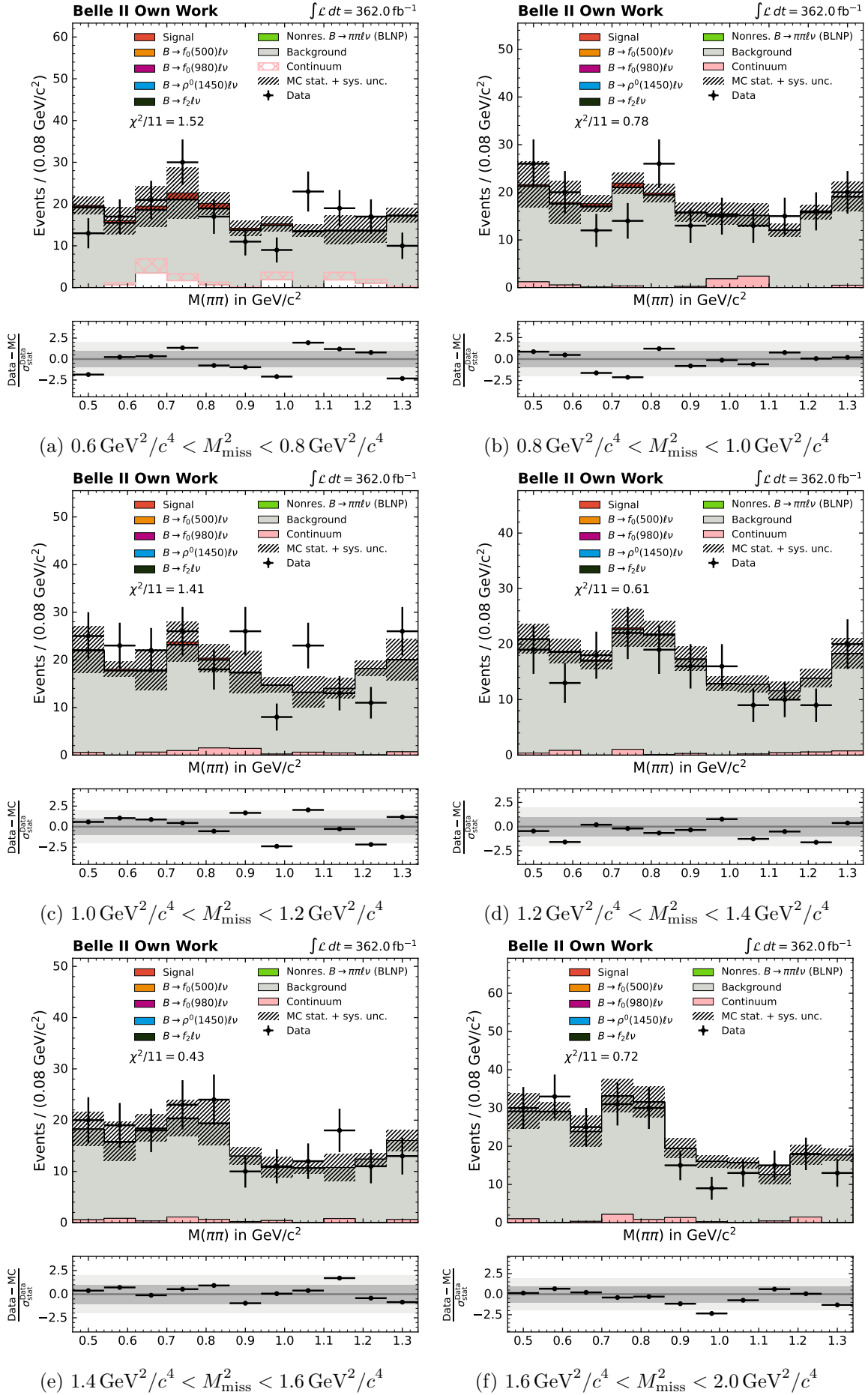


Figure B.8: Fit result to data in $M(\pi\pi)$ shown in bins of M_{miss}^2 in the range 0.2 to $2.0 \text{ GeV}^2/c^4$.

Appendix C

Post-fit Nuisance Parameters for All Templates in the Fit

C.0.1 $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ Nuisance Parameters

Figure C.1 shows the values and errors of the bin-wise nuisance parameters after the fit to data in the $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ channel for all templates supplemental to the illustrations given in Section 9.1. Small deviations of the nuisance parameters around zero can be observed for the continuum background template and the $B\bar{B}$ background template due to statistical fluctuations caused by the small simulated sample and the sizeable systematic uncertainties on the $B\bar{B}$ template shape. Nuisance parameters for all other templates cannot be constrained by the fit and observe no pulls.

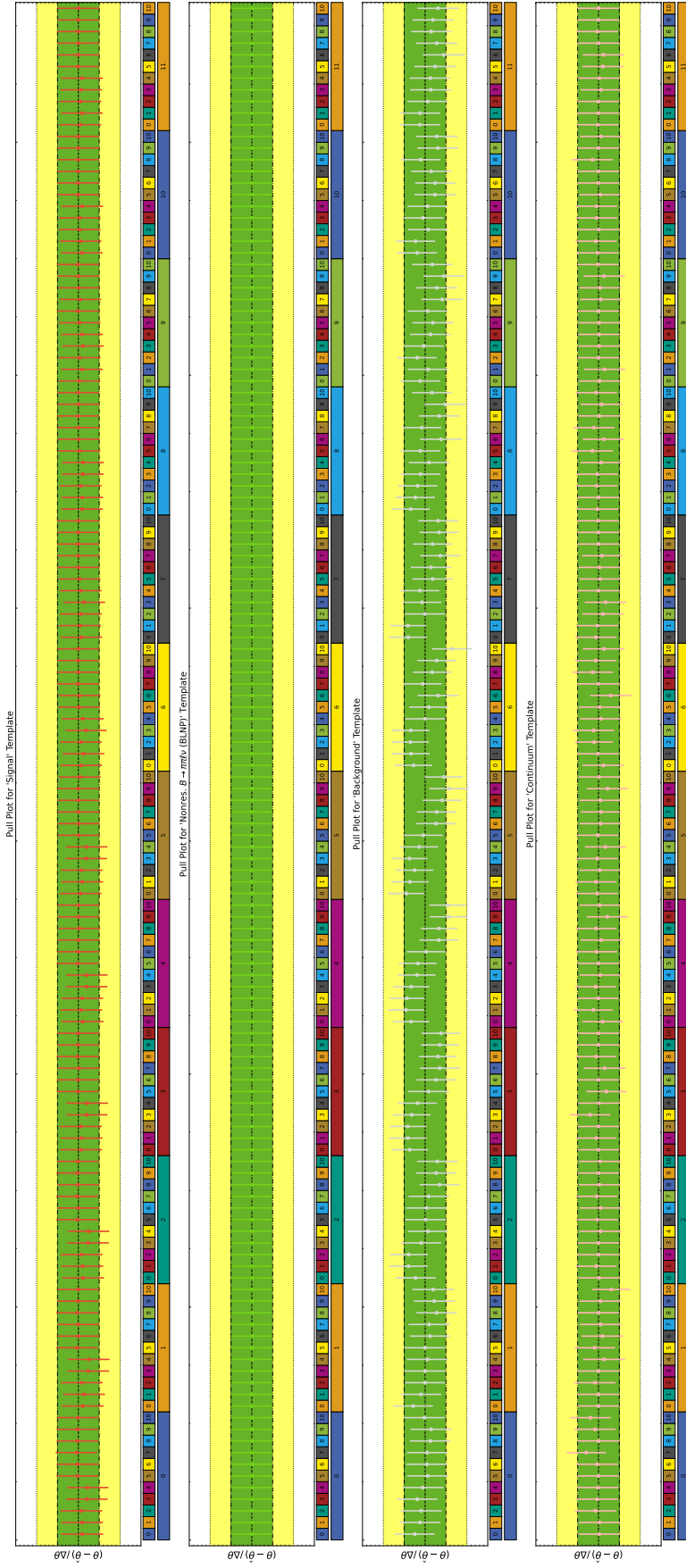


Figure C.1: Nuisance parameter pulls of the templates for (from top to bottom) $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ signal, nonresonant X_u dipion backgrounds, BB backgrounds, and continuum backgrounds in the fit to data in the $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ channel. The top row of indices below each template indicates the bins in $M(\pi\pi)$, the bottom row bins in M_{miss}^2 .

C.0.2 $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ Nuisance Parameters

Figures C.2 and C.3 show the values and errors of the bin-wise nuisance parameters after the fit to data in the $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ channel for all templates supplemental to the illustrations given in Section 9.1. Small deviations of the nuisance parameters around zero can be observed for the continuum background template and the $B\bar{B}$ background template due to statistical fluctuations caused by the small simulated sample and the sizeable systematic uncertainties on the $B\bar{B}$ template shape. As described in Section 9.1, the signal template also shows pulls of the central value away from zero which is attributed to $\rho - \omega$ - interference which is not included in the nominal template model due to its unknown size. Nuisance parameters for all other templates cannot be constrained by the fit and observe no pulls.

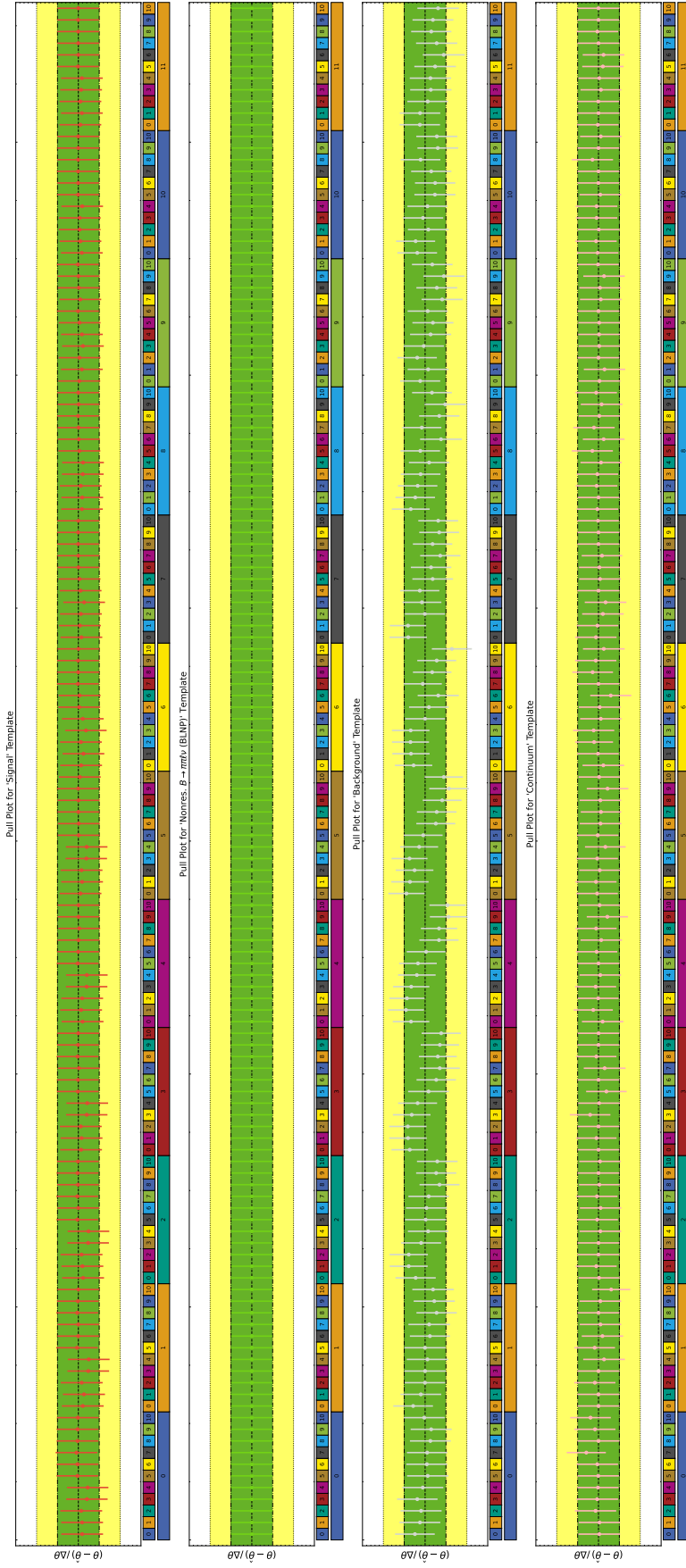


Figure C.2: Nuisance parameter pulls of the templates for (from top to bottom) $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ signal, nonresonant X_u dipion backgrounds, BB backgrounds, and continuum backgrounds in the fit to data in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ channel. The top row of indices below each template indicates the bins in $M(\pi\pi)$, the bottom row bins in M_{miss}^2 .

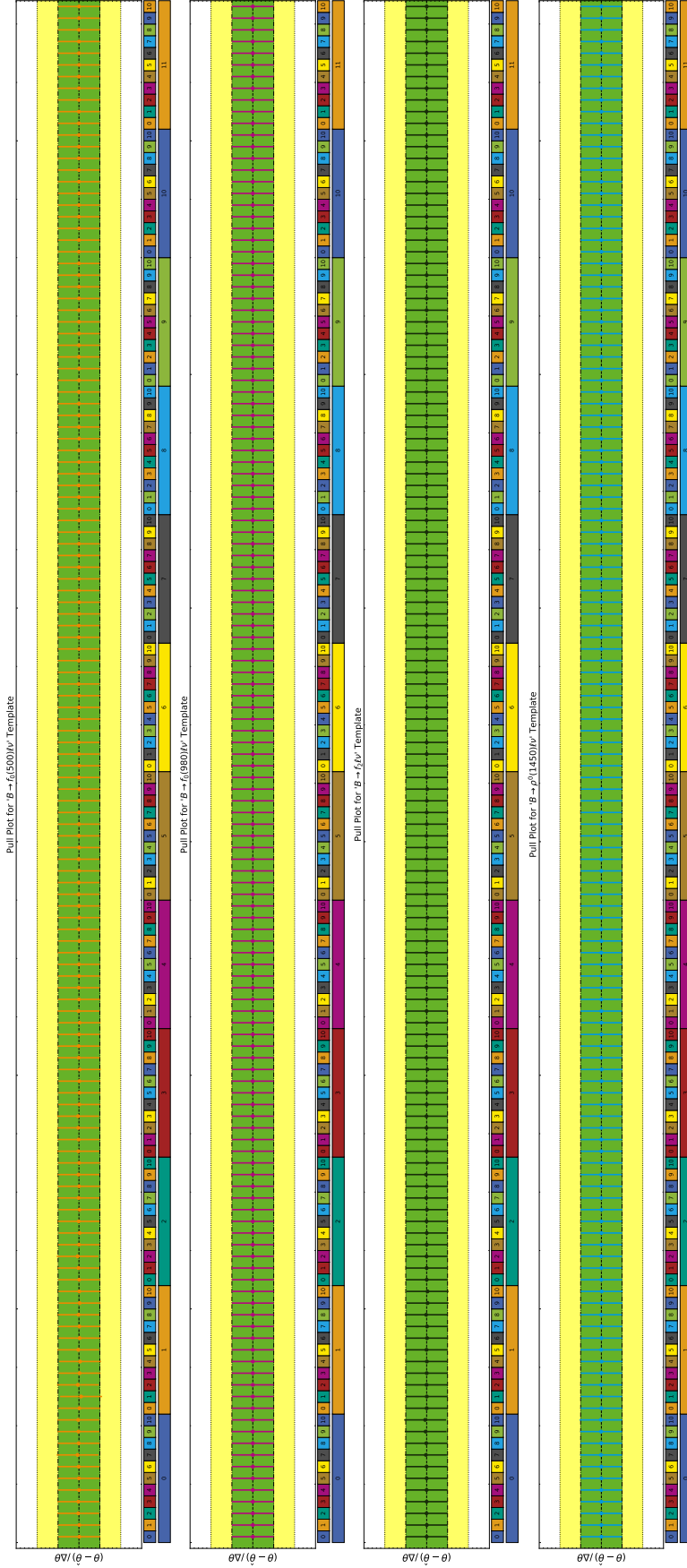


Figure C.3: Nuisance parameter pulls of the templates for (from top to bottom) $B \rightarrow f_0(500)\ell\nu_\ell$ backgrounds, $B \rightarrow f_0(980)\ell\nu_\ell$ backgrounds, $B \rightarrow f_2(1270)\ell\nu_\ell$ backgrounds, and $B \rightarrow \rho^0(1450)\ell\nu_\ell$ backgrounds in the fit to data in the $B^+ \rightarrow \rho^0\ell^+\nu_\ell$ channel. The top row of indices below each template indicates the bins in $M(\pi\pi)$, the bottom row bins in M_{miss}^2 .

Appendix D

Correlation Matrices between the FEI Calibration Factors

The FEI calibration factors introduced Section 6.2.2 are given with uncertainties and the correlation between the factors. This is used as systematic uncertainty as described in Section 8.2.2. The correlation is illustrated in Fig. D.1 for charged B_{tag} candidates and in Fig. D.2 for neutral ones.

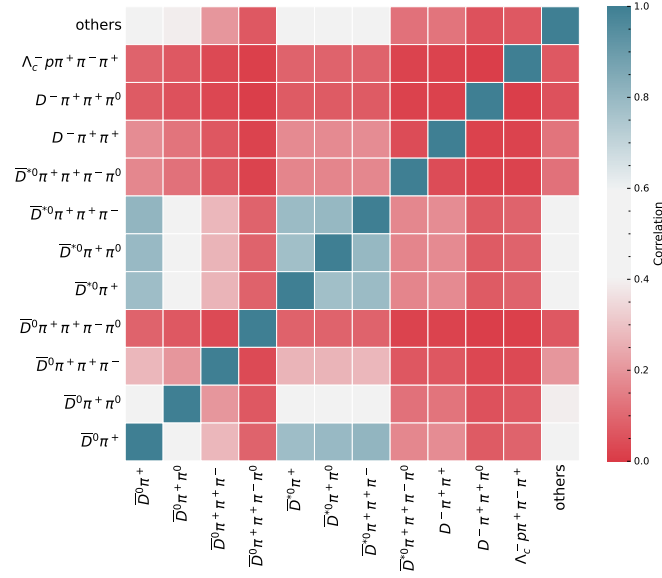


Figure D.1: Correlation matrices for the FEI factors for B^+ tag candidates. Large correlation is observed for FEI channels dominated by the shared systematic uncertainty of the $B \rightarrow X_c \ell \nu_\ell$ branching fraction.

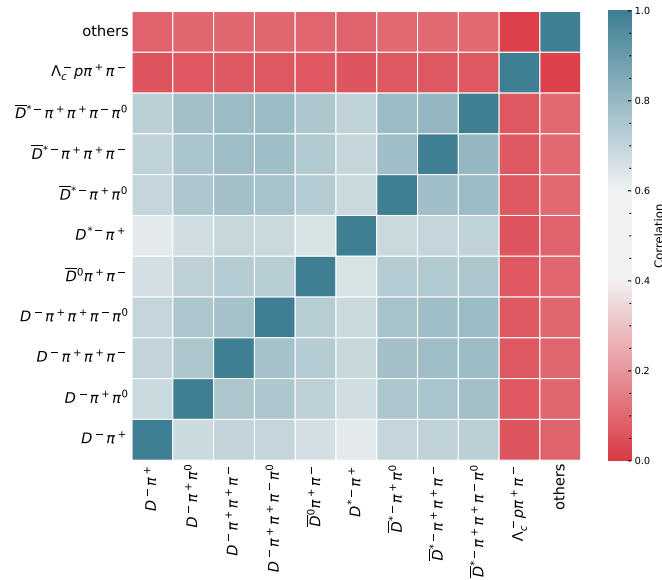


Figure D.2: Correlation matrices for the FEI factors for B^0 tag candidates. Large correlation is observed for FEI channels dominated by the shared systematic uncertainty of the $B \rightarrow X_c \ell \nu_\ell$ branching fraction.

Appendix E

Hybrid MC for Neutral $B \rightarrow X_u \ell \nu_\ell$ Decays

Illustrations of the lepton energy in the B frame E_l^B , the momentum transfer q^2 and the mass of the hadronic system M_X after application of the hybrid weights determined for neutral $B \rightarrow X_u \ell \nu_\ell$ decays are shown in Fig. E.1. For charged $B \rightarrow X_u \ell \nu_\ell$ decays, this is illustrated in Fig. 5.3.

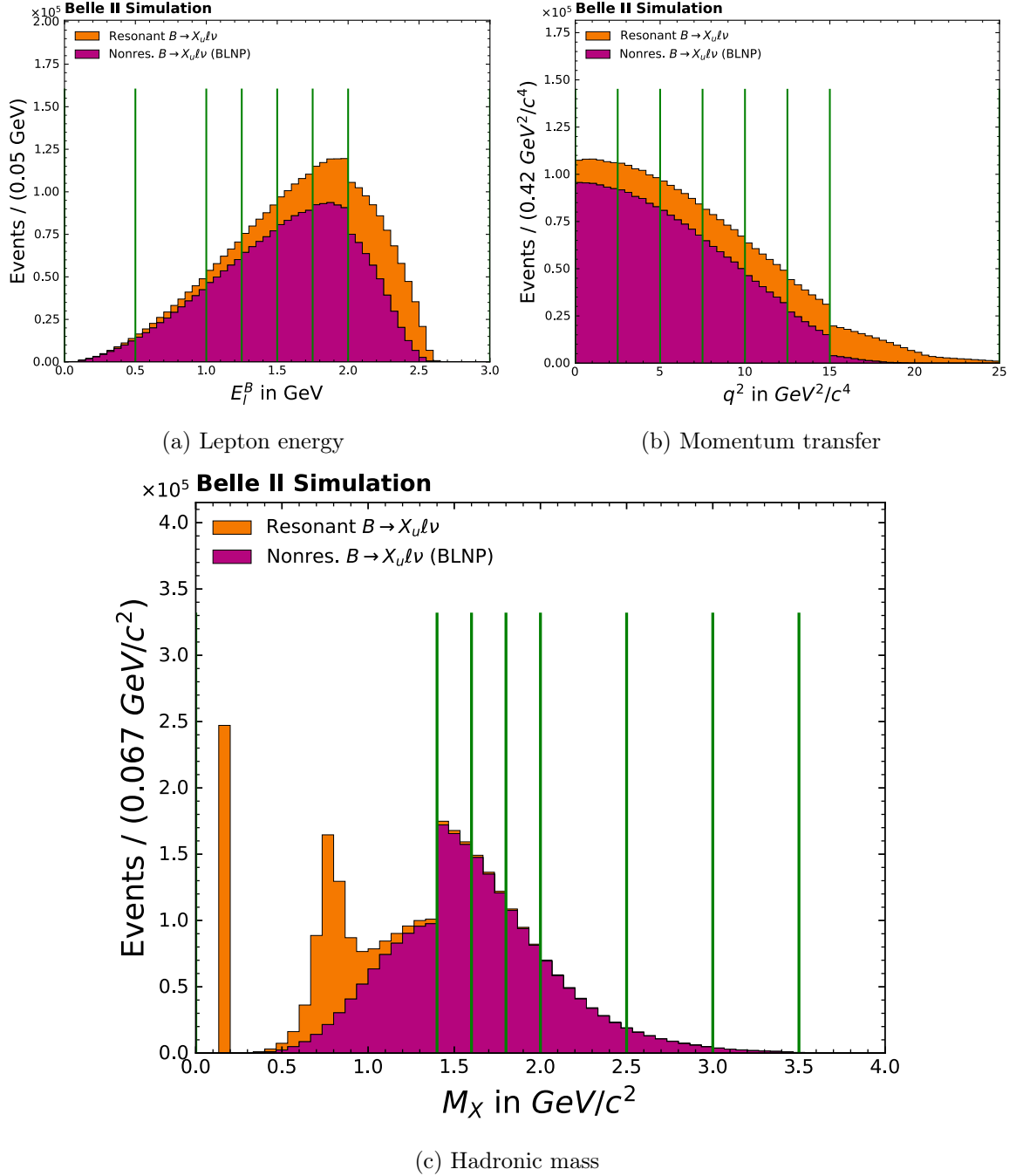


Figure E.1: Hybrid MC model combination for neutral $B \rightarrow X_u \ell \nu_\ell$ decays projected to the reweighting variables with bins of the reweighting procedure indicated by green vertical lines. Shown are the lepton energy in the B frame E_l^B , the momentum transfer q^2 and the mass of the hadronic system M_X .

Glossary

BaBar A first-generation B factory, located at the Stanford Linear Accelerator Center. 1, 7, 10, 16, 24, 94

Belle A first-generation B factory and the predecessor of the Belle II experiment. 1, 7, 10–13, 15, 16, 25, 46, 94, 145

Belle II A second-generation B factory and the successor of the Belle experiment. 9–13, 15–17, 44, 145, 150, 158

CLEO A general purpose detector located at the CESR. More at Ref. [130]. 7, 9, 44, 45, 48, 150

continuum A class of background processes at the B Factories in which no $\Upsilon(4S)$ resonance is created and instead a light-quark pair interaction occurs.. 37, 49, 152

DIRAC A software package used to supply heterogeneous computing resources to particle physics collaborations. See Ref. [44]. 17, 19, 20

EvtGen A Monte Carlo Generator suited for the decay of heavy flavor particles. More at Ref. [70]. 29, 30, 35, 37, 67, 80, 84, 85, 151, 153, 159

FastBDT Implementation of SGBDTs using the Eigen matrix library. See Ref. [59]. 23, 25, 26, 44, 47, 151

GEANT4 A toolkit for simulating the passage of particles through matter using a wide variety of phenomenological models. See Ref. [69]. 28

ISGW2 A generic model for semileptonic meson decays. See Ref. [78]. 30

KEKB The electron-positron collider at which the Belle experiment was located.. 11, 12, 158

KKMC A Monte Carlo generator specifically for lepton and quark pair production at lepton colliders. See Ref. [100].. 37

-
- POCA** The point of closest approach (POCA) is a tracks' signed distance to the z axis as determined by the tracking algorithm. 40
- PYTHIA** A general-purpose Monte Carlo generator used to describe hard and soft interactions, parton distributions, initial- and final-state parton showers, multiparton interactions, fragmentation and decay. See Ref. [80]. 31, 32, 37, 80, 151, 158
- SuperKEKB** An upgrade of the KEKB electron-positron collider and the accelerator at which the Belle II experiment is located.. 9, 11, 12, 14, 17, 150, 158
- TreeFitter** A decay chain fitting algorithm developed for Belle II. More at Ref. [63]. 24, 42
- XCache** An XRootD-based disk caching mechanism. More at Ref. [49]. 18–21, 97, 151
- XRootD** A file transfer protocol and software framework to transfer payloads typical in high-energy physics. More at Ref. [47]. 18–20, 97, 146, 151

Acronyms

BGL Boyd–Grinstein–Lebed. 35

ISGW2 Updated Isgur–Scora–Grinstein–Wise Model. 30, 80, 82, *Glossary*: ISGW2

ALICE A Large Ion Collider Experiment. 18

ARICH Aerogel Ring-Imaging Cherenkov detector. 14, 15, 23

basf2 Belle II Analysis Software Framework. 23–25, 28, 29, 35

BDT Boosted Decision Tree. 22, 24, 38, 40, 50

BLNP Bosch–Lange–Neubert–Paz. 31–33, 63, 67, 80–82, 151, 153, 154

BSZ Bharucha–Straub–Zwicky. 28, 30, 151

CDC Central Drift Chamber. 14, 40, 41, 44, 45

CESR Cornell Electron Storage Ring. 9, 145, 150

CKM Cabibbo–Kobayashi–Maskawa. 1, 3–5, 10, 35, 38, 98, 150

cms center-of-mass system. 12, 17, 42, 45, 46, 150, 152

CUSB Columbia University-Stony Brook. 9, 150

DEPFET Depleted Field Effect Transistor. 14

ECL Electromagnetic Calorimeter. 15, 23, 24, 41

FEI Full Event Interpretation. 25, 27, 36, 38, 51–54, 57, 64, 83, 86, 141, 142, 152, 157

FSP Final State Particle. 40

FSR Final State Radiation. 37

HER High-Energy Ring. 11, 12

HFLAV Heavy Flavor Averaging Group. 8, 94–97, 150, 155, 160

HQE Heavy Quark Expansion. 5, 6, 31, 80

HQET Heavy Quark Effective Theory. 5

IJS Institut Jožef Stefan. 20

IP Interaction Point. 11, 12, 15, 25, 40

ISR Initial State Radiation. 37

KEK High Energy Accelerator Research Organization. 11, 150

KLM K_L^0 and muon detector. 15, 23

KSFW Kakuno Super Fox–Wolfram. 46–48

LCSR Light-Cone Sum Rules. 7, 28, 30

LER Low-Energy Ring. 11, 12

LHC Large Hadron Collider. 12, 14

LHCb Large Hadron Collider beauty. 1, 10

LQCD Lattice Quantum Chromodynamics. 30

MC Monte Carlo. 20, 37, 47, 49, 60, 152

MVA multivariate analysis. 47, 49, 152

OZI Ōkubo–Zweig–Iizuka. 10

PDF probability density function. 23, 58–60, 63

PDG Particle Data Group. 5, 8, 94–96, 150, 155, 160

PID particle identification. 13, 15

POCA Point of closest approach. 40, *Glossary*: POCA

PXD Pixel Detector. 13, 14, 23

ROC receiver operating characteristic. 49, 152

ROE Rest of the Event. 23, 46

RPC resistive plate chamber. 15

SGBDT Stochastic Gradient Boosted Decision Tree. 22, 23, 25, 26, 47, 145

SM Standard Model of particle physics. 1, 3–5

SVD Silicon Vertex Detector. 14

TOP Time-Of-Propagation counter. 14, 15, 23

List of Figures

2.1	Inclusive and exclusive average of $ V_{ub} $ over time as calculated by the PDG. Also given is the expectation from CKM unitarity determined in a fit performed by the CKMfitter group [6] without the input from $ V_{ub} $ measurements. . . .	5
2.2	Comparison of statistically independent, previous determinations of $\mathcal{B}(B \rightarrow \rho l \nu_\ell)$ together with the PDG average and the HFLAV evaluation. The PDG average is independent between the charged and neutral decay mode while the HFLAV evaluation uses isospin relations between the modes.	8
3.1	The bottomonium resonances $\Upsilon(1S)$, $\Upsilon(2S)$, $\Upsilon(3S)$ and $\Upsilon(4S)$ in the total hadronic cross section, given in nb as measured at the CESR by the CUSB detector [23]. The inset figure shows additional results for $\Upsilon(4S)$, $\Upsilon(5S)$ and $\Upsilon(6S)$ obtained by the CLEO Collaboration [24]. These are given as the ratio $R = \sigma_{\text{had}}/\sigma_{\mu\mu}$. This plot was initially published in Ref. [25]. This figure is adapted from [26].	9
3.2	Illustration of the SuperKEKB accelerator complex at KEK. The labels “Tsukuba,” “Nikko,” “Fuji”, and “Oho” refer to straight sections of the accelerator, the Belle II experiment is located in the Tsukuba section where the beams collide. Illustration taken from Ref. [33].	11
3.3	The Belle II detector with its seven subdetectors. Adapted from Ref. [37]. . .	13
3.4	The weekly and total integrated luminosity recorded by the Belle II experiment since 2019. The label “ $\Upsilon(4S)$ runs” refers to data taking periods in which SuperKEKB operated with a cms energy of $\sqrt{s} \approx 10.58$ GeV, the label “non- $\Upsilon(4S)$ runs” refers to data taking periods at which the accelerator operated 60 MeV below the $\Upsilon(4S)$ resonance or scanned various energy points. Values taken from Ref. [43].	17
3.5	Number of waiting jobs at different Belle II grid computing sites, captured during one week. Unequal resource demands lead to significantly more than 100,000 jobs waiting at four sites while the remaining sites have much fewer waiting jobs. This uneven amount of waiting jobs can reduce the efficiency of the grid.	19

3.6	Schematic depiction of file access through an XRootD XCache. The label “Remote SE” refers to Storage Elements located at a computing sites not physically close to the server on which the XCache is hosted.	20
4.1	The hierarchical nature of the Full Event Interpretation. Particle candidates are reconstructed in six stages and a FastBDT classifier is applied at each stage to rank the candidates. Initial candidates are reconstructed from tracks and clusters, intermediate particles from combinations of candidates from previous stages. The classifier in each stage uses the classifier output of the previous stage as input. Adapted from Ref. [67].	26
4.2	Receiver operating characteristic of the Full Reconstruction and the Full Event Interpretation applied to B^0 meson and B meson candidates in recorded data. The tag-side efficiency and purity are measured by fitting the beam-constrained mass spectrum. Taken from Ref. [67].	27
5.1	Left: Differential decay rates for the fit of BSZ coefficients to $B \rightarrow \rho \ell \nu_\ell$ data performed in Ref. [19]. Figure adapted from Ref. [19]. Line shapes of the ρ meson in the EvtGen simulation and modeled according to Ref. [5, Chapter 50] by Ref. [74].	30
5.2	The spectrum of the hadronic mass m_X as generated by the BLNP model and fragmented by PYTHIA. As illustrated in the left figure, the $B \rightarrow \pi \pi \ell \nu_\ell$ distribution exhibits a step-like feature at the kinematic threshold of $B \rightarrow \pi \rho \ell \nu_\ell$ production. As this process is not measured, the spectrum is corrected as illustrated on the right.	32
5.3	Hybrid MC model combination for charged $B \rightarrow X_u \ell \nu_\ell$ decays projected to the reweighting variables with bins of the reweighting procedure indicated by green vertical lines. Shown are the lepton energy in the B frame E_ℓ^B , the momentum transfer q^2 , and the mass of the hadronic system M_X	34
6.1	Signature of a B_{sig} candidate decaying in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ channel on the right and a representative B_{tag} decay on the left. In the $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ channel, both the ρ meson on the signal side and the D meson on the tag side must be charged instead of neutral.	39
6.2	Variants of variables to estimate the squared missing mass of the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ signal process in the event, determined from simulation. $M_{\text{miss, simple}}^2$ uses no beam energy information, $M_{\text{miss, rescaled}}^2$ rescales the missing momentum to the expected momentum of a B meson and $M_{\text{miss, nominal}}^2$ follows Eq. (6.5).	43

- 6.3 Illustration of the topological differences between events with light-quark pairs and events with $B\bar{B}$ production in the cms. Final-state particles from light-quark pairs are highly boosted and produce a jet-like signature. $B\bar{B}$ pairs are produced almost at rest which distributes the final-state particles uniformly in the cms. Adapted from Refs. [84, 104]. 45
- 6.4 Properties of the MVA method used to reject continuum backgrounds in the (left) $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ and (right) $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ reconstruction. Shown are (top) the ROC curve for test and training samples together with the AUC value, (middle) the separation of continuum background and $B \rightarrow \rho \ell \nu_\ell$ signal in the test sample and (bottom) the agreement between Data and MC distributions. The hatched region in the bottom corresponds to rejected data, with the red line indicating the selection of $\mathcal{P}_{CS} > 0.6$ 49
- 6.5 Channel-dependent calibration factors [114, 110, 113] for (top) neutral and bottom (charged) B_{tag} candidates to correct for efficiency differences of the FEI in recorded and simulated data. Shown are factors derived in two independent analyses as well as a combination of them which takes into account correlated uncertainties and differences between the two calibration approaches. Empty markers correspond to the individual measurements while solid markers give the combined factors. The dashed line indicates unity with which calibration factors of several well-modelled FEI reconstruction channels are consistent. 54
- 6.6 Signal-extraction variables (top) M_{miss}^2 and (bottom) $M(\pi\pi)$ for the (left) $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ and (right) $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ reconstruction channels in recorded and simulated data with all normalizations set to the expectation in simulation as described above. The $B \rightarrow X_c \ell \nu_\ell$ component exhibits distinctly peaking behavior at the nominal ρ mass in $M(\pi\pi)$ as ρ mesons are reconstructed in background events. While the simulated data can reproduce the necessary structures in data, the comparison also indicates that substantial systematic uncertainties must be assigned to the simulated sample to e.g. account for the overestimation above the ρ mass in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ channel. The comparison in the $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ channel also shows that either the signal or the nonresonant $B \rightarrow \pi\pi \ell \nu_\ell$ component are overestimated. 56
- 6.7 Agreement between recorded and simulated data for $B \rightarrow D(\rightarrow \pi^\pm \pi^\mp) \ell \nu_\ell$ decays in very fine bins of (left) $M(\pi\pi)$ and (right) M_{miss}^2 . For $M(\pi\pi)$, an additional selection with $|M_{\text{miss}}^2| < 1.0 \text{ GeV}^2/c^4$ is applied to ensure only well-reconstructed decays. 57
- 7.1 Fit templates in the $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ reconstruction channel in M_{miss}^2 and $M(\pi\pi)$ for (from top left) signal, nonresonant $B \rightarrow \pi^+ \ell \nu_\ell$, $q\bar{q}$ and $B\bar{B}$ background processes. The normalization is set to the expectation in simulation. Illustrations of the templates with uncertainties are given in Chapter A. 66

- 7.2 Composition of the background templates in both $B \rightarrow \rho \ell \nu_\ell$ reconstruction channels, grouped in categories likely to produce peaking structures in $M(\pi\pi)$. Decays are identified by the mother particle of the dipion candidate in simulation. $B \rightarrow X_c \ell \nu_\ell$ processes modeled with EvtGen's Dalitz models are shown in **blue**, $B \rightarrow X_c \ell \nu_\ell$ decays differently are shown in **red** and combinatorial misreconstruction and misidentification is shown in **green**. The nonresonant component shown in **dark purple** originates from the BLNP model. The component shown in **light purple** contains, among others, resonant $B \rightarrow X_u \ell \nu_\ell$ backgrounds, beam background, and hadronic B decays with secondary leptons. More details on the simulation of these processes can be found in Section 5.3. 67
- 7.3 Four out of eight fit templates in M_{miss}^2 and $M(\pi\pi)$ in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ reconstruction channel with normalization set to the expectation in simulation. Illustrations of the templates with uncertainties are given in Chapter A. . . . 68
- 7.4 Four out of eight fit templates in M_{miss}^2 and $M(\pi\pi)$ in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ reconstruction channel with normalization set to the expectation in simulation. Illustrations of the templates with uncertainties are given in Chapter A. . . . 69
- 7.5 Projections of the fit variables (top) M_{miss}^2 in the full $M(\pi\pi)$ range and (bottom) $M(\pi\pi)$ in signal-rich regions of M_{miss}^2 , shown for the Asimov data sets in the (left) $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ and (right) $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ reconstruction channels. 71
- 7.6 Likelihood scans to simulated data in the extraction of the signal yield parameter for (top) $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ and (bottom) $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ decays. Shown are the nominal fit setup and a fit setup in which no systematic uncertainties other than MC statistics are considered. Also shown is the Hesse approximation for the nominal fit setup. The dashed vertical lines indicate the 1σ and 2σ intervals of the nominal model, the dashed horizontal line indicates the point at which the likelihood value increases by a factor of 0.5. 72
- 7.7 Normalized distributions of event yield and pull in the (left) $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ and (right) $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ signal extraction of the signal yield parameter, conducted using Poisson sampling of the nominal templates and template normalizations from MC expectation. Nuisance parameters in the fit are fixed to zero. The fit can reproduce the uncertainties but only reproduces the parameter of interest within 2.5 % of the total event yield. 73
- 7.8 Normalized distributions of event yield and pull in the (left) $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ and (right) $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ signal extraction of the signal yield parameter, conducted using Poisson sampling of templates modified by sampling the nuisance parameters and template normalizations from MC expectation. Nuisance parameters in the fit are floating. The fit can reproduce the uncertainties for both $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ channels and can completely reproduce the $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ signal yield. For $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$, the toys reproduce the signal yield within 1.6%. 76

- 8.1 Overestimation of the number of events in several bins of simulated data in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ reconstruction, amplitude fit to Ref. [79] and varied versions of the fit parameters [74]. Eigenvariation 3, which has the largest effect, is used as a systematic uncertainty estimate in the fit. 79
- 8.2 m_x , E_l^B and q^2 distributions in simulated data without reconstruction effects for 12 variations of the parameters used in the BLNP model, normalized to the distributions obtained with the nominal parameters. The region $m_X < 1.5 \text{ GeV } c^{-1}$ exhibits large systematic uncertainty which is reflected in the $B \rightarrow \pi \pi \ell \nu_\ell$ templates in the signal extraction (see Figs. A.2 and A.6). 81
- 8.3 Systematic uncertainty assigned to the shape of the $B \rightarrow \pi \pi \ell \nu_\ell$ template partially originating from fragmentation of the X_u system generated in the BLNP model. The template is allowed to fluctuate between the nominal shape and the overall m_X distribution. Full correlation is assumed between the bins. The figure shows a slight discrepancy between the assigned uncertainty and the sum of the stacked histograms which originates from binning effects. 82
- 8.4 19 independent (dashed) **up** and (solid) **down** variations of the parameters used in the $B \rightarrow \rho$ form factor, normalized to the form factor with best-fit parameters. 83
- 9.1 Likelihood profiles in the fit to data, determined by pointwise evaluation of the likelihood and fit with a bifurcated parabola. Good agreement is observed between pointwise evaluation and the Hesse matrix approximation within the 1σ interval in both $B \rightarrow \rho \ell \nu_\ell$ channels. In the $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ signal extraction, slight disagreement between the pointwise determination and the Hesse matrix approximation is observed in the interval -2σ to -1.5σ which is attributed to the parameter limits. 89
- 9.2 Projections of the fit variables (top) M_{miss}^2 in the full $M(\pi\pi)$ range and (bottom) $M(\pi\pi)$ in signal-rich regions of M_{miss}^2 , shown for the fit to data in the (left) $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ and (right) $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ reconstruction channels. The signal-rich regions correspond to -0.1 to $0.6 \text{ GeV}^2/c^4$ for $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ and -0.4 to $0.4 \text{ GeV}^2/c^4$ for $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ 90
- 9.3 Pulls on the 132 bin-wise, strongly correlated nuisance parameters for the signal and background templates in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ fit to data. The error bars correspond to the uncertainty determined in the fit. The solid vertical lines from top to bottom indicate bins in M_{miss}^2 . The signal template parameters exhibit pulls of up to 0.55σ in the third and fourth $M(\pi\pi)$ bin in each M_{miss}^2 bin due to the interference with ω mesons. The background template exhibits pulls with up to 1.28σ due to the uncertainty on the D meson decays and the large uncertainty from the MC sample size. 92

9.4	Pulls on the 121 bin-wise, strongly correlated nuisance parameters for the signal and background templates in the $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ fit to data. The error bars correspond to the uncertainty determined in the fit. The solid vertical lines from top to bottom indicate bins in M_{miss}^2 . The signal template parameters exhibit no strong pulls. The background template parameters exhibit pulls with up to 1.05σ due to the uncertainty on the D meson decays and the large uncertainty from the MC sample size.	93
9.5	The measured branching fractions of (top) $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ and (bottom) $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ compared to previous determinations, the PDG average and the HFLAV evaluation.	95
A.1	Two-dimensional projection of the signal template in the $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ signal extraction with relative systematic uncertainties in each bin. Further description is given in the text.	113
A.2	Two-dimensional projection of the nonresonant $B \rightarrow \pi\pi\ell\nu_\ell$ template in the $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ signal extraction with relative systematic uncertainties in each bin. Further description is given in the text.	114
A.3	Two-dimensional projection of the combined background template in the $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ signal extraction with relative systematic uncertainties in each bin. Further description is given in the text.	115
A.4	Two-dimensional projection of the continuum template in the $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ signal extraction with relative systematic uncertainties in each bin. Further description is given in the text.	116
A.5	Two-dimensional projection of the signal template in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ signal extraction with relative systematic uncertainties in each bin. Further description is given in the text.	118
A.6	Two-dimensional projection of the nonresonant $B \rightarrow \pi\pi\ell\nu_\ell$ template in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ signal extraction with relative systematic uncertainties in each bin. Further description is given in the text.	119
A.7	Two-dimensional projection of the combined background template in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ signal extraction with relative systematic uncertainties in each bin. Further description is given in the text.	120
A.8	Two-dimensional projection of the continuum template in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ signal extraction with relative systematic uncertainties in each bin. Further description is given in the text.	121
A.9	Two-dimensional projection of the $B \rightarrow f_0(500)\ell\nu_\ell$ template in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ signal extraction with relative systematic uncertainties in each bin. Further description is given in the text.	122

A.10	Two-dimensional projection of the $B \rightarrow f_0(980)\ell\nu_\ell$ template in the $B^+ \rightarrow \rho^0\ell^+\nu_\ell$ signal extraction with relative systematic uncertainties in each bin. Further description is given in the text.	123
A.11	Two-dimensional projection of the $B \rightarrow f_2(1270)\ell\nu_\ell$ template in the $B^+ \rightarrow \rho^0\ell^+\nu_\ell$ signal extraction with relative systematic uncertainties in each bin. Further description is given in the text.	124
A.12	Two-dimensional projection of the $B \rightarrow \rho^0(1450)\ell\nu_\ell$ template in the $B^+ \rightarrow \rho^0\ell^+\nu_\ell$ signal extraction with relative systematic uncertainties in each bin. Further description is given in the text.	125
B.1	Fit result to data in M_{miss}^2 shown in bins of $M(\pi\pi)$ in the range 0.48 to $0.80 \text{ GeV } c^{-2}$	127
B.2	Fit result to data in M_{miss}^2 shown in bins of $M(\pi\pi)$ in the range 0.80 to $1.07 \text{ GeV } c^{-2}$	128
B.3	Fit result to data in $M(\pi\pi)$ shown in bins of M_{miss}^2 in the range -1.0 to $0.6 \text{ GeV}^2/c^4$	129
B.4	Fit result to data in $M(\pi\pi)$ shown in bins of M_{miss}^2 in the range -1.0 to $0.6 \text{ GeV}^2/c^4$	130
B.5	Fit result to data in M_{miss}^2 shown in bins of $M(\pi\pi)$ in the range 0.48 to $0.80 \text{ GeV } c^{-2}$	132
B.6	Fit result to data in M_{miss}^2 shown in bins of $M(\pi\pi)$ in the range 0.48 to $0.80 \text{ GeV } c^{-2}$	133
B.7	Fit result to data in $M(\pi\pi)$ shown in bins of M_{miss}^2 in the range -1.0 to $0.2 \text{ GeV}^2/c^4$	134
B.8	Fit result to data in $M(\pi\pi)$ shown in bins of M_{miss}^2 in the range 0.2 to $2.0 \text{ GeV}^2/c^4$	135
C.1	Nuisance parameter pulls of the templates for (from top to bottom) $B^0 \rightarrow \rho^-\ell^+\nu_\ell$ signal, nonresonant X_u dipion backgrounds, $B\bar{B}$ backgrounds, and continuum backgrounds in the fit to data in the $B^0 \rightarrow \rho^-\ell^+\nu_\ell$ channel. The top row of indices below each template indicates the bins in $M(\pi\pi)$, the bottom row bins in M_{miss}^2	137
C.2	Nuisance parameter pulls of the templates for (from top to bottom) $B^+ \rightarrow \rho^0\ell^+\nu_\ell$ signal, nonresonant X_u dipion backgrounds, $B\bar{B}$ backgrounds, and continuum backgrounds in the fit to data in the $B^+ \rightarrow \rho^0\ell^+\nu_\ell$ channel. The top row of indices below each template indicates the bins in $M(\pi\pi)$, the bottom row bins in M_{miss}^2	139
C.3	Nuisance parameter pulls of the templates for (from top to bottom) $B \rightarrow f_0(500)\ell\nu_\ell$ backgrounds, $B \rightarrow f_0(980)\ell\nu_\ell$ backgrounds, $B \rightarrow f_2(1270)\ell\nu_\ell$ backgrounds, and $B \rightarrow \rho^0(1450)\ell\nu_\ell$ backgrounds in the fit to data in the $B^+ \rightarrow \rho^0\ell^+\nu_\ell$ channel. The top row of indices below each template indicates the bins in $M(\pi\pi)$, the bottom row bins in M_{miss}^2	140

-
- D.1 Correlation matrices for the FEI factors for B^+ tag candidates. Large correlation is observed for FEI channels dominated by the shared systematic uncertainty of the $B \rightarrow X_c l \nu_\ell$ branching fraction. 142
- D.2 Correlation matrices for the FEI factors for B^0 tag candidates. Large correlation is observed for FEI channels dominated by the shared systematic uncertainty of the $B \rightarrow X_c l \nu_\ell$ branching fraction. 142
- E.1 Hybrid MC model combination for neutral $B \rightarrow X_u l \nu_\ell$ decays projected to the reweighting variables with bins of the reweighting procedure indicated by green vertical lines. Shown are the lepton energy in the B frame E_l^B , the momentum transfer q^2 and the mass of the hadronic system M_X 144

List of Tables

3.1	Key machine parameters of KEKB and SuperKEKB. The KEKB parameters are those achieved while the SuperKEKB parameters are the design parameters.	12
3.2	Overview of the data set recorded by the Belle II experiment between 2019 and 2022. The first uncertainty is statistical, the second is systematic. The methodology with which the luminosity is measured is described in Ref [42]. The main data set for B flavor physics is recorded at the $\Upsilon(4S)$ resonance, with a supporting off-resonance data set recorded at beam energies 60 MeV below the resonance. The two “scan” data sets consist of data recorded at various beam energies and are needed to study heavy quarkonia and measure the mass and width of the $\Upsilon(4S)$ resonance.	16
3.3	Cache performance metrics, collected over 14 months and separated into types of jobs.	21
5.1	Branching fractions of exclusive $B \rightarrow X_u \ell \nu_\ell$ processes used in the simulated sample. The bottom four exclusive processes are completely unmeasured, the values here only divide the difference between the $B^+ \rightarrow \pi^\pm \pi^\mp \ell \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ branching fractions assuming a $2\times$ larger fraction for $B \rightarrow f_2(1270) \ell \nu_\ell$. For the $B^0 \rightarrow \pi^- \pi^0 \ell \nu_\ell$ branching fraction the value from PYTHIA fragmentation is used.	31
5.2	Branching fractions of exclusive $B \rightarrow X_c \ell \nu_\ell$ processes determined in Ref. [85] such that they sum to the measured inclusive $B \rightarrow X_c \ell \nu_\ell$ branching fraction given in the last row. The bottom two exclusive processes involving η mesons are completely unmeasured so 100% uncertainties must be assigned.	36
6.1	Momentum regions p_e^{\min} , p_e^{\max} and parameters of the bremsstrahlung correction as determined by optimizing the momentum resolution in simulated data.	41
6.2	Variables used to train the Continuum Suppression MVA method. A check mark indicates that the variable is derived from the corresponding candidate, a cross mark indicates that it is only used for the other B meson candidate in the event.	48

6.3	Expected number of events and reconstruction efficiencies in the reconstructed $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ sample. The reconstruction efficiencies are given with uncertainties purely originating from the size of the MC sample.	55
6.4	Expected number of events and reconstruction efficiencies in the reconstructed $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ sample. The reconstruction efficiencies are given with uncertainties purely originating from the size of the MC sample.	55
7.1	Expected yields in both $B \rightarrow \rho \ell \nu_\ell$ signal extraction channels with uncertainties determined from a fit to Asimov data.	70
7.2	Toy fit configurations in $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ for different factors $C_B \rightarrow X_u \ell \nu_\ell$ to the template normalizations expected from MC as well as averages determined from 300 toys. The table header identifies the templates by their hadronic component X_u	74
7.3	Ratio of fitted to expected template yields in the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ signal extraction in simulation. The fitted values are determined from 1000 toys in which only the resonant (Variant 1) or nonresonant (Variant 2) background $B \rightarrow \pi \pi \ell \nu_\ell$ template normalizations are allowed to float.	75
8.1	Branching fractions in the simulated sample and uncertainties for D^0 decays. The column ‘‘MC Model’’ corresponds to the decay model in EvtGen (also see Ref. [71]). If this column is ‘‘D_DALITZ’’ the given decay mode and branching fraction should be interpreted to include intermediate resonances, otherwise the branching fraction should be interpreted as exclusive.	84
8.2	Branching fraction in the simulated sample and uncertainties for D^+ decays. The column ‘‘MC Model’’ corresponds to the decay model in EvtGen. If this is ‘‘D_DALITZ’’ the given decay mode and branching fraction should be interpreted to include intermediate resonances. If the MC model is not a Dalitz model, the branching fraction should be interpreted as exclusive.	85
8.3	Sources of systematic uncertainty, quoted as a percentage of the $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ measured branching fractions. The template shape uncertainty is determined by fixing all shape nuisance parameters in the fit to data to zero and subtracting in quadrature the resulting uncertainty from the uncertainty of the nominal fit to data. The value given in the row labeled ‘‘Total (without shape)’’ is used as an uncertainty on the reconstruction efficiency in Table 9.3.	87
9.1	Ratio of obtained to expected template normalizations in a fit to data. As expected, the normalizations of several templates used for unmeasured background processes deviate from the expected normalization or are at the limit. As parameter values are not the final signal yields and are only used to re-determine the background template, they are given without uncertainties.	89

- 9.2 Yields from reconstructed $B \rightarrow \rho \ell \nu_\ell$ decays obtained from maximum likelihood fits to 362 fb^{-1} of data, with uncertainties including template shape uncertainties determined via likelihood profiling with MINOS. The normalizations of the nonresonant templates as well as that of the resonant $B \rightarrow f_0(980) \ell \nu_\ell$ processes are at their limits so likelihood profiling cannot be used to estimate the uncertainties. 91
- 9.3 Measured total branching fractions of $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ and $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ decays with statistical and systematic uncertainties, compared with world averages from the PDG [5] and the HFLAV [16]. For $B^0 \rightarrow \rho^- \ell^+ \nu_\ell$ the first uncertainty in the HFLAV evaluation is statistical and the second one is systematic, for the $B^+ \rightarrow \rho^0 \ell^+ \nu_\ell$ evaluation and the PDG average the given uncertainty combines both. The values of the parameters used in the measurement are also given with the uncertainties on $N_{\text{sig}}^{\text{data}}$ corresponding to the statistical and systematic uncertainty in the fit. The former is determined in a fit with fixed nuisance parameters, the latter is determined by subtracting in quadrature the statistical uncertainty from the value obtained with floating nuisance parameters. The uncertainty given for the efficiency incorporates all systematic effects without influence on the template shape. 96

