Partially censored
(contains unpublished work with Belle data)

Construction of Angular Observables Sensitive to New Physics in $\bar{B} \longrightarrow D^* \tau^- \bar{\nu}_\tau$ Decays

and

Measurements of Differential Cross Sections of $\bar{B} \longrightarrow D^* \ell^- \bar{\nu}_\ell$ Decays with Hadronic Tagging at Belle

Thesis submitted for the degree of Master of Science

Department of Physics
Ludwig Maximilian University of Munich

submitted by
Kilian Lieret

supervised by
Prof. Dr. Thomas Kuhr, Dr. Martin Jung, Prof. Dr. Gerhard Buchalla

submitted
Thursday, 6th September, 2018

Konstruktion Winkelabhängiger Observablen Sensitiv auf Neue Physik in $\bar{B} \longrightarrow D^* \tau^- \bar{\nu}_\tau$ Zerfällen

und

Messung Differenzieller Wirkungsquerschnitte des $\bar{B} \longrightarrow D^* \ell^- \bar{\nu}_\ell$ Zerfalls mit Hadronischen Tags am Belle Experiment

Masterarbeit eingereicht zur Erlangung des akademischen Grades Master of Science

Department für Physik
Ludwig-Maximilians-Universität München

eingereicht von
Kilian Lieret

betreut von
Prof. Dr. Thomas Kuhr, Dr. Martin Jung, Prof. Dr. Gerhard Buchalla

eingereicht
Donnerstag, 6. September 2018

Abstract

Recent measurements of $\bar{B}\longrightarrow D^{(*)}l^-\bar{\nu}_l$ cross sections at Belle, BaBar, and LHCb challenge lepton universality and thus the Standard Model at a combined confidence level close to four standard deviations. New measurements of differential decay rates could contribute to the understanding of these anomalies.

The differential cross section of the decay $\bar{B}\longrightarrow D^*(\rightarrow D\pi)l^-\bar{\nu}_l$ is parametrized according to different dependencies on the three decay angles and the coupling constants of potential new physics contributions. Observables using binned measurements of the differential cross section are characterized and explicitly constructed. Based on an estimate for the obtainable sensitivity, optimal binnings for such measurements are discussed. The discriminatory power of the thus constructed observables is discussed based on a basis of dimension six operators with renormalizable couplings contributing to $\bar{B}\longrightarrow D^*l^-\bar{\nu}_l$.

Furthermore, continuing work on an analysis of the $\bar{B}\longrightarrow D^*(\rightarrow D\pi)\ell^-\bar{\nu}_\ell$ decay channel for $\ell = e, \mu$ using data from the Belle detector at KEKB is presented. The events are selected from 772 million $e^+e^-\rightarrow\Upsilon(4S)\rightarrow B\bar{B}$ events, where one B meson is fully reconstructed in hadronic modes. Unfolded differential decay rates in four kinematic variables are presented separately for $\ell = e$, $\ell = \mu$ and a combined fit, allowing for precise calculations of V_{cb} and $B\longrightarrow D^*$ form factors. The new lepton flavor specific results are also expected to impact the discussion about potential light lepton flavor universality violations prompted by measurements of $B\longrightarrow K^{(*)}\ell\ell$ decays.

Conventions

Throughout this document, natural units $\hbar = c = 1$ are assumed. The convention for the weak hypercharge Y is fixed by the relation $Q = T_3 + Y$ to the electrical charge, Q, and the third component of the weak isospin, T_3. Unless otherwise declared, ℓ denotes either electrons or muons, whereas we take the letter l to mean any lepton.

A list of frequently used acronyms is given on page 157.

Contents

Chapter 1

Introduction

After being finalized in the 1970s, the Standard Model of particle physics (SM) [1, 2] has become the most successful theory to describe the observed particles and their interactions. However, the SM leaves a series of notable observations unexplained and is therefore considered to be only an incomplete theory. In particular, the SM fails to account for the total matter and energy content of the universe as estimated from astronomical observations, leading to the concept of *dark matter* and *dark energy* [3, 4]: The larger part of the universe appears to be still shrouded in mystery.

Despite these shortcomings, the predictions of the SM have been tested and confirmed to astonishing precision over the last 40 years [5], with the last building block, the Higgs boson, finally being discovered in 2012 [6, 7].

As far as collider experiments are concerned, one of the very few concrete tensions between measurements and the SM predictions can be found in decays of B mesons (particles consisting of two quarks), in particular in *semileptonic* decays $\bar{B} \longrightarrow D^{(*)} l^- \bar{\nu}_l$, where $l = e, \mu, \tau$ are *leptons*. In the SM, these decays are mediated by the exchange of a W boson, a particle that couples to all leptons with equal strength (*Lepton Flavor Universality*). Three experiments have recently performed precision measurements of this decay channel: BaBar [8] at the Stanford Linear Accelerator Center [9] (Stanford, USA), Belle [10] at the KEKB accelerator [11] (Tsukuba, Japan) and LHCb [12] at the Large Hadron Collider [13] (near Geneva, Switzerland). Testing the lepton flavor universality principle, they compared the relative frequency of this decay with tauons in the final state ($l = \tau$) to that of light leptons in the final state ($l = e, \mu$), that is, measured the ratio

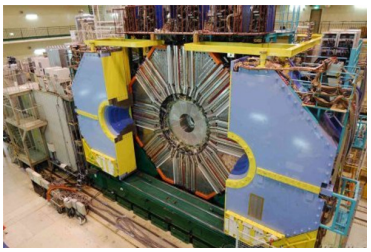
$$R(D^{(*)}) := \frac{\mathrm{BR}(\bar{B} \longrightarrow D^{(*)} \tau^- \bar{\nu}_\tau)}{\mathrm{BR}(\bar{B} \longrightarrow D^{(*)} \ell^- \bar{\nu}_\ell)}, \qquad \ell = e, \mu$$

of the corresponding branching fractions (BR, a measure of how often the B meson decays through a specific channel). Each of the three independent experiments reported an excess over the SM expectation [14–19]. Simply put, there are more tauons than we would expect based on the lepton flavor universality of the SM. Combining the measurements from all experiments, the significance of this anomaly has come close to the level of four standard deviations [20].

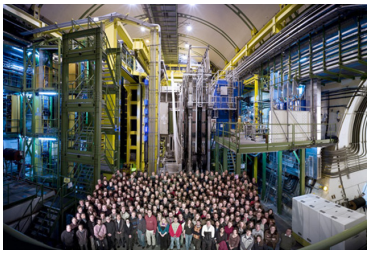
This observation could be a sign of new physics (NP) [21, 22], a hint at the involvement of so far unknown particles, possibly revolutionizing our understanding of particle physics!

In these exciting times, countless attempts have been made to explain the observations by extending the SM with additional particles: additional Higgs bosons [29], leptoquarks [30] or "copies" of the W bosons, to name just a few. While the available data is yet insufficient to discriminate between the different models, the upcoming Belle II experiment [31] at Super KEKB [32] and the continued operation of the LHCb experiment are expected to significantly increase the amount of data available for new analyses.

(a) BaBar [23]



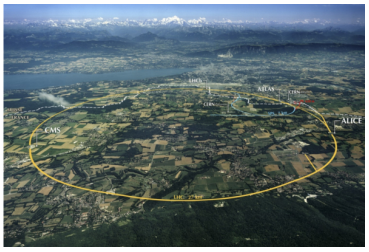
(b) Belle [24]



(c) LHCb [25]

(d) SLAC [26]

(e) KEKB [27]



(f) LHC [28]

Figure 1.1: Accelerators and detectors involved in measurements of $R(D^{(*)})$

First and foremost, this will allow for even more precise measurements of $R(D^{(*)})$, that will hopefully increase the statistical significance of the anomaly. However, the larger dataset will also allow considering new and related observables in this decay channel that are able to differentiate between contributions from specific NP models. Therefore, the construction of observables with high discriminatory power regarding NP contributions is an active field of research.

In this context, chapter 2 considers the specific decay channel $\bar{B}\longrightarrow D^*(\rightarrow D\pi)l^-\bar{\nu}_l$ and studies the dependency of the decay rate on three angles between the decay products (the *differential cross section*), taking into account potential NP. This angular dependency is characterized by a set of twelve coefficients (each depending on the NP coupling constants). Given a future measurement of the angular structure of the decay (given as a histogram with bins in the three angles), the construction of observables to extract these coefficients is studied.
In contrast to previous studies that only gave explicit examples of such observables as a proof of concept, the construction is formalized and requirements for the number of bins in the three angles of the measurement are discussed. Furthermore, the expected level of uncertainty on the observables is estimated and minimized over certain degrees of freedom in the construction of the observables. The influence of the bin spacings of the measurement histogram on the obtainable sensitivity is studied and optimal bin spacings for experimental measurements are determined. Furthermore, the usefulness of the studied observables to discriminate between potential NP contributions is discussed based on a basis of dimension six operators with renormalizable couplings contributing to $\bar{B}\longrightarrow D^*\ell^-\bar{\nu}_\ell$.

In chapter 3, ongoing work on an analysis of the $\bar{B}\longrightarrow D^*(\rightarrow D\pi)\ell^-\bar{\nu}_\ell$ decay channel for light leptons ($\ell = e, \mu$) using data from the Belle detector is presented. While potential NP contributions that could explain the $R(D^*)$ tension are generally thought to mostly affect the decay with τ leptons, measurements of this channel are used to constrain certain parameters (*form factors*) necessary for the theoretical prediction of $R(D^*)$. In fact, the published preliminary results of the analysis updated here are currently being used as experimental input for all $R(D^*)$ predictions that are included in the $R(D^*)$ average calculated by the HFLAV collaboration. In the analysis, the events are selected from 772 million $e^+e^- \rightarrow \Upsilon(4S) \rightarrow B\bar{B}$ events, where one B meson is fully reconstructed in hadronic modes. While the results had previously been reported without distinguishing between the lepton flavor, this update presents the unfolded differential decay rates in four kinematic variables separately for $\ell = e$ and $\ell = \mu$ in the final state. These new lepton flavor specific results will allow for additional tests of potential non-flavor universal

NP in the light lepton modes. This is in particular interesting in light of recent measurements of $B \longrightarrow K^{(*)}\ell\ell$ by the LHCb collaboration [33, 34], that are in tension with the light lepton flavor universality of the SM.

In this first chapter, some basic concepts are introduced, starting with a quick review of the SM and its limitations and followed by a more detailed introduction to the anomalies in semileptonic B decays. After some general remarks about extensions of the SM and NP in $\bar{B} \longrightarrow D l^- \bar{\nu}_l$, three prominent NP models that offer to explain the $R(D^{(*)})$ tension are introduced.

1.1 The Standard Model of Particle Physics

The Standard Model of particle physics (SM) [1, 2, 35] is a renormalizable gauge quantum field theory built around the internal symmetry group $\mathsf{SU}(3)_C \times \mathsf{SU}(2)_W \times \mathsf{U}(1)_Y$. Despite several shortcomings, it is the most successful theory of elementary particles and their interactions to date.

1.1.1 Field content of the SM

The particle content of the SM can be divided into two groups corresponding to their spin: *Fermions* (particles with half-integer spin) and *Bosons* (particles with integer spin):

- The SM fermions have spin 1/2 and appear in two groups: *Leptons* (integer electrical charge, no color charge) and *quarks* (fractional electrical charge, color charge). All fermions can be subdivided in three different generations, each containing two quarks and two leptons of different electrical charge Q: u, c, t (*up-type quarks*, $Q = +2/3$) and d, s, b (*down-type quarks*, $Q = -1/3$), e, μ, τ (*charged leptons*, $Q = -1$) and ν_e, ν_μ, ν_τ (*neutrinos*, $Q = 0$). These six different types of leptons and quarks are called the lepton/quark *flavors*. Furthermore, each fermion has two chirality components, called *left-handed* and *right-handed* and denoted with lower index L and R, respectively. However, note that the SM does not include right-handed neutrinos. Each fermion is partnered with an *antifermion* of opposite charge and chirality.

- Bosons can be further subdivided into the spin one (*vector*) gauge bosons and the spin zero (*scalar*) Higgs particle H. The gauge bosons W and B are responsible for the electroweak interactions, the gauge bosons G (*gluons*) for the strong force.

1.1.2 Gauge symmetries and spontaneous symmetry breaking

The gauge symmetry group of the SM is given by $\mathsf{SU}(3)_C \times \mathsf{SU}(2)_W \times \mathsf{U}(1)_Y$, where $\mathsf{SU}(\mathsf{n}) := \{A \in \mathbb{C}^{n\times n} \,|\, \det A = 1\}$ and $\mathsf{U}(1) := \{z \in \mathbb{C} \,|\, |z| = 1\}$. Each of the three symmetry groups corresponds to gauge bosons via the gauge principle: $\mathsf{SU}(3)$ to G, $\mathsf{SU}(2)$ to W and $\mathsf{U}(1)$ to B.

The transformation properties of a field under this symmetry group dictate its interactions with other particles. In the following, we use the shorthand notation $(\mathbf{m}, \mathbf{n})_Y$ for a particle which transforms as an m-tuplet under $\mathsf{SU}(3)$, an n-tuplet under $\mathsf{SU}(2)$ and carries $\mathsf{U}(1)$ hypercharge Y. The fields of the SM together with their representations under $\mathsf{SU}(3)_C \times \mathsf{SU}(2)_W \times \mathsf{U}(1)_Y$ are summarized in table 1.1.

To avoid breaking gauge symmetries explicitly, masses in the SM are introduced via the Higgs mechanism, proposed almost simultaneously by three independent groups in 1964 [36–38]: Instead of directly including mass terms, an additional complex scalar field H (*Higgs field*) is introduced, with gauge-invariant couplings to both fermions and gauge bosons. The Higgs potential $V(H^\dagger H)$ is chosen to have its global minima away from the origin, such that H acquires a non-zero vacuum expectation value, thereby effectively breaking the gauge symmetry (*Spontaneous Symmetry Breaking* $\mathsf{SU}(3)_C \times \mathsf{SU}(2)_W \times \mathsf{U}(1)_Y \longrightarrow \mathsf{SU}(3)_C \times \mathsf{U}(1)_{\mathrm{em}}$) and giving rise to terms that can be

Fermions ($S = 1/2$)		Gauge Fields ($S = 1$)		Scalars ($S = 0$)	
L_L	$(\mathbf{1},\mathbf{2})_{-1/2}$	B	$(\mathbf{1},\mathbf{1})_0$	H	$(\mathbf{1},\mathbf{2})_{1/2}$
Q_L	$(\mathbf{3},\mathbf{2})_{1/6}$	W	$(\mathbf{1},\mathbf{3})_0$		
ν_R	$(\mathbf{1},\mathbf{1})_0$	G	$(\mathbf{8},\mathbf{1})_0$		
e_R	$(\mathbf{1},\mathbf{1})_{-1}$				
u_R	$(\mathbf{3},\mathbf{1})_{2/3}$				
d_R	$(\mathbf{3},\mathbf{1})_{-1/3}$				

Table 1.1: The fields of the standard model and their representation under $\mathsf{SU}(3)_C \times \mathsf{SU}(2)_W \times \mathsf{U}(1)_Y$. The weak hypercharge convention is fixed by $Q = T_3 + Y$, where Q is the electrical charge and T_3 is the third component of the weak isospin (zero for $\mathsf{SU}(2)$ singlets and $\pm 1/2$ for $\mathsf{SU}(2)$ doublets). Note that by CPT symmetry, there is another set of charge-conjugated fermions of opposite chirality and quantum numbers.

identified with the gauge boson and fermion masses (however, no such coupling is introduced for neutrinos in the SM).

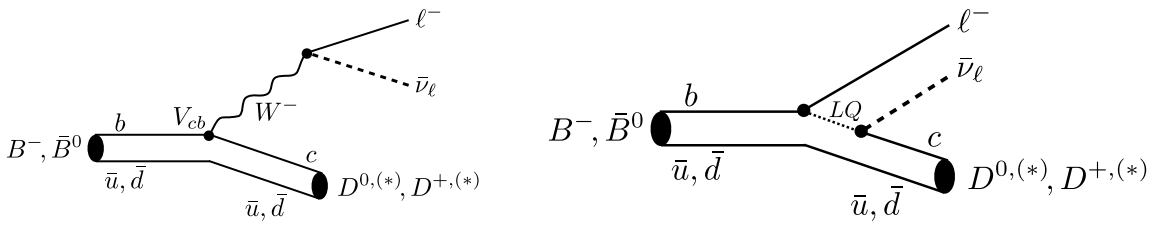
The presence of quartic interactions between one degree of freedom of the W boson (usually denoted as W^3), the B boson and two Higgs bosons imply that W^3 and B are not mass-eigenstates. Rather, they mix to the linear combinations Z and γ, the (electrically) neutral massive gauge boson and the photon. The other degrees of freedom of the W boson then give rise to the $W^\pm$ boson mass eigenstates.

Similarly, the coupling of quarks to the Higgs field as parametrized by *Yukawa matrices* implies that mass eigenstates are not flavor eigenstates (which are also called *interaction eigenstates*, as they are defined by their couplings to the gauge bosons). Rather, each mass eigenstate is a linear combination of the flavor eigenstates and vice-versa. In contrast, because the SM does not include right-handed neutrinos or neutrino masses, the distinction between mass and flavor eigenstates does not have to be made for charged leptons.

In the SM, the couplings of the (flavor eigenstate) leptons to the gauge bosons are flavor-independent (*lepton universality*). Interactions with Z bosons (*neutral current*) cannot change fermion flavor, whereas the $W^\pm$ boson (*charged current*) mediates between charged leptons and neutrinos as well as up and down-type quarks of the same generation. Because of the mismatch of mass and flavor eigenstates in the quark sector, the charged current also effectively mixes quark generations. For the quark sector, this effect is described by the Cabibbo–Kobayashi–Maskawa (CKM) matrix.[1]

Processes that change fermion flavor but not charge are called *flavor changing neutral currents* (FCNCs). In the SM, they only occur beyond tree level (that is, in decay topologies containing loops) and are strongly suppressed by the Glashow–Iliopoulos–Maiani (GIM) mechanism [40]. Experimental results tightly constrain additional sources of FCNCs, making the suppression of FCNCs an important criterion for extensions of the SM.

[1] If V_u and V_d are the unitary matrices that rotate a left-handed up and down-type quark in flavor basis to the corresponding state in mass basis, then the CKM matrix is given as $V = V_u^\dagger V_d$. This also points to the reason of why the neutral current cannot mix quark generations: As the neutral current does not mix between up- and down-type quarks, the rotations between mass and flavor eigenstates cancel in the description of the process: $V_u^\dagger V_u = \mathbb{1}$ and $V_d^\dagger V_d = \mathbb{1}$. In extensions of the standard model that feature massive neutrinos of both chiralities, the same rationale leads to mismatched mass and flavor eigenstates for the leptons and the pendant of the CKM matrix, the Pontecorvo–Maki–Nakagawa–Sakata matrix (PMNS matrix) [39].



(a) SM process (b) Mediated by leptoquarks

Figure 1.2: Two possible Feynman diagrams of $\bar{B} \longrightarrow D^* l^- \bar{\nu}_l$ [46].

1.1.3 Shortcomings of the SM

While the SM is the most successful theory of particle physics to date, it is commonly accepted that it is not a complete theory. Besides not including a description of gravity, the failure to accommodate neutrino oscillations[2] due to the absence of neutrino masses is a striking flaw.[3] Furthermore, the SM conflicts with cosmological studies of the mass and energy content of the universe (leading to the concept of *dark matter* and *dark energy* [3, 4]) as well as with the observed matter-antimatter asymmetry [43]. From a theoretical perspective, the SM also fails to provide satisfactory explanations for some of its features that appear to be unnaturally "fine tuned" (e.g. the hierarchy problem or the strong CP problem).

On the side of most concrete experimental measurements, the SM has been extremely successful. However, tensions with the SM predictions have been found in the anomalous magnetic dipole moment of the muon [44], measurements of the proton radius with muonic hydrogen[4] [45] and in B meson decays (described in more detail in the next section).

1.2 Anomalies in B decays

Semileptonic decays of B mesons are well understood in the SM as processes being mediated by a W boson. Particularly precise measurements of the branching ratios to $\bar{B} \longrightarrow D^{(*)} \ell^- \bar{\nu}_\ell$ for $\ell = e, \mu$ have been performed at the B factories Belle and BaBar, with the measurement of the CKM matrix element V_{cb} as one of its primary motivations.

Measurements of the analog decay channel with tauons in the final state, $\bar{B} \longrightarrow D^{(*)} \tau^- \bar{\nu}_\tau$, are experimentally more challenging due to the elusiveness of the short-lived τ lepton that leads to additional neutrinos. While some inclusive and semi-inclusive measurements of $\mathrm{BR}(b \longrightarrow X \tau^- \bar{\nu}_\tau)$ had been performed at the LEP collider [47] from 1996 on [48], the exclusive decay channel $\bar{B} \longrightarrow D^* \tau^- \bar{\nu}_\ell$ was first measured in 2007 by the Belle and BaBar collaborations [49, 50].

The semitauonic decay channel is sensitive to additional amplitudes that are suppressed by the lepton masses in the light lepton channels and allows to probe for additional new physics contributions. In particular, the possibility to constrain models with an extended Higgs sector served as a prime motivation to compare semileptonic and semitauonic B decays.

It is of advantage to consider the ratio of the branching fractions of the tau and light lepton

[2] First predicted by Bruno Pontecorvo [41], neutrino oscillations have been confirmed by studies of solar and atmospherical neutrinos as well as by reactor and accelerator experiments. For the experimental status as of late 2018 see e.g. [42].

[3] However, it should be noted that very simple extensions of the SM with massive neutrinos exist. Therefore some authors use a slightly more general definition of the SM that also includes neutrino masses, even though their structure has not been completely understood so far.

[4] The proton radius can be separately calculated from studies of hydrogen atoms and from studies of muon-proton systems (*muonic hydrogen*). However, current measurements find a 4% difference between both methods, a tension of more than five standard deviations.

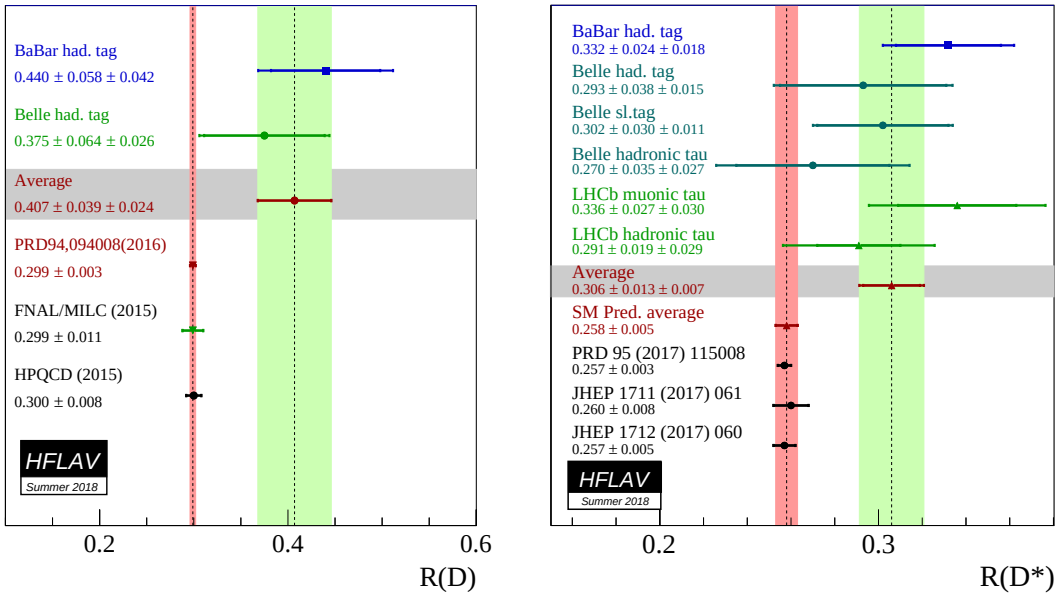


Figure 1.3: Separate averages for $R(D)$ and $R(D^*)$ as calculated by the HFLAV group [20].

mode

$$R(D^{(*)}) := \frac{\mathrm{BR}(\bar{B}\longrightarrow D^{(*)}\tau^-\bar{\nu}_\tau)}{\mathrm{BR}(\bar{B}\longrightarrow D^{(*)}\ell^-\bar{\nu}_\ell)}, \tag{1.2.1}$$

as several sources of systematic error cancel in both theoretical predictions[5] (V_{cb}, form factor normalizations etc) and experimental measurements (detector efficiencies etc).

As of summer 2018, the HFLAV collaboration quotes the theoretical predictions [20]

$$R(D)_{\mathrm{theo}} = 0.299 \pm 0.003, \qquad R(D^*)_{\mathrm{theo}} = 0.258 \pm 0.005, \tag{1.2.2}$$

an arithmetic average of [51] ($R(D)$ only), [52] ($R(D^*)$ only) and [53, 54]. In these analyses, the $R(D)$ value is calculated from the lattice results [55, 56], but a fit to $\bar{B}\longrightarrow D\ell^-\bar{\nu}_\ell$ measurements from BaBar [57] and Belle [58] allowed for significantly reduced uncertainties [51]. The 2017 value of $R(D^*)$ had been based on the analysis [59], which performed a parametrization of form factors based on the work of Caprini, Neubert, Lellouch [60] (*CLN parametrization*). In contrast, the 2018 value is obtained using a fit of the unfolded spectrum of the Belle analysis [61] (subject of chapter 3) to form factors using the parametrization of Boyd, Grinstein and Lebed [62–64] (*BGL parametrization*).

The experimental values are averaged from analyses of Belle, Babar and LHCb [14–19], resulting in [20]:

$$R(D)_{\mathrm{exp}} = 0.407 \pm 0.039_{\mathrm{stat}} \pm 0.024_{\mathrm{syst}}, \qquad R(D^*)_{\mathrm{exp}} = 0.306 \pm 0.013_{\mathrm{stat}} \pm 0.007_{\mathrm{syst}}, \tag{1.2.3}$$

with a correlation of -20% between both results. Theoretical predictions and measurements of $R(D)$ and $R(D^*)$ are shown separately in figure 1.3, a joint fit in figure 1.4. The values of $R(D)$ and $R(D^*)$ exceed the SM predictions by 2.3σ and 3.0σ, respectively and the combined fit currently is in a 3.78σ tension with the SM [20].

Similar studies comparing the branching ratios with electrons and muons in the final states did so far not find tensions to e/μ universality [65, 66].

[5] Table A.2 (page 109) compares the errors on $R(D^*)$ with those of the corresponding branching ratios as used in chapter 2, demonstrating this effect.

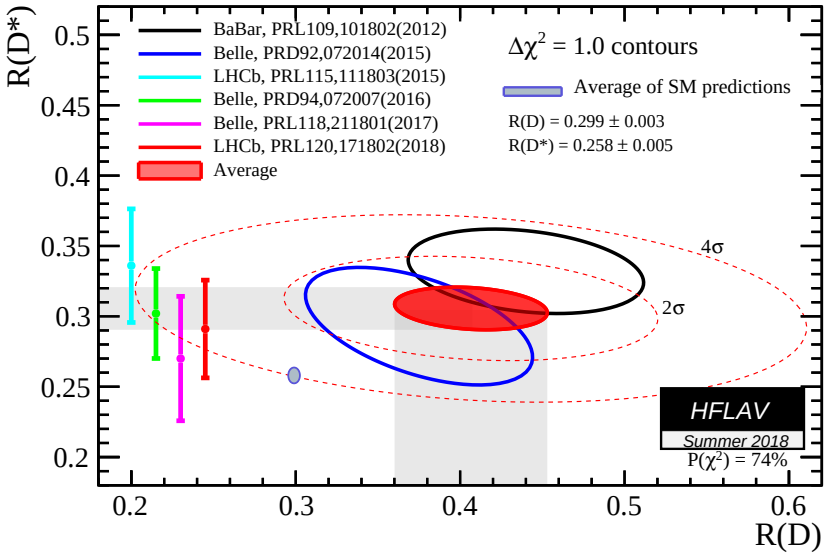
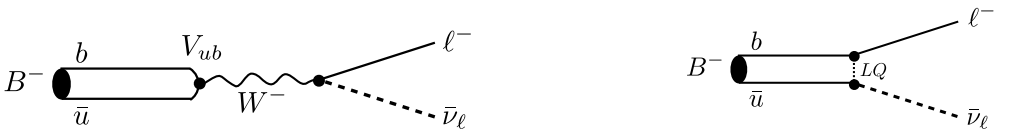


Figure 1.4: Joint fit to $R(D^{(*)})$ by the HFLAV group [20].



(a) SM process (b) Mediated by leptoquarks

Figure 1.5: Two possible Feynman diagrams of $\bar{B}\longrightarrow D^* l^- \bar{\nu}_l$ [46].

A different, but related decay channel that also drew some attention is $B^-\longrightarrow\tau^-\bar{\nu}_\tau$, depicted in figure 1.5. The average of $\mathrm{BR}(B^-\longrightarrow\tau^-\bar{\nu}_\tau)$ is $(1.09\pm0.24)\times10^{-4}$ [5] based on [67, 68], exceeding the SM expectation of $\mathrm{BR}(B^-\longrightarrow\tau^-\bar{\nu}_\tau) = 0.75^{+0.10}_{-0.05}\times10^{-4}$ [46, p. 2].[6]

A similar variable to $R(D^{(*)})$ can be measured in leptonic decays of the B meson with $b\longrightarrow s$ transitions [70]:

$$R(K^{(*)}) := \frac{\mathrm{BR}(B\longrightarrow K^{(*)}\mu\mu)}{\mathrm{BR}(B\longrightarrow K^{(*)}ee)}.$$

The theoretical SM expectation of this ratio has been calculated to 1 with errors of $\mathcal{O}(1\%)$. The most precise experimental measurements are from the LHCb collaboration and are in 2.6σ ($R(K)$, [33]) and more than 2σ ($R(K^*)$, [34]) tension with the SM. The $B\longrightarrow K^{(*)}\ell\ell$ decay channel was also studied by Belle [71], CMS [72] and ATLAS [73], albeit with lower precision. The implications of these results for searches of NP are for example discussed in [74].

1.3 Physics beyond the Standard Model

Due to the shortcomings described in 1.1.3, the SM is considered to be only an effective theory, valid for energies below a certain energy scale Λ. A model describing new physics (NP) beyond the SM is commonly expected to satisfy the following conditions [75]:

1. Its gauge group should contain the SM symmetry group $\mathsf{SU}(3)_C\times\mathsf{SU}(2)_L\times U(1)_Y$
2. The degrees of freedom of the SM should be included as composite or fundamental fields

[6] Note that the value of $(1.44\pm0.31)\times10^{-4}$ [69] as currently given by the HFLAV collaboration seems to be the result of an incorrect averaging of the same two analyses cited by [5].

3. At energies $\ll \Lambda$, the SM should be recovered (but for yet to be discovered weakly coupling light particles, e.g. sterile neutrinos).

Most NP models implement the third property by a perturbative approach motivated by the Appelquist-Carazzone theorem [76], by which the low energy effects of heavy, yet to be discovered particles can be taken into account by shifts of the particle masses and interaction strengths. This means that the Lagrangian is in fact given by the following expansion

$$\mathcal{L} = \mathcal{L}^{(4)}_{\text{SM}} + \frac{1}{\Lambda}\sum_k \mathcal{C}^{(5)}_k \mathcal{O}^{(5)}_k + \frac{1}{\Lambda^2}\sum_k \mathcal{C}^{(6)}_k \mathcal{O}^{(6)}_k + \mathcal{O}(\Lambda^{-3}), \tag{1.3.1}$$

where $\mathcal{L}^{(4)}_{\text{SM}}$ is the "usual" part of the SM Lagrangian, containing only operators with (energy) dimension two and four. NP then leads to higher order operators $\mathcal{O}^{(n)}_k$ with dimensionless coupling constants (Wilson coefficients) $\mathcal{C}^{(n)}_k$.

A full set of dimension five and six operators under these assumptions can be found in [75].

1.3.1 New physics in $\bar{B}\longrightarrow D^*\tau^-\bar{\nu}_\tau$

In the following, we will focus on NP models that offer explanations to the anomaly found in $\bar{B}\longrightarrow D^{(*)}l^-\bar{\nu}_l$ decays. Note that $\bar{B}\longrightarrow D^{(*)}l^-\bar{\nu}_l$ actually denotes two decay channels:

$$(\bar{u}b)B^- \longrightarrow (\bar{u}c)D^{0(*)}l^-\bar{\nu}_l \quad \text{and} \quad (\bar{d}b)\bar{B}^0 \longrightarrow (\bar{d}c)D^{+(*)}l^-\bar{\nu}_l \tag{1.3.2}$$

and analogously for the CP conjugated mode $B\longrightarrow D^{(*)}l^+\nu_l$.

Both decays feature $b\longrightarrow c$ transitions, while the $\bar{u}$ or $\bar{d}$ quark seems to remain unchanged. Throughout the thesis, we therefore assume that

1. NP contributions only affect the transition $b\longrightarrow cl^-\bar{\nu}_l$.

In fact, most analyses make use of *Heavy Quark Effective Field Theory* (HQET), where the decay $\bar{B}\longrightarrow D^*\ell^-\bar{\nu}_\ell$ is described by an expansion in $\Lambda_{\text{QCD}}/m_{b,c}$ and the effect of the *spectator quark* ($\bar{u}$ or $\bar{d}$) is neglected to first order [77].

If we furthermore make the following two assumptions regarding possible Feynman diagrams of the considered NP contributions:

2. All vertices are 3-vertices (3-couplings)[7]
3. There is only one intermediate (NP) particle X,

only three different interaction topologies are possible, depicted in figure 1.6:

$$b\bar{c}\longrightarrow X\longrightarrow \ell\bar{\nu}, \tag{1.3.3a}$$
$$b\nu\longrightarrow X\longrightarrow c\ell, \tag{1.3.3b}$$
$$b\bar{\ell}\longrightarrow X\longrightarrow c\bar{\nu}, \tag{1.3.3c}$$

where X is a particle of charge $Q = -1$, $-1/3$ or $2/3$ respectively. Considering the quantum numbers of the SM particles interacting with the NP particle for different chiralities in each diagram, the hypercharge Y of the NP particle can be read of directly. The results are shown in table 1.2.

[7] Note that this requirement also follows from dimensional considerations of the Lagrangian.

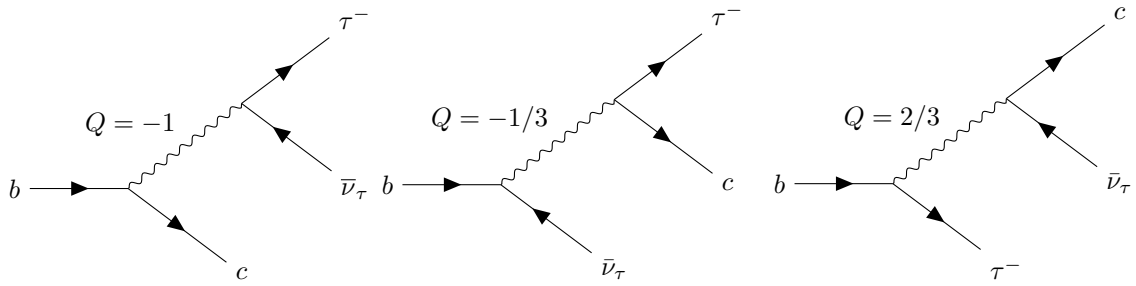


Figure 1.6: Feynman diagrams for possible production of a NP particle in $b \longrightarrow c\tau\bar{\nu}_\tau$.

	Hypercharge Y for		
Chirality	$Q = -1$	$Q = 1/3$	$Q = 2/3$
LL	0	$-1/3$	$2/3$
LR	$-1/2$	$-5/6$	$1/6$
RL	$-1/2$	$1/6$	$11/6$
RR	-1	$-1/3$	$2/3$

Table 1.2: Hypercharge of the new physics particle X in the interactions depicted in figure 1.6 based on the chirality structure of the interaction.

	Color Charge	
Spin	Yes	No
0 ("Scalar")	LQ	2HDM
1 ("Vector")		W'

Table 1.3: New Physics Models relevant for $\bar{B} \longrightarrow D^{(*)}\tau^-\bar{\nu}_\tau$ grouped by spin and color charge.

In order to explain $R(D^{(*)})$, the coupling of the NP particle X furthermore has to favor τ over light leptons.

The paper [22] characterizes all SM gauge group invariant interactions of dimension six relevant for $\bar{B} \longrightarrow D^*\tau^-\bar{\nu}_\tau$ by their transformation under $\mathsf{SU}(3)_C \times \mathsf{SU}(2)_W \times \mathsf{U}(1)_Y$. For the interaction (1.3.3a) (which has the same topology as the SM channel mediated by the W boson), the only options are $(\mathbf{1}, \mathbf{3})_0$ (same as the SM W boson) and $(\mathbf{1}, \mathbf{2})_{1/2}$. Additional W-like bosons are commonly denoted as W' and are discussed in more detail in section 1.3.3. Particles transforming as $(\mathbf{1}, \mathbf{2})_{1/2}$ can arise from extended Higgs sectors, in particular *Two Higgs Doublet Models* that are considered in section 1.3.2.

The other two interaction topologies (1.3.3b) and (1.3.3c) require the new exchange boson to have quark-lepton couplings. The transformation properties in this case are $(\mathbf{3}, \mathbf{3})_{2/3}$, $(\mathbf{3}, \mathbf{1})_{2/3}$, $(\mathbf{3}, \mathbf{2})_{7/6}$, $(\bar{\mathbf{3}}, \mathbf{2})_{5/6}$ (mistakenly referred to as $(\bar{\mathbf{3}}, \mathbf{2})_{5/3}$ in [22]), $(\bar{\mathbf{3}}, \mathbf{3})_{1/3}$ and $(\bar{\mathbf{3}}, \mathbf{1})_{1/3}$. Particles with these properties are generally called *leptoquarks*, which can be both scalar or vector particles and are discussed in section 1.3.4.

A quick categorization of these three models is shown in table 1.3.

Table 1.4 shows all vector fields that can have linear and renormalizable interactions to SM fermions and are invariant under the full SM gauge group.

Vector	Representation	Note
$\mathcal{B}_\mu$	$(\mathbf{1},\mathbf{1})_0$	(Z'; $Q=0$)
$\mathcal{B}^1_\mu$	$(\mathbf{1},\mathbf{1})_1$	Candidate W' [78]
$\mathcal{W}_\mu$	$(\mathbf{1},\mathrm{Adj})_0$	Candidate W'
$\mathcal{W}^1_\mu$	$(\mathbf{1},\mathrm{Adj})_1$	(no fermion couplings)
$\mathcal{G}_\mu$	$(\mathrm{Adj},\mathbf{1})_0$	(only quark couplings)
$\mathcal{G}^1_\mu$	$(\mathrm{Adj},\mathbf{1})_1$	(only quark couplings)
$\mathcal{H}_\mu$	$(\mathrm{Adj},\mathrm{Adj})_0$	(only quark couplings)
$\mathcal{L}_\mu$	$(\mathbf{1},\mathbf{2})_{-3/2}$	(no renorm. couplings to q [78])
$\mathcal{U}^2_\mu$	$(\mathbf{3},\mathbf{1})_{2/3}$	U_1 LQ
$\mathcal{U}^5_\mu$	$(\mathbf{3},\mathbf{1})_{5/3}$	($\tilde{U}_1$ LQ; $Q=5/3$)
$\mathcal{Q}^1_\mu$	$(\mathbf{3},\mathbf{2})_{1/6}$	($\tilde{V}_2$ LQ; only coupling $\bar{\ell}_L\gamma^\mu u^c_R$ [79])
$\mathcal{Q}^5_\mu$	$(\mathbf{3},\mathbf{2})_{-5/6}$	$\bar{V}_2$ LQ
$\mathcal{X}_\mu$	$(\mathbf{3},\mathrm{Adj})_{2/3}$	U_3 LQ
$\mathcal{Y}^1_\mu$	$(\bar{\mathbf{6}},\mathbf{2})_{1/6}$	(only quark couplings)
$\mathcal{Y}^5_\mu$	$(\bar{\mathbf{6}},\mathbf{2})_{-5/6}$	(only quark couplings)

Table 1.4: All irreducible representations of vector fields that can have linear and renormalizable interactions and are invariant under the full SM gauge group [80, p. 4]. Hypothetical particles that are interesting for $\bar{B}\longrightarrow D^{(*)}\tau^-\bar{\nu}_\tau$ are highlighted.

1.3.2 Two Higgs Doublet Models

Extensions of the SM Higgs sector were among of the first models to be discussed in the context of the anomalies in semileptonic B decays. A particularly simple such extension is given by Two Higgs Doublet Models. A thorough introduction can be found in [81], which serves as a source throughout this section.

The SM Higgs sector has the simplest possible structure, featuring only a single $\mathsf{SU}(2)$ doublet. An important constraint for any extension is the parameter ρ, defined as [82, p. 52]

$$\rho = \frac{M_W^2}{M_Z^2\cos^2\theta_\mathrm{W}}\,. \tag{1.3.4}$$

In the SM at tree level, this value is $\rho = 1$ which is in good agreement with the 2017 world average of 1.00037 ± 0.00023, only about $1.6\,\sigma$ above the SM expectation [83, p. 36]. Extensions of the scalar sector by additional multiplets modify this value to [83, p. 36]

$$\rho = \frac{\sum_{i=1}^n [I_i(I_i+1) - Y_i^2]|v_i|^2}{2\sum_{i=1}^n Y_i^2|v_i|^2}, \tag{1.3.5}$$

where I_i is the weak isospin, Y_i the weak hypercharge and v_i the vacuum expectation value of the neutral component of the multiplet. Thus, $\mathsf{SU}(2)$ singlets with $Y=0$ and $\mathsf{SU}(2)$ doublets with $Y=\pm1/2$ both do not change the expectation of $\rho=1$, as $I_i(I_i+1)-Y_i^2 = 2Y_i^2$. Therefore, the simplest way to extend the scalar sector of the SM while staying compatible with $\rho=1$ is to add additional scalar singlets or doublets. Two Higgs Doublet Models (2HDMs) are a particularly simple extension, adding only one more color neutral doublet, $(\mathbf{1},\mathbf{2})_{1/2}$. 2HDMs are a common ingredient of supersymmetric models, axion models and could shed light on the baryon asymmetry of the universe by introducing additional sources of CP symmetry breaking [81].

Assuming renormalizability, the potential of a 2HDM is given by [84]

$$\begin{aligned} V_{\text{2HDM}} = & m_{11}^2 \Phi_1^\dagger \Phi_1 + m_{22}^2 \Phi_2^\dagger \Phi_2 - \left[m_{12}^2 \Phi_1^\dagger \Phi_2 + \text{h.c.} \right] \\ & + \frac{1}{2}\lambda_1 \left(\Phi_1^\dagger \Phi_1 \right)^2 + \frac{1}{2}\lambda_2 \left(\Phi_2^\dagger \Phi_2 \right)^2 + \lambda_3 \left(\Phi_1^\dagger \Phi_1 \right) \left(\Phi_2^\dagger \Phi_2 \right) + \lambda_4 \left(\Phi_1^\dagger \Phi_2 \right) \left(\Phi_2^\dagger \Phi_1 \right) \\ & + \left\{ \frac{1}{2}\lambda_5 \left(\Phi_1^\dagger \Phi_2 \right)^2 + \left[\lambda_6 \left(\Phi_1^\dagger \Phi_1 \right) + \lambda_7 \left(\Phi_2^\dagger \Phi_2 \right) \right] \left(\Phi_1^\dagger \Phi_2 \right) + \text{h.c.} \right\}, \end{aligned} \tag{1.3.6}$$

with a total of 14 real parameters.

Usually, it is assumed that the Higgs sector is free of CP violation (necessary to even distinguish between scalars and pseudoscalars) and that the CP symmetry is also not spontaneously broken. Another common assumption is $\lambda_6 = 0 = \lambda_7$ (by enforcing discrete symmetries that are only potentially softly broken in the quadratic combinations of the fields) . Under these assumptions, the potential is simplified to

$$\begin{aligned} \tilde{V}_{\text{2HDM}} = & m_{11}^2 \, \Phi_1^\dagger \Phi_1 + m_{22}^2 \, \Phi_2^\dagger \Phi_2 - m_{12}^2 \left(\Phi_1^\dagger \Phi_2 + \Phi_2^\dagger \Phi_1 \right) + \frac{\lambda_1}{2} \left(\Phi_1^\dagger \Phi_1 \right)^2 + \frac{\lambda_2}{2} \left(\Phi_2^\dagger \Phi_2 \right)^2 \\ & + \lambda_3 \, \Phi_1^\dagger \Phi_1 \, \Phi_2^\dagger \Phi_2 + \lambda_4 \, \Phi_1^\dagger \Phi_2 \, \Phi_2^\dagger \Phi_1 + \frac{\lambda_5}{2} \left[\left(\Phi_1^\dagger \Phi_2 \right)^2 + \left(\Phi_2^\dagger \Phi_1 \right)^2 \right], \end{aligned} \tag{1.3.7}$$

with only eight real parameters. Some basic constraints on these parameters, e.g. based on the stability of the potential are discussed in e.g. in [84] and [81, p. 91].

It is clear that $\tilde{V}_{\text{2HDM}}$ is invariant under rotations of Φ_a. After suitable rotations, we can assume without loss of generality that the minimization of the potential results in the vacuum expectation values (VEVs)

$$\langle \Phi_1 \rangle_0 = \begin{pmatrix} 0 \\ \dfrac{v_1}{\sqrt{2}} \end{pmatrix}, \quad \langle \Phi_2 \rangle_0 = \begin{pmatrix} 0 \\ \dfrac{v_2}{\sqrt{2}} \end{pmatrix}. \tag{1.3.8}$$

Expanding around the VEVs, we write

$$\Phi_a = \begin{pmatrix} \phi_a^+ \\ \left(v_a + \rho_a + \mathrm{i}\eta_a \right) \big/ \sqrt{2} \end{pmatrix}, \quad a = 1, 2, \tag{1.3.9}$$

where ρ_a, η_a, $\phi_a^+ \pm \overline{\phi_a^+}$ are eight degrees of freedom. We write $\phi_a^- := \overline{\phi_a^+}$. To find the corresponding mass eigenstates, write $\tilde{V}_{\text{2HDM}} = \vec{\Phi}^{\mathrm{T}} \mathcal{M} \vec{\Phi} + \mathcal{O}(\vec{\Phi}^3)$ (where $\vec{\Phi}^{\mathrm{T}} = \begin{pmatrix} \Phi_1 & \Phi_2 \end{pmatrix}$ and $\mathcal{O}(\vec{\Phi}^3)$ are terms proportional to $\Phi_1^m \Phi_2^n$ with $m + n \geq 3$) and diagonalize $\mathcal{M}$. A derivation can be found in [85]. As a result, one finds the neutral bosons, A^0 (photon), G_Z^0, h^0, H^0, and the charged bosons $G_W^\pm$ (W boson) and $H^\pm$. For the charged bosons one finds:

$$\begin{pmatrix} G_W^\pm \\ H^\pm \end{pmatrix} = \begin{pmatrix} \cos\beta & \sin\beta \\ -\sin\beta & \cos\beta \end{pmatrix} \begin{pmatrix} \phi_1^\pm \\ \phi_2^\pm \end{pmatrix}, \tag{1.3.10}$$

where β, the "perhaps single most important parameter in studies of 2HDMs" [81, p. 7] is given by

$$\tan\beta := \frac{v_2}{v_1}. \tag{1.3.11}$$

The overall size of the VEVs is fixed by the relation [81, p. 49]

$$v_1^2 + v_2^2 = \frac{1}{\sqrt{2} G_F} \approx 246\,\text{GeV}, \tag{1.3.12}$$

	Transformation					Coupling		
Type	U_R	D_R	L_R	Φ_1	Φ_2	u_R^i	d_R^i	e_R^i
I	+	+	+	−	+	Φ_2	Φ_2	Φ_2
II	+	−	−	−	+	Φ_2	Φ_1	Φ_1
X	+	+	−	−	+	Φ_2	Φ_2	Φ_1
Y	+	−	+	−	+	Φ_2	Φ_1	Φ_2

Table 1.5: The four types of 2HDMs with natural flavor conservation. Type X and Y have also been referred to as type III and IV, Ia and IIb or "lepton specific" and "flipped". Table adapted from [84, p. 7] and [81, p. 11].

where G_F is the Fermi constant. The Yukawa couplings of the 2HDM are given by [84]

$$\begin{aligned} -\mathcal{L}_Y = \overline{Q}_L \tilde{\Phi}_1 \eta_1^U U_R + \overline{Q}_L \Phi_1 \eta_1^D D_R + \overline{Q}_L \Phi_1 \eta_1^L L_R + \overline{Q}_L \tilde{\Phi}_2 \eta_2^U U_R + \\ + \overline{Q}_L \Phi_2 \eta_2^D D_R + \overline{Q}_L \Phi_2 \eta_2^L L_R, \end{aligned} \tag{1.3.13}$$

where $\tilde{\Phi}_i = \mathrm{i}\sigma_2 \Phi_i^*$ and η_a^F $(F = U, L, D)$ are 3×3 Yukawa matrices.

One of the challenges of 2HDM model building is to avoid tree-level FCNCs, which do not exist in the SM and are experimentally severely constrained. Models with *natural flavor conservation* therefore completely forbid FCNCs. The Paschos-Glashow-Weinberg theorem [86, 87] states that the absence of FCNCs is equivalent to the following three conditions: Fermions of given charge and helicity

1. transform according to the same irreducible $\mathsf{SU}(2)$ representation,
2. correspond to the same T_3 eigenvalue,
3. receive their mass from a single source (for a choice of basis common to all fermions of given charge and helicity).

These conditions only leave four types, each of which can be enforced via a $\mathbb{Z}_2$ symmetry. The four types together with their couplings and symmetries are shown in table 1.5.

Prominently appearing in supersymmetric models, the type II 2HDM has been the most studied model so far. In particular, the Minimal Supersymmetric Standard Model (MSSM) features exactly this structure. However, a 2HDM of this type has been ruled out as a joint explanation of the $R(D)$ and $R(D^*)$ results by a 2012 BaBar analysis [88]. This is demonstrated in figure 1.7: The value of $R(D)$ corresponds to a value of $\tan\beta/m_{H^+} = 0.44\,\mathrm{GeV}^{-1}$, whereas the value of $R(D^*)$ points to a value of $\tan\beta/m_{H^+} = 0.75\,\mathrm{GeV}^{-1}$. This rules out the type II 2HDM as a joint explanation for $R(D^{(*)})$ at confidence level[8] of 99.8% [88].

A different possibility that excludes FCNCs, but still leaves room for additional sources of CP violation (which are excluded by the discrete symmetries describe above) was proposed by Pich and Tuzón [90]. Here, the mass matrices of the Yukawa flavor space coupling matrices of Φ_1 and Φ_2 are required to be proportional to each other (Yukawa alignment), meaning they are simultaneously diagonalizable and hence do not lead to FCNCs. Some phenomenological studies of these Aligned 2HDMs (A2HDMs) have for example been performed in [91].

[8] This confidence level was obtained under the assumption of $m_{H^+} > 10\,\mathrm{GeV}$, however $m_{H^+} \leq 10\,\mathrm{GeV}$ had already been ruled out by $B \longrightarrow X_s\gamma$ measurements in [89].

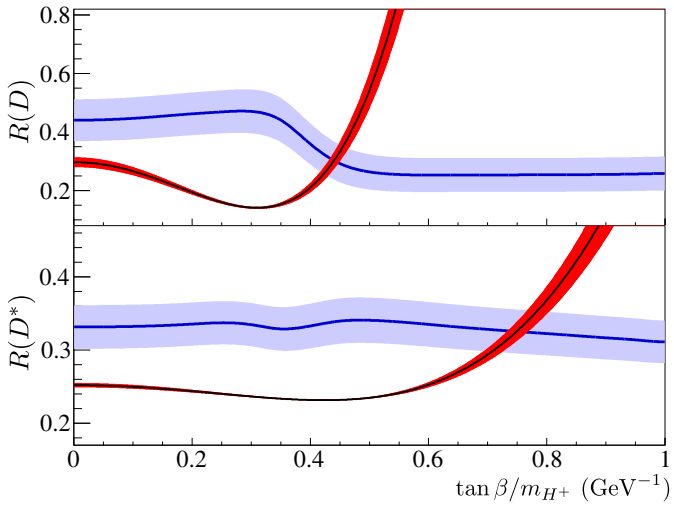


Figure 1.7: $R(D^{(*)})$ measurement of the BaBar collaboration under the assumption of specific values of $\tan\beta$ (dark blue line with light blue band) compared with the theory prediction of a type II 2HDM (dark line, red band). The SM corresponds to $\tan\beta = 0$. [88]

Another possibility to suppress FCNCs is to require the non-diagonal Yukawa couplings to be proportional to the geometric mean of the corresponding fermion masses [92] (type III 2HDM) or to simply set lower bounds on the scalar boson masses to avoid experimental constraints (which however also reduces the phenomenological impact of the model in general).

A 2017 review of models with extended Higgs sectors can be found in [93]. Analyses that concentrate on the impact on $\bar{B} \longrightarrow D^* \tau^- \bar{\nu}_\tau$ are for example [94], [95] and [22].

1.3.3 W' Bosons

W' bosons are hypothetical massive color singlets with electric charge ± 1 and spin 1, that is, they can be considered as "heavy copies of the W boson". The most general coupling of the W' boson under the assumption of Lorentz invariance can be written as [96, p. 1]

$$\mathcal{L} = \frac{g}{\sqrt{2}} \overline{f}_i \gamma_\mu (C^R_{f_i f_j} P_R + C^L_{f_i f_j} P_L) W' f_j + \text{h.c.}, \tag{1.3.14}$$

where $C^{L/R}_{f_i, f_j}$ are matrices parametrizing the couplings to SM fermions, $P_{R/L} = (1 \pm \gamma_5)/2$ and g is the SM coupling constant, such that the SM W coupling corresponds to $C^R = 0$ and $C^L = V_{CKM}$ for quarks and a diagonal matrix for the leptons.

Most W' models are associated with extensions of the SM gauge group. As the W' carries charge, this extension of the SM gauge group has to be non-abelian [78]. Introducing a W' boson via an extension of the SM gauge group will also automatically introduce a Z' boson, a gauge boson with similar properties to the SM Z boson and thereby to additional experimental constraints (such as dilepton searches if the Z' couples to SM leptons) [97].

A popular class of models with W' bosons are left-right symmetric models (LR models), in which the SM gauge group $\mathsf{SU}(3)_C \times \mathsf{SU}(2)_W \times \mathsf{U}(1)_Y$ is extended by another $\mathsf{SU}(2)_R$ group, applying to right-handed particles in the same way that $\mathsf{SU}(2)_L$ applies to left-handed particles. The extended symmetry group is then reduced to the SM gauge group by spontaneous symmetry breaking at low energies, so that the observed chirality asymmetry in the SM is explained as an emergent property at low energies. Explanations for the small mass of left-handed neutrinos via see-saw mechanisms and automatic R parity conservation in supersymmetric extensions [98] further add to the motivation. LR models have been extensively studied since the 1970s [99–102], in particular in the context of the Pati-Salam model $\mathsf{SU}(4)_C \times \mathsf{SU}(2)_L \times \mathsf{SU}(2)_R$ or $^{(\mathsf{SU}(4)_C \times \mathsf{SU}(2)_L \times \mathsf{SU}(2)_R)}/_{\mathbb{Z}_2}$,

one of the first attempts at a grand unified theory. A particularly simple LR model is the minimal LR model (MLRM) with the gauge group $\mathsf{SU}(3) \times \mathsf{SU}(2)_L \times \mathsf{SU}(2)_R \times \mathsf{U}(1)_{L-R}$.

As for the transformation properties of W' under the SM symmetry group, it was shown [80] showed that the only representations of color singlet vector bosons with renormalizable couplings to SM fermions are $(\mathbf{1},\mathbf{1})_1$, $(\mathbf{1},\mathbf{3})_0$ and $(\mathbf{1},\mathbf{2})_{-3/2}$. The representations $(\mathbf{1},\mathbf{2})_{-3/2}$ and $(\mathbf{1},\mathbf{1})_1$ do not have renormalizable couplings to quarks and leptons, respectively, and are therefore not relevant for the B anomalies. Thus, the only transformation property for a W' contributing to the B anomalies is $(\mathbf{1},\mathbf{3})_0$.

In general however, results of ATLAS and CMS tightly constrain color singlet W' models in a way that seems to leave fermionic couplings too weak to explain the B anomalies [22].

1.3.4 Leptoquarks

While leptons and quarks are described as independent fields in the SM, the consistency of the SM as a quantum field theory requires a cancellation of triangle (gauge) anomalies [82, p. 46], leading to deep connections between the quark and lepton sectors. This has been taken as a pointer towards more fundamental theories that include quark-lepton transitions by means of new mediators, *leptoquarks* (LQs), that carry both lepton and baryon numbers.

In fact, LQs occur in many extensions of the SM, in particular in most (attempts at) grand unified theories, such as Pati-Salam $\mathsf{SU}(4)$ [99], Georgi-Glashow $\mathsf{SU}(5)$[9] [104], $\mathsf{SO}(10)$ [105] and $\mathsf{SO}(15)$ [106]. Other models with scalar leptoquarks include extended technicolor (in which LQs appear as bound states of techni-fermions) [107], R parity violating SUSY models and composite models (models in which quarks and leptons are composite particles) [108].

An early classification was performed in 1987 by Bruchmüller, Rückl and Wyler (BRW model) [109]. Under the assumption that

1. LQ couplings are invariant under the SM gauge group $\mathsf{SU}(3)_C \times \mathsf{SU}(2)_W \times \mathsf{U}(1)_Y$,
2. LQ couplings are renormalizable,
3. LQ couplings separately conserve baryon number (B) and lepton number (L),
4. LQ couple *only* to SM fermions and gauge bosons,

only ten representations under $\mathsf{SU}(3)_C \times \mathsf{SU}(2)_W \times \mathsf{U}(1)_Y$ are possible. Different assignments become possible if there are no direct couplings to SM fermions [110], or when the baryon number is not conserved (which however usually leads to rapid proton decay) [111]. Additionally, it is often assumed that LQ couplings to fermions are chiral[10] and that LQs only couple to a single generation of SM fermions[11].

In addition to the notation from the BRW model [109], the notation from [114] (sometimes referred to as "Aachen notation") has gained some popularity. Frequently however, simply the transformation properties under $\mathsf{SU}(3)_C \times \mathsf{SU}(2)_W \times \mathsf{U}(1)_Y$ are used as labels. All three notations are compared in table 1.7. Some examples of how LQs can arise from spontaneous symmetry breaking of extended gauge symmetry groups are shown in table 1.8.

[9] While the original model is no longer considered to be viable due to a range of issues, there are many propositions for extensions that solve these [103].

[10] Chiral couplings mean that a LQ couples *either* to left-handed *or* to right-handed quarks, but not both. Non-chiral couplings lead to four-Fermi interactions $\propto \bar{u}_R d_L \bar{e}_R \nu_L$ that contribute to $\pi \longrightarrow e\nu$, which is helicity suppressed in the SM, resulting in very strong constraints [112].

[11] This property, also called "diagonal" couplings, protects against tree-level flavor changing neutral currents, that are strongly constrained [112]. It should however be noted that strict diagonality is impossible for LQs that couple to left-handed quarks [113].

BRW	Aachen	Repr.	S	F	Q	Allowed Couplings
S_1	S_0	$(\bar{\mathbf{3}},\mathbf{1})_{1/3}$	0	-2	1/3	$\bar{q}^c_{\mathrm{L}}\ell_{\mathrm{L}}$ or $\bar{u}^c_{\mathrm{R}}e_{\mathrm{R}}$
$\tilde{S}_1$	$\tilde{S}_0$	$(\bar{\mathbf{3}},\mathbf{1})_{4/3}$	0	-2	4/3	$\bar{d}^c_{\mathrm{R}}e_{\mathrm{R}}$
S_3	S_1	$(\bar{\mathbf{3}},\mathbf{3})_{1/3}$	0	-2	1/3	$\bar{q}^c_{\mathrm{L}}\ell_{\mathrm{L}}$
V_2	$V_{1/2}$	$(\bar{\mathbf{3}},\mathbf{2})_{5/6}$	1	-2	1/3, 4/3	$\bar{q}^c_{\mathrm{L}}\gamma^\mu e_{\mathrm{R}}$ or $\bar{d}^c_{\mathrm{R}}\gamma^\mu\ell_{\mathrm{L}}$
$\tilde{V}_2$	$\tilde{V}_{1/2}$	$(\bar{\mathbf{3}},\mathbf{2})_{-1/6}$	1	-2	$-2/3$, 1/3	$\bar{u}^c_{\mathrm{R}}\gamma^\mu\ell_{\mathrm{L}}$
R_2	$S_{1/2}$	$(\mathbf{3},\mathbf{2})_{7/6}$	0	0	2/3, 5/3	$\bar{q}_{\mathrm{L}}e_{\mathrm{R}}$ or $\bar{u}_{\mathrm{R}}e_{\mathrm{L}}$
$\tilde{R}_2$	$\tilde{S}_{1/2}$	$(\mathbf{3},\mathbf{2})_{1/6}$	0	0	$-1/3$, 2/3	$\bar{d}_{\mathrm{R}}\ell_{\mathrm{L}}$
U_1	V_0	$(\mathbf{3},\mathbf{1})_{2/3}$	1	0	2/3	$\bar{q}_{\mathrm{L}}\gamma^\mu\ell_{\mathrm{L}}$ or $\bar{d}_{\mathrm{R}}\gamma^\mu e_R$
$\tilde{U}_1$	$\tilde{V}_0$	$(\mathbf{3},\mathbf{1})_{5/3}$	1	0	5/3	$\bar{u}_{\mathrm{R}}\gamma^\mu e_{\mathrm{R}}$
U_3	V_1	$(\mathbf{3},\mathbf{3})_{2/3}$	1	0	$-1/3$, 5/3	$\bar{q}_{\mathrm{L}}\gamma^\mu\ell_{\mathrm{L}}$

Table 1.7: Classification of $\mathsf{SU}(3)_C \times \mathsf{SU}(2)_W \times \mathsf{U}(1)_Y$ invariant LQs in the BRW model [79, 109]. Furthermore, the equivalent in the Aachen notation [114] is presented (note that the mirrored particle-antiparticle convention). Highlighted are all LQs that could contribute to $\bar{B}\longrightarrow D^*\ell^-\bar{\nu}_\ell$. This table uses the hypercharge convention $Q = Y + T^3$. $F = 3B + L$ denotes the fermion number, S the spin and Q the (electric) charge of the LQ.

Leptoquark	Symmetry Breaking Pattern
$(\mathbf{3},\mathbf{1})_{2/3}$, $(\mathbf{3},\mathbf{1})_{5/3}$	$SU(4) \to SU(3) \times U(1)$
$(\mathbf{3},\mathbf{2})_{1/6}$, $(\mathbf{3},\mathbf{2})_{-5/6}$	$SU(5) \to SU(3)_c \times SU(2)_L \times U(1)_Y$
$(\mathbf{3},\mathbf{3})_{2/3}$	$SU(6) \to U(1) \times SU(3) \times (SU(3) \to SU(2))$

Table 1.8: Examples of symmetry breaking patterns giving rise to different types of leptoquarks [80, p. 5].

A 2016 review of leptoquarks with a focus on their collider phenomenology is presented in [115].

Chapter 2

Observables Sensitive to New Physics in $\bar{B} \longrightarrow D^*(\rightarrow D\pi) l^- \bar{\nu}_l$

In this chapter, the differential cross section of the decay $\bar{B} \longrightarrow D^*(\rightarrow D\pi) l^- \bar{\nu}_l$ is parametrized according to different dependencies on the three decay angles and the coupling constants of potential NP contributions. Observables using binned measurements of the cross section are constructed and optimal binnings for such measurements are discussed. Furthermore, a basis of dimension six operators with renormalizable couplings contributing to $\bar{B} \longrightarrow D^* l^- \bar{\nu}_l$ is considered in this framework.

2.1 Introduction

The full differential cross section for the $\bar{B} \longrightarrow D^*(\rightarrow D\pi) l^- \bar{\nu}_l$ decay can be written as (similar to [116])

$$\frac{\mathrm{d}^4\Gamma(\bar{B} \longrightarrow D^*(\rightarrow D\pi) l^- \bar{\nu}_l)}{\mathrm{d}q^2 \, \mathrm{d}\chi \, \mathrm{d}\theta_\ell \, \mathrm{d}\theta_{D^*}} = \frac{9}{32\pi} N_F(q^2) \sum_{a \in A} W_a(q^2, c_i) B_a(\chi, \theta_\ell, \theta_{D^*}), \tag{2.1.1}$$

with 12 linearly independent[1] terms B_a that depend on the three decay angles χ, θ_ℓ and θ_{D^*} (defined in figure 2.1) and q^2 dependent coefficients $W_a(q^2, c_i)$ that also depend on parameters c_i ($c_i \in \mathbb{C}$, $1 \leq i \leq 13$) that parametrize NP contributions. Here, q^2 is the squared four momentum transfer $q^2 = (p_B - p_{D^*})^2$ between the B and the D^*.[2] The four variables χ, θ_ℓ, θ_{D^*} and q^2 are referred to as the *kinematic variables* of this decay channel. Plots of the corresponding single differential branching fractions $\mathrm{d}\Gamma/\,\mathrm{d}q^2$, $\mathrm{d}\Gamma/\,\mathrm{d}\chi$ etc. are shown in figure 2.2.

Following the notation of [116], the summands $W_a B_a$ are split in three categories based on the helicity configuration $\lambda_{D^*}\lambda_W$ of the interactions they describe. Here, λ_{D^*} and λ_W are the helicities of the D^* and the $(\ell\bar{\nu})$ systems, respectively. The configuration $\lambda_{D^*}\lambda_W = 00$ is called "longitudinal" and the corresponding terms are indicated with an upper index '0': $W_1^0 B_1^0$, $W_2^0 B_2^0$ and $W_3^0 B_3^0$. The lower indices 1, 2 and 3 are arbitrary and enumerate terms with different angle dependencies. Similarly, the terms $W_1^T B_1^T$, $W_2^T B_2^T$, $W_3^T B_3^T$, $W_4^T B_4^T$ and $W_5^T B_5^T$ correspond to $\lambda_{D^*}\lambda_W = \pm\pm'$ ("transverse") and the terms W_1^{0T}, W_2^{0T}, W_3^{0T} and W_4^{0T} to $\lambda_{D^*}\lambda_W = 0\pm$ or ± 0 ("mixed").

Note that the amplitudes W_5^T, W_3^{0T} and W_4^{0T} drop out in the CP average (which is being assumed in most experimental measurements of $\bar{B} \longrightarrow D^* l^- \bar{\nu}_l$). In order to stay as general as possible (and to stay sensitive to new sources of CP violation), we do consider the full cross section *without* CP averaging.

[1] When we talk about linearly independent functions $\{f_i \mid 1 \leq i \leq n\}$ in the following, we mean that there exist no $\alpha_i \in \mathbb{R}$, not all zero, such that $\sum_i \alpha_i f_i \equiv 0$.

[2] Sometimes the linearly related quantity $w := \frac{m_B^2 + m_{D^*}^2 - q^2}{2 m_B m_{D^*}}$ is used instead of q^2.

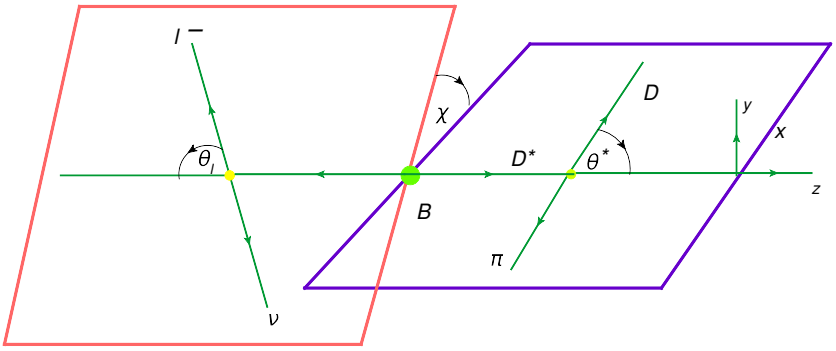


Figure 2.1: Definition of the angle χ and the helicity angles θ_ℓ and θ_{D^*} in the decay $\bar{B}\longrightarrow D^*(\to D\pi)\tau^-\bar{\nu}_\tau$: The red plane is spanned by the momenta of the leptons ℓ and $\bar{\nu}_\ell$, with θ_ℓ the angle between the opposite D^* momentum and the ℓ^- momentum (as seen from the $\ell\bar{\nu}_\ell$ rest frame). The blue plane is spanned by the π and D momenta with θ_{D^*} the angle between the D and the π (as seen from the D^* rest frame). Finally, χ is the angle between both decay planes as seen from the B rest frame. Graphic from [116, p. 6].

$$
\begin{aligned}
B_1^0 &= \sin(\theta_{D^*})\cos^2(\theta_{D^*})\sin(\theta_\ell)\\
B_2^0 &= \sin(\theta_{D^*})\cos^2(\theta_{D^*})\sin^3(\theta_\ell)\\
B_3^0 &= \sin(\theta_{D^*})\cos^2(\theta_{D^*})\sin(\theta_\ell)\cos(\theta_\ell)
\end{aligned}
\qquad
\begin{aligned}
B_1^T &= \sin^3(\theta_{D^*})\sin(\theta_\ell)\\
B_2^T &= \sin^3(\theta_{D^*})\sin^3(\theta_\ell)\\
B_3^T &= \sin^3(\theta_{D^*})\sin(\theta_\ell)\cos(\theta_\ell)\\
B_4^T &= \cos(2\chi)\sin^3(\theta_{D^*})\sin^3(\theta_\ell)\\
B_5^T &= \sin(2\chi)\sin^3(\theta_{D^*})\sin^3(\theta_\ell)
\end{aligned}
$$

$$
\begin{aligned}
B_1^{0T} &= \cos(\chi)\sin(\theta_{D^*})\sin(2\theta_{D^*})\sin(\theta_\ell)\sin(2\theta_\ell)\\
B_2^{0T} &= \cos(\chi)\sin(\theta_{D^*})\sin(2\theta_{D^*})\sin^2(\theta_\ell)\\
B_3^{0T} &= \sin(\chi)\sin(\theta_{D^*})\sin(2\theta_{D^*})\sin^2(\theta_\ell)\\
B_4^{0T} &= \sin(\chi)\sin(\theta_{D^*})\sin(2\theta_{D^*})\sin(\theta_\ell)\sin(2\theta_\ell)
\end{aligned}
$$

Table 2.1: The functions $B_a(\chi,\theta_\ell,\theta_{D^*})$

To use sum notations, we sometimes join the lower and upper index, e.g. write W_a, $a = 10T$ instead of W_1^{0T}. Thus, we write for the limits of the sum in (2.1.1):

$$A := \{10, 20, 30, 1T, 2T, 3T, 4T, 5T, 10T, 20T, 30T, 40T\}. \tag{2.1.2}$$

The full implementation of the differential cross section follows [116] (with some adjustment to the dissection into summands) and can be found in detail in appendix A. The explicit angle dependencies B_a are shown in table 2.1.

As we will concentrate on the angular distribution, we mostly consider equation (2.1.1) integrated over q^2:

$$\frac{\mathrm{d}^3\Gamma}{\mathrm{d}\chi\,\mathrm{d}\theta_\ell\,\mathrm{d}\theta_{D^*}} = \sum_{a\in A}\overline{W}_a(c_i)B_a(\chi,\theta_\ell,\theta_{D^*}), \tag{2.1.3}$$

where we defined

$$\overline{W}_a(c_i) := \frac{9}{32\pi}\int_{q^2_{\min}}^{q^2_{\max}}\mathrm{d}q^2\, N_F(q^2)W_a(q^2,c_i), \tag{2.1.4}$$

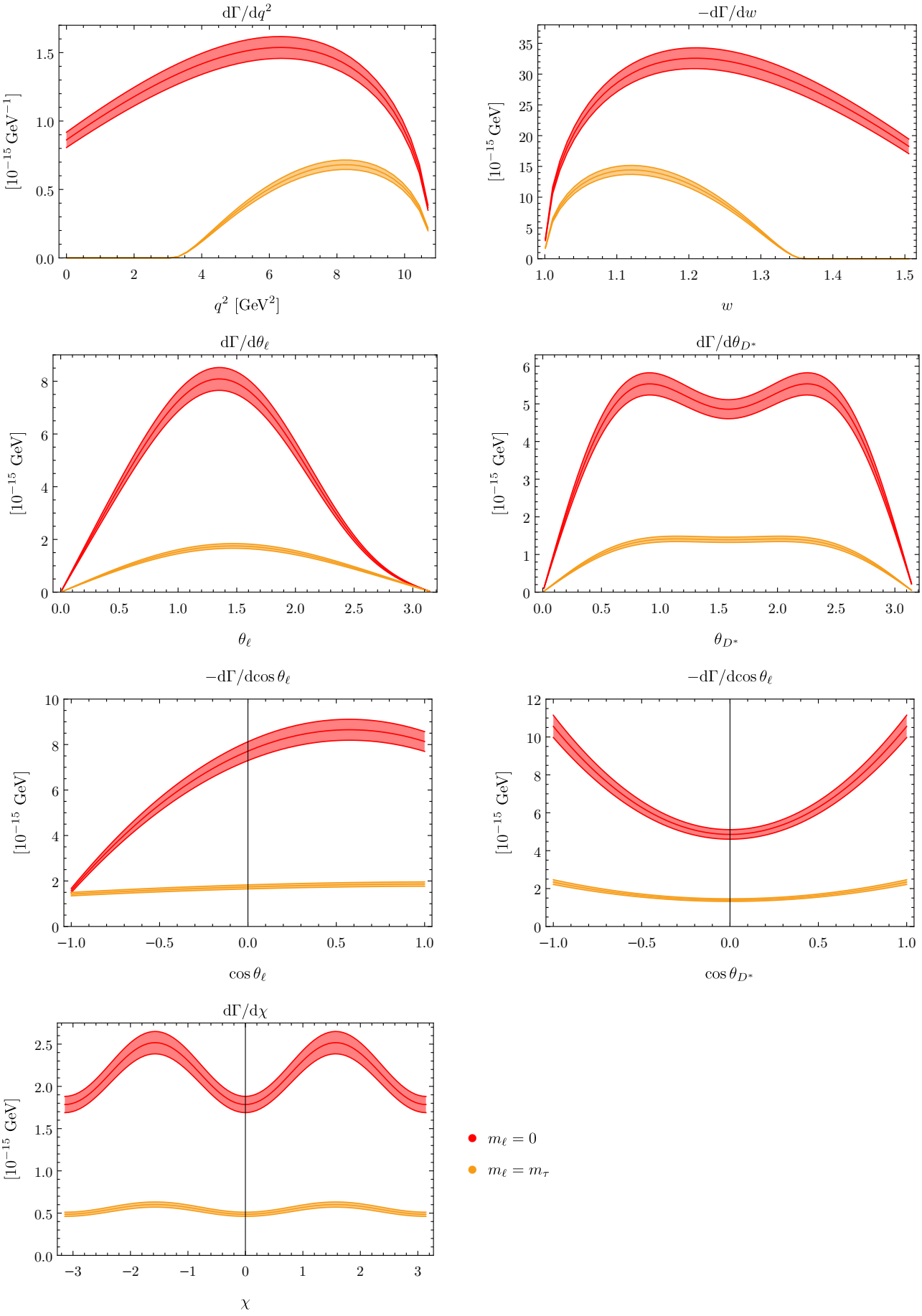


Figure 2.2: Single differential cross sections in $q^2, w, \theta_\ell, \theta_{D^*}, \cos\theta_\ell, \cos\theta_{D^*}$ and χ for the SM (red: light leptons, orange: tauons). The error bands show the pointwise $\pm 1\sigma$ systematic errors (see appendix A.4). Note that an additional minus sign is introduced for the differential cross sections in $\cos\theta_\ell$, $\cos\theta_{D^*}$ and w in order to show positive values for an easier comparison with histograms. This is due to the flipping of integral limits (which would require to integrate from "right to left"), as the cosine and $w(q^2)$ decrease with increasing angles and q^2, respectively.

with $q^2_{\text{min}} = m_l^2$ and $q^2_{\text{max}} = (m_B - m_{D^*})^2$ as limits.[3] The values of $\overline{W}_a$ for the SM are visualized in figure 2.3.

An experimental analysis measuring the differential cross section with respect to several of the kinematic variables will publish their results *binned* (i.e. integrated over subsets of the phase space) in bins $\Gamma(i)$. Thus, the goal is to use these values $\Gamma(i)$ to gain insight on NP contributions.

We proceed in three steps:

1. In section 2.2, we consider equation (2.1.3) binned in N bins, such that it takes the form of a set of N linear equations with the $\overline{W}_a$ as unknown variables. For a suitable binning, we can then find weights $\omega_a(i)$, such that

$$\overline{W}_a = \sum_i \omega_a(i)\Gamma(i), \tag{2.1.5}$$

   giving rise to an observable measuring $\overline{W}_a$, which is simply a weighted sum of the measured bin contents $\Gamma^{\text{exp}}(i)$:

$$\mathcal{O}_a\colon\ \Gamma^{\text{exp}} \longmapsto \sum_{i=1}^{N} \omega_a(i)\,\Gamma^{\text{exp}}(i). \tag{2.1.6}$$

   In section 2.3 we construct such observables for different assumptions on the bins and weights and discuss their performance. This will result in recommendations for the number and spacing of bins in the three kinematic variables in future analyses, summarized in section 2.3.5.

2. The $\overline{W}_a(c_i)$ contain the maximal information that can be inferred from the angularly binned differential cross section. The next step is to inspect their dependency on NP coupling constants. In section 2.4, we make this dependency explicit by parametrizing $\overline{W}_a(c_i) = \sum_i c_i \overline{W}_a^{(i)}$, where the c_i are linearly independent quadratic combinations of the coupling constants g_A, g_V, g_S, g_P, T_L and the $\overline{W}_a^{(i)}$ are NP independent and considered to be constant (though they of course still depend on masses, form factor parameters etc).

3. Considering general NP without any additional assumptions or constraints, it is in general difficult to make concrete statements about the NP coupling constants given the values of $\overline{W}_a(c_i)$. In section 2.5, we therefore consider a base of all dimension six operators with renormalizable couplings relevant to the $\bar{B}\longrightarrow D^* l^- \bar{\nu}_l$ anomaly. Each of these specific operators introduces relationships between the coupling constants that reduce the dimensionality of the problem. This allows us to more concretely consider the discriminatory power of potential new observables based on the strategy described here.

The idea to consider such observables in the context of the B anomalies has been discussed to some extent in [117] and a particularly simple angular observable, the forward-backward asymmetry

$$A_\theta(q^2) := \left[\int_{-1}^{0} \mathrm{d}\cos\theta_\ell - \int_{0}^{1} \mathrm{d}\cos\theta_\ell\right] \frac{\mathrm{d}\Gamma}{\mathrm{d}q^2\,\mathrm{d}\cos\theta_\ell} \quad \text{and} \quad \overline{A}_\theta := \int_{q^2_{\text{min}}}^{q^2_{\text{max}}} \mathrm{d}q^2 A_\theta(q^2) \tag{2.1.7}$$

has gained widespread use and has been considered for a wide range of different NP models. However, so far no attempt has been made to formalize the construction of such observables and to systematically investigate their performance and binning requirements, which is the aim of this chapter. A more detailed comparison with literature results can be found in section 2.6.

[3] These limits follow out of kinematic constraints: As q^2 is Lorentz invariant, it can be evaluated in the B rest frame as $q^2 = (p_B - p_{D^*})^2 = m_B^2 + m_{D^*}^2 - 2(E_B E_{D^*} - 0\cdot\vec{p}_{D^*}) \leq m_B^2 + m_{D^*}^2 - 2m_B m_{D^*} = (m_B - m_{D^*})^2$. On the other hand, we have $q^2 = (p_B - p_{D^*})^2 = (p_\ell + p_\nu)^2 = m_\ell^2 + 2(E_\ell E_\nu - \vec{p}_\ell\vec{p}_\nu) \geq m_l^2 + 2(E_\ell E_\nu - |\vec{p}_\ell||\vec{p}_\nu|) \geq m_\ell^2$, because $E \geq |\vec{p}|$.

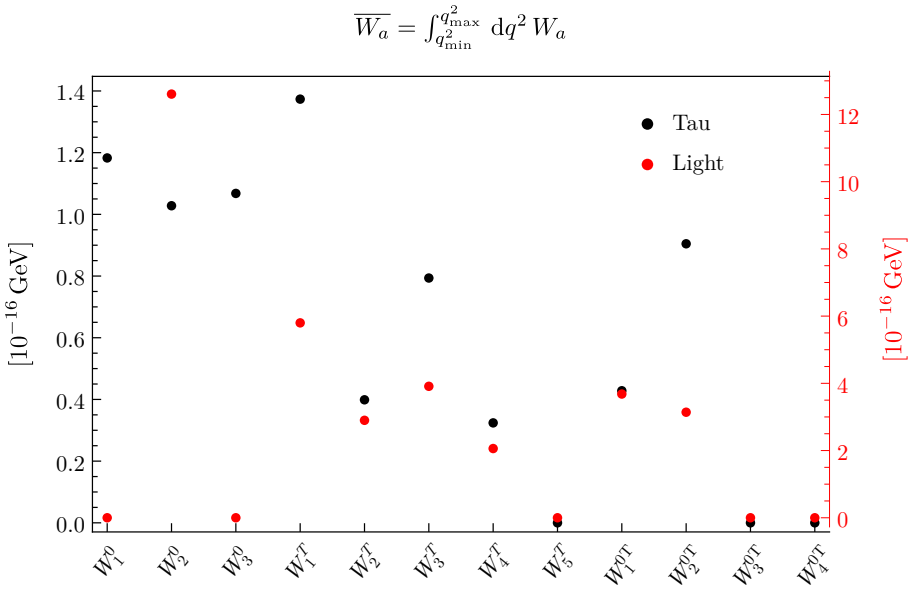


Figure 2.3: Integrated values $\overline{W}_a$ for the SM.

Constructing observables $\mathcal{O}_a$ that directly map onto measurements of $\overline{W}_a$ is not the only way to determine $\overline{W}_a$. A conceptually much simpler approach to extract all $\overline{W}_a$ at once is given by minimizations of expressions of the discrepancy between the measured binned data and the parametrized expectation for each bin.

However, such fits usually constitute "all-or-nothing" approaches: Either the binning is sufficiently fine to extract every $\overline{W}_a$ or the minimization will not arrive at unique values for any $\overline{W}_a$ (because there are more degrees of freedom than constraints). The minimal value of the minimized function Δ_f gives a benchmark for the overall agreement of data and the theoretically expected form of the distribution.

In contrast, the approach of angular observables considers each $\overline{W}_a$ (or combinations thereof) separately, arriving at different conditions for the binning to extract each $\overline{W}_a$ and also at different associated experimental errors. This allows for a more tailored analysis, even with a low number of bins and generally sparse data.

2.2 Our Approach

Let us assume that we have access to the q^2-integrated differential cross section in the three angles, given in bins

$$U(i) \subseteq [-\pi, \pi] \times [0, \pi] \times [0, \pi], \quad 1 \leq i \leq N, \tag{2.2.1}$$

Define the binned cross section as

$$\Gamma(i) := \int_{U(i)} \mathrm{d}\chi\, \mathrm{d}\theta_\ell\, \mathrm{d}\theta_{D^*} \int_{q^2_{\min}}^{q^2_{\max}} \mathrm{d}q^2 \frac{\mathrm{d}^4\Gamma}{\mathrm{d}q^2\, \mathrm{d}\chi\, \mathrm{d}\theta_\ell\, \mathrm{d}\theta_{D^*}}. \tag{2.2.2}$$

and furthermore introduce the notation

$$\overline{B}_a(i) := \int_{U(i)} \mathrm{d}\chi\, \mathrm{d}\theta_\ell\, \mathrm{d}\theta_{D^*} B_a(\chi, \theta_\ell, \theta_{D^*}), \tag{2.2.3}$$

the binned angle dependencies. With this notation we have by virtue of (2.1.1):

$$\Gamma(i) = \sum_{a \in A} \overline{B}_a(i)\, \overline{W}_a, \tag{2.2.4}$$

which is a matrix equation that can be solved for the $\overline{W}_a$ if and only if

$$\operatorname{rank}\left(\overline{B}_a(i)_{a\in A, 1\leq i\leq N}\right) \overset{!}{=} \#A = 12. \tag{2.2.5}$$

In particular, this gives us the constraint $N \geq 12$ for the number of required bins. Due to the linear character of (2.2.4), a solution will always have the form

$$\overline{W}_a = \sum_{i=1}^{N} \omega_a(i)\, \Gamma(i), \tag{2.2.6}$$

i.e. the q^2-integrated amplitudes $\overline{W}_a$ are given as a simple weighted sum over the bin contents of the cross section, giving rise to the observables $\mathcal{O}_a$ (cf. (2.1.6)). The numbers $\omega_a(i)$ are called *weights* and are elements of generalized inverses of the matrix $(\overline{B}_a(i))$.

The problem of finding appropriate bins, weights and the parametrization of their degrees of freedom is discussed from a slightly more mathematical standpoint in appendix B.

In the following, we are interested in the following three questions:

1. How many bins do we need to construct $\mathcal{O}_a$? How many to construct all 12 $\mathcal{O}_a$?
2. Given a fixed set of bins, how can we choose weights such that we obtain the lowest expected uncertainty on $\mathcal{O}_a$?
3. How can we find the most "beautiful" weights? That is, how can we construct conceptually simple observables $\mathcal{O}_a$?

Answers to these questions will be given under several different assumptions for the bins and for the weights in section 2.3.

2.2.1 Expected uncertainty on the observables

To assess the usefulness of an observable $\mathcal{O}$ for experimental purposes, we need to estimate the level of expected uncertainty.

Let us consider a general expression for an observable $\mathcal{O}_{\vec{\alpha}}$,

$$\mathcal{O}_{\vec{\alpha}}\colon\ \Gamma^{\text{exp}} \longmapsto \sum_{i=1}^{N} \omega(i)\, \Gamma^{\text{exp}}(i), \tag{2.2.7}$$

which maps the results of a binned measurement $\Gamma^{\text{exp}}(i)$ to a measurement of a linear combination $\sum_a \alpha_a \overline{W}_a$ (a slight generalization from our previous definition of $\mathcal{O}_a$, which now corresponds to $\alpha_{a'} = \delta_{aa'}$).

In the following we take the theoretical SM expectation $\Gamma^{\text{theo}}_{\text{SM}}(i)$ as an estimate for the expected value of $\Gamma^{\text{exp}}(i)$. Let $N_{B\bar{B}}$ be the number of generated $B\bar{B}$ events in the experiment and assume that there is a flat reconstruction efficiency ϵ in such a way, that the bin yields are given as

$$N^{\text{exp}}(i) = N_{B\bar{B}} \epsilon \hbar^{-1} \tau_B \Gamma^{\text{exp}}(i), \tag{2.2.8}$$

where τ_B is the mean life time of the B meson[4] and $\hbar = 6.58 \times 10^{-25}\,\text{GeV}\cdot\text{s}$ [5] is the reduced Planck constant. For more information about this, see section 3.9. We write $N_{B\bar{B}} \epsilon \hbar^{-1} \tau_B =: A$.

Let us now assume that the event yield $N^{\text{exp}}(i)$ is Poisson distributed around its true value for repeated measurements, such that

$$A^2 \operatorname{Var}(\Gamma^{\text{exp}}(i)) = \operatorname{Var}(N^{\text{exp}}(i)) = E(N^{\text{exp}}(i)) = A \cdot \Gamma^{\text{theo}}_{\text{SM}}(i). \tag{2.2.9}$$

[4] The PDG averages are $\tau_{B^\pm} = (1.638 \pm 0.004)\,\text{ps}$ and $\tau_{B^0} = (1.520 \pm 0.004)\,\text{ps}$ [5].

Based on this, we make the assumption $\mathrm{Var}(\Gamma^{\mathrm{exp}}(i) = A^{-1}\Gamma^{\mathrm{theo}}_{\mathrm{SM}}(i)$. Since we are only interested in the effects of different binnings and weights, we set $A = 1$. Since $\mathrm{Var}(aX + bY) = a^2\,\mathrm{Var}(X) + b^2\,\mathrm{Var}(Y)$ for $a, b \in \mathbb{R}$ and random variables X, Y, our estimate for the variance of $\mathcal{O}_a$ thus ends up being

$$\begin{aligned}\mathrm{Var}(\mathcal{O}_{\vec{\alpha}}) &= \sum_{i=1}^{N} \omega(i)^2\,\mathrm{Var}(\Gamma^{\mathrm{exp}}(i)) = \sum_{i=1}^{N} \omega(i)^2\,\Gamma^{\mathrm{theo}}_{\mathrm{SM}}(i) = \\ &= \sum_{i=1}^{N} \omega(i)^2 \sum_{a' \in A} \overline{B}_{a'}(i)\overline{W}^{\mathrm{SM}}_{a'}.\end{aligned} \tag{2.2.10}$$

Finding the optimally performing weights then corresponds to the following minimization:

$$\min\left\{ \sum_{i=1}^{N} \omega(i)^2 \sum_{a' \in A} \overline{B}_{a'}(i)\overline{W}^{\mathrm{SM}}_{a'} \;\middle|\; \omega(i) \in \mathbb{R},\ \sum_{i=1}^{N} \omega(i) \sum_{a' \in A} \overline{B}_{a'}(i)\,\hat{e}_{a'} = \vec{\alpha} \right\}, \tag{2.2.11}$$

In practice, this minimization is done by parametrizing all solutions for the weights $\omega^{(\tau)}(i)$ (explained in appendix B) and then simply minimizing the expression for the variance with respect to the degrees of freedom τ.

For an even more thorough investigation of *optimal* observables, especially when considering fine binnings, it would be necessary to test the stability of the performance of such optimal observables under variation of the expected event yield per bin. That is, rather than using just the standard model expectation, the figure of merit should be calculated from the variance for a distribution of possible bin yields varied around the standard model expectation.

At the same time it should be stressed that the assumption of the SM distributions $\Gamma^{\mathrm{theo}}_{\mathrm{SM}}(i)$ (or in fact any kind of numbers for $\Gamma(i)$) for the optimization of the weights does in principle not introduce any kind of bias, because the expected value of $\mathcal{O}_{\vec{\alpha}}$ remains unchanged when varying the degrees of freedom of the weights.

Intuitively one can understand how functions that are hard to distinguish from each other lead to observables with high associated uncertainty as calculated via equation (2.2.10): If two functions B_a and $B_{b\neq a}$ are very similar, it also means that $\overline{B}_a(i) \approx \overline{B}_b(i)$, i.e. the matrix $(\overline{B}_a(i))_{ai}$ has two very similar (almost collinear) rows, that is has a very small determinant and therefore (right)inverses with matrix elements of high absolute value. Since the weights are elements of (right)inverses of this matrix, this means that the squared weights in the calculation of the variance will lead to a high uncertainty of the corresponding observable by equation (2.2.10).

Note that it is impossible to calculate all *relative* errors $\sqrt{\mathrm{Var}(\mathcal{O}_{\vec{\alpha}})}/E(\mathcal{O}_{\vec{\alpha}})$, because some of the observables will have vanishing SM expectation value. Nonetheless, we can use $\sqrt{\mathrm{Var}(\mathcal{O}_{\vec{\alpha}})}$ to compare the performances of differently constructed observables extracting the same feature, e.g. comparing $\sqrt{\mathrm{Var}(\mathcal{O}_a)}$ for $\mathcal{O}_a$ constructed with respect to different binnings.

To put the numbers into perspective, it can be instructive to apply equation (2.2.10) to the measurement of the total cross section Γ, which corresponds to $\omega(i) \equiv 1$ for any binning and gives $\sqrt{\mathrm{Var}(\Gamma)} = 5.85 \times 10^{-8}\,\mathrm{GeV}$ $(\ell = \tau)$ and $1.16 \times 10^{-7}\,\mathrm{GeV}$ $(\ell = e/\mu)$. As $\Gamma = \frac{8}{9}\pi(3\overline{W}^0_1 + 6\overline{W}^T_1 + 2\overline{W}^0_2 + 4\overline{W}^T_2)$, we can compare

$$\frac{8}{9}\pi\sqrt{9\,\mathrm{Var}(\mathcal{O}^0_1) + 36\,\mathrm{Var}(\mathcal{O}^T_1) + 4\,\mathrm{Var}(\mathcal{O}^0_2) + 16\,\mathrm{Var}(\mathcal{O}^T_2)} \tag{2.2.12}$$

to this value.

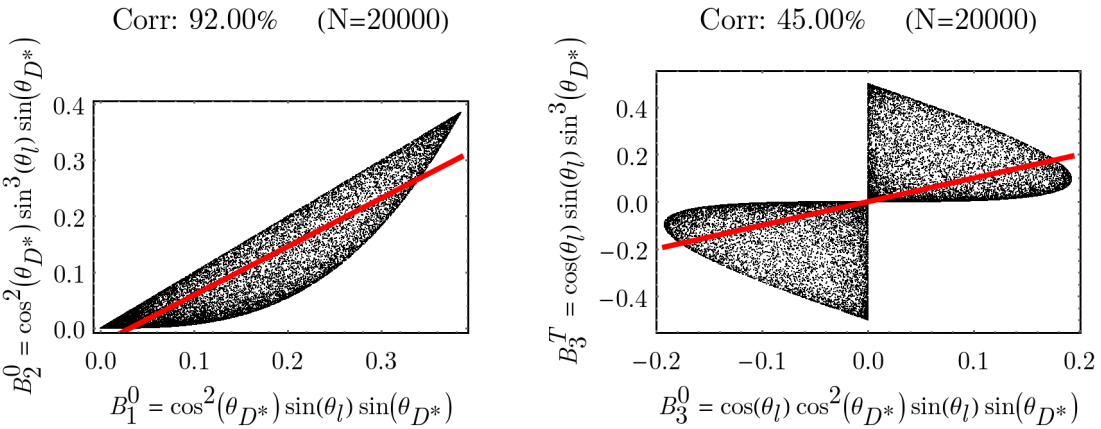


Figure 2.4: Scatter plots and correlation between two different angle dependencies B_a. The plot for B_1^0 vs B_2^0 is almost identical to the one for B_1^T vs B_2^T.

2.2.2 Expected correlation of the observables

Of equal importance is the correlation between two observables. Adding the observable

$$\mathcal{O}_{\vec{\alpha}'}: \ \Gamma^{\text{exp}} \longmapsto \sum_{i=1}^{N} \omega'(i)\, \Gamma^{\text{exp}}(i) \tag{2.2.13}$$

and using the bilinearity of the covariance, we get:

$$\text{Cov}(\mathcal{O}_{\vec{\alpha}}, \mathcal{O}_{\vec{\alpha}'}) = \sum_i \sum_{i'} \omega(i)\omega(i')\, \text{Cov}(\Gamma^{\text{exp}}(i), \Gamma^{\text{exp}}(i')). \tag{2.2.14}$$

Assuming that the measurements of different bins are strictly uncorrelated (which will in fact not be true experimentally, cf. section 3.7), and making similar assumptions about the variance as above, we get

$$\begin{aligned} \text{Cov}(\Gamma^{\text{exp}}(i), \Gamma^{\text{exp}}(i')) &= \underbrace{\text{Corr}(\Gamma^{\text{exp}}(i), \Gamma^{\text{exp}}(i'))}_{\delta_{ii'}} \sqrt{\text{Var}(\Gamma^{\text{exp}}(i))} \sqrt{\text{Var}(\Gamma^{\text{exp}}(i'))} = \\ &= \sum_i \omega(i)\omega'(i) \Gamma^{\text{theo}}_{\text{SM}}(i), \end{aligned} \tag{2.2.15}$$

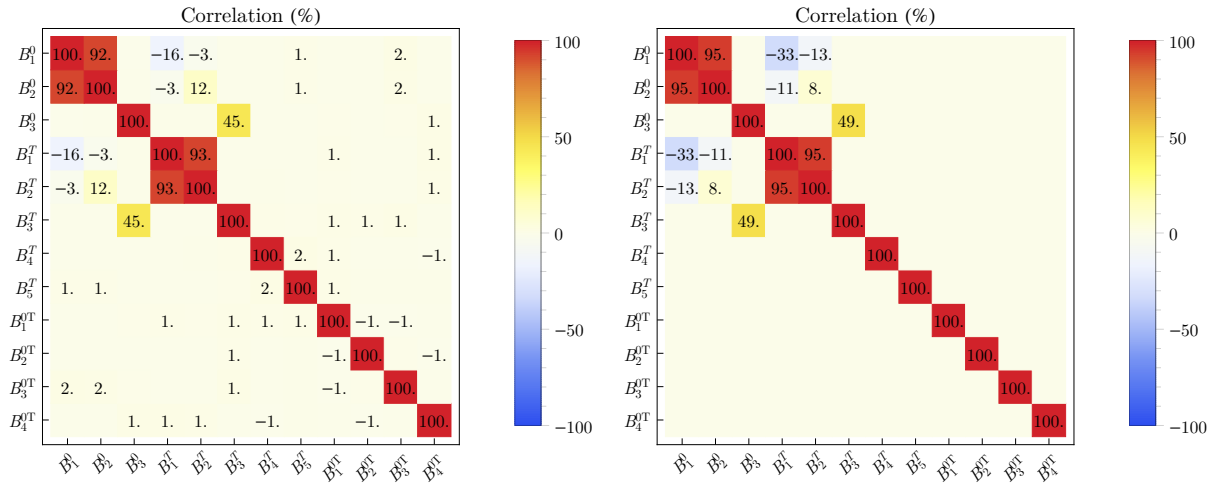
and thus for the correlation

$$\text{Corr}(\mathcal{O}_{\vec{\alpha}}, \mathcal{O}_{\vec{\alpha}'}) = \frac{\sum_{i=1}^{N} \omega(i)\omega'(i) \Gamma^{\text{theo}}_{\text{SM}}(i)}{\sqrt{\sum_{i=1}^{N} \omega(i)^2\, \Gamma^{\text{theo}}_{\text{SM}}(i)} \sqrt{\sum_{i'=1}^{N} \omega'(i')^2\, \Gamma^{\text{theo}}_{\text{SM}}(i')}}. \tag{2.2.16}$$

For an ideal set of observables, we would prefer low correlation between the observables. In principle, even a vanishing correlation can be obtained by considering mutually exclusive parts of the dataset, i.e. by ensuring, that for any two observables $\mathcal{O}_{\vec{\alpha}}$ and $\mathcal{O}_{\vec{\alpha}'}$ as above, $\omega(i)$ and $\omega(i')$ are non-zero on mutually exclusive subsets of the phase space. However, this in turn is likely to increase the variance for both variables, as lower amounts of data can be used per observable.

Considering the observables $\mathcal{O}_a$ with $\mathcal{O}_a(\Gamma^{\text{theo}}) = \overline{W}_a \Longleftrightarrow \mathcal{O}_a(\overline{B}_{a'}) = \delta_{aa'}$ and assuming that they are optimized for low variance independently from each other, it is also instructive to look into the correlations of the binned or unbinned angle dependencies $\overline{B}_a$ or B_a, respectively.

This is not only interesting for later comparison with the correlation matrices of the observables, but it should also be noted that similar angle dependencies (which are often connected to high correlation) will lead to high weights and thus a high variance as well (as discussed before).



(a) Unbinned

(b) Binned in $5 \times 5 \times 3$ uniform bins

Figure 2.5: Correlation between angle dependencies calculated with a uniform sample of 50,000 points.

For the unbinned B_a, one can consider a sample of uniformly randomly distributed points of the phase space $[-\pi, \pi] \times [0, \pi] \times [0, \pi]$ and evaluate two functions B_a and $B_{a'}$ on all of these points. The result can be illustrated in scatter plots as in figure 2.4. It is however important to note that taking the correlation coefficient of two variables as a measure of their independence can be misleading. This caveat is particularly important here, where symmetries in the trigonometric functions can lead to the set $\{(B_a, B_b)\}$ being symmetric around the $(x, 0)$-axis, causing the correlation to vanish.

For the binned values $\overline{B}_a$, a binning $U(i)$ is fixed and the tuples $(\overline{B}_a(i), \overline{B}_a(i'))$ are calculated and then used to determine the correlation coefficients. For a fine binning (e.g. with $10 \times 10 \times 10$ bins), the binned correlation coefficients will converge to the unbinned correlation coefficients. Figure 2.5 gives examples of a matrix of binned and unbinned correlation coefficients.

2.3 Binnings and Weights

In this section, observables $\mathcal{O}_a$ will be constructed based on four different assumptions on the bins $U(i)$ and weights $\omega(i)$. Subsection 2.3.5 compares the results.

If not stated otherwise, all sensitivity optimizations have been carried out for the $l = \tau$ case.

2.3.1 General subsets as bins

Based on the results of appendix B, we know that there exists a binning in 12 bins $U(1), \ldots, U(12) \subset [-\pi, \pi] \times [0, \pi] \times [0, \pi]$, such that we can distinguish between all B_a, i.e. extract all $\overline{W}_a$ using the 12 data points $\Gamma(1), \ldots, \Gamma(12)$.

As a simple proof of concept (that however only uses $1/12^2$ of the data), one can for example consider[5]

$$U_i := \left[-\pi + \frac{i-1}{6}\,\pi, -\pi + \frac{i}{6}\,\pi\right] \times \left[\frac{i-1}{12}\,\pi, \frac{i}{12}\,\pi\right] \times \begin{cases} \left[\frac{i-1}{12}\,\pi, \frac{i}{12}\,\pi\right] & 1 \leq i \leq 10 \\ \left[\frac{11}{12}\,\pi, \frac{12}{12}\,\pi\right] & i = 11 \\ \left[\frac{10}{12}\,\pi, \frac{11}{12}\,\pi\right], & i = 12. \end{cases} \tag{2.3.1}$$

[5] In fact it is even possible to consider $[\frac{i-1}{12}\pi, \frac{i}{12}\pi]$ for the last bin, but the resulting matrix will be so close to being singular, that the non-singularity is hard to verify numerically.

and verify that the corresponding 12×12 matrix $(\overline{B}_a(i))_{a,i}$ indeed has full rank.

Subsequent sections will all place additional constraints on the binning and the weights, thus (unpractical as it may be) this general scenario will in principle yield the best performing observables for a fixed number of bins, N, and could thus serve as a benchmark to test further simplifications against.

While the binning in full generality consists of arbitrary subsets of the phase space, we only consider a simplified approach: First, numbers $n_\chi, n_{\theta_\ell}, n_{\theta_{D^*}} \geq 2$ (called the resolution) are fixed and $n_\chi - 1$ points $\chi^i \in [-\pi, \pi]$, $n_{\theta_\ell} - 1$ points $\theta_\ell^j \in [0, \pi]$ and $n_{\theta_{D^*}} - 1$ points $\theta_{D^*}^k \in [0, \pi]$ are selected from a uniform distribution and then sorted such that $\chi^i \leq \chi^{i+1}$, $\theta_\ell^i \leq \theta_\ell^{i+1}$ and $\theta_{D^*}^i \leq \theta_{D^*}^{i+1}$. Note that the distribution of these edge points is no longer uniform after the sorting. Additionally, we define $\chi^0 = -\pi$, $\theta_\ell^0 = \theta_{D^*}^0 = 0$ and $\chi^{n_\chi} = \theta_\ell^{n_{\theta_\ell}} = \theta_{D^*}^{n_{\theta_{D^*}}} = \pi$. From these edge points, we define the *cubes* $[\chi^i, \chi^{i+1}] \times [\theta_l^j, \theta_l^{j+1}] \times [\theta_{D^*}^k, \theta_{D^*}^{k+1}]$. Finally, each of the cubes is randomly assigned to one of the N bins.

As the computing effort scales directly with the resolution $n_\chi n_{\theta_\ell} n_{\theta_{D^*}}$, a high resolution has to be traded for a low sample size (for a fixed number of bins), even though a higher sample would be required to account for the number of possibilities to assign each cube. This was also the reason why the values χ^i, $\theta_{D^*} {}_l^j$ and $\theta_{D^*}^k$ were randomly chosen instead of in an equidistant fashion, because it allows testing for a more inhomogeneous sample of binnings for coarse resolutions.

As for setting a benchmark, a binning with 27 bins was considered with a $5 \times 5 \times 5$ resolution (figure 2.7b), however the best value for the obtained average expected error $\frac{1}{12} \sum_a \sqrt{\mathrm{Var}(\mathcal{O}_a)} = 1.76 \times 10^{-8}$ is still above the results later obtained in more constrained scenarios (as the amount of generated random binnings is nowhere near sufficient to sample the large parameter space).

Furthermore, the scenario of 12 bins was considered. Note that in this case, no degrees of freedom in the weights remain, thus no optimization of the sensitivity following section 2.2.1 can be performed. Nonetheless, the best value of the average expected error on the measurement of the $\mathcal{O}_a$, $\frac{1}{12} \sum_a \sqrt{\mathrm{Var}(\mathcal{O}_a)} = 2.25 \times 10^{-8}$ is in the same order of magnitude as the results with 27 bins. Samples for $3 \times 3 \times 3$ resolution were tested both with random cube edge points (figure 2.6a) and with equidistantly spaced cube edge points ("uniform resolution", figure 2.6b). Furthermore, samples with an increased resolution of $5 \times 5 \times 5$ were considered (figure 2.7a). The expected errors on the observables for the uniform $3 \times 3 \times 3$ resolution are significantly lower than for the other two cases. However, this shows not much more than that – at such low statistics compared to the dimensionality of the configuration space – the assignment of the cubes to the bins is of much more importance than the number and spacing of the cubes themselves.

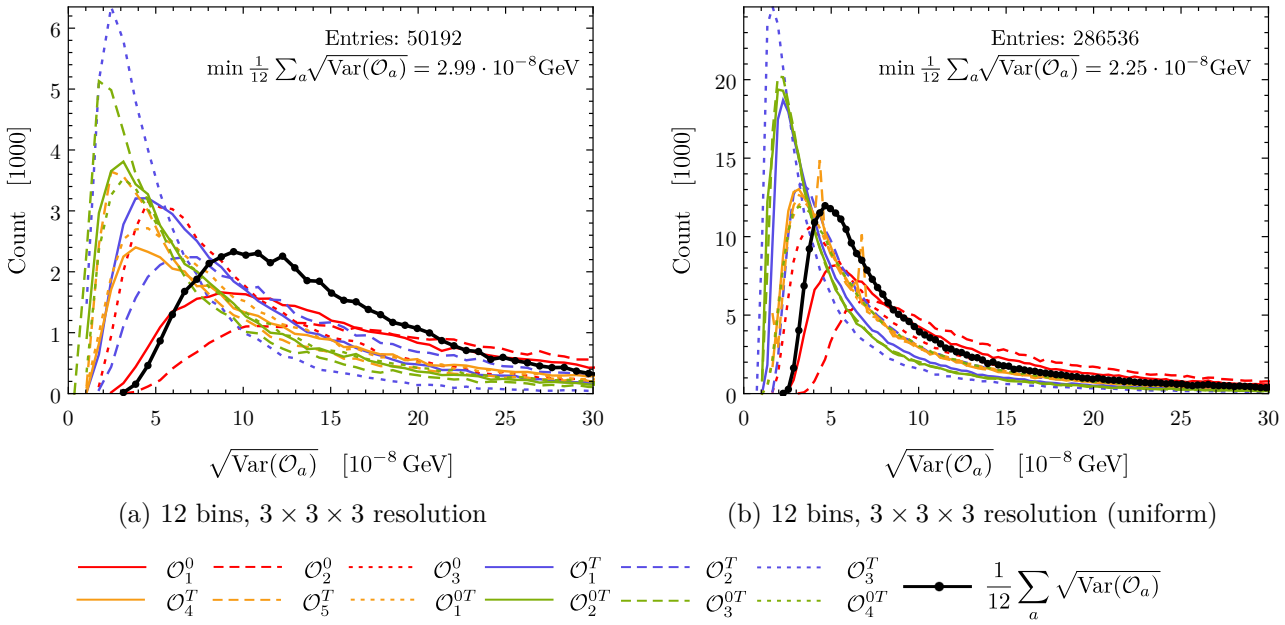
It should also be noted that computing results for different cube edge points is much more costly computation power wise than testing the result of different cube to bin assignments (because the former requires a reintegration of the angle functions). This fact should be taken into account in future studies by testing a larger number of cube to bin assignments per fixed cube edge points.

In general, getting further significantly results would require some insight (such as a well-motivated symmetry or a rough figure of merit) to consider a promising lower dimensional subspace of the configuration space. However, this seems to be particularly hard for the assignment of the cubes to the bins, as this property is hard to condense into a set of real parameters, because there is no simple concept of "neighboring" configurations that could be used for classification.

2.3.2 Product bins

In the following, we consider the more practically relevant scenario of bins in (Cartesian-) product form, i.e.

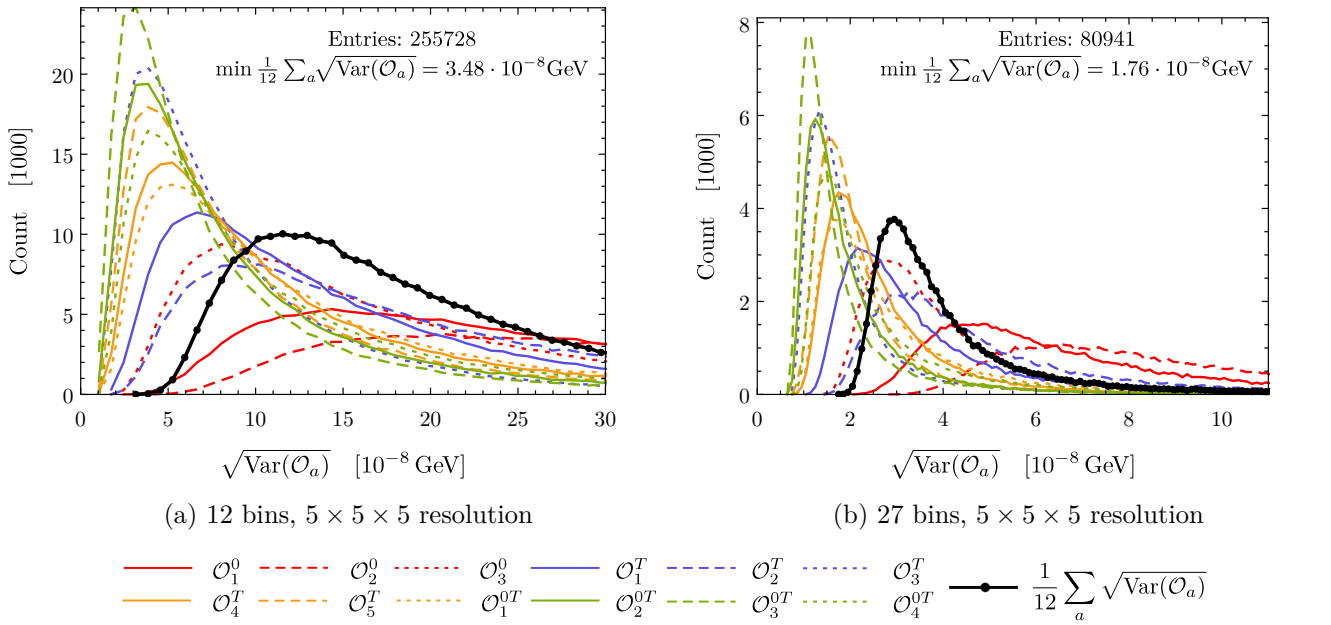
$$U(i,j,k) := U_\chi(i) \times U_{\theta_\ell}(j) \times U_{\theta_{D^*}}(k) \subseteq [-\pi, \pi] \times [0, \pi] \times [0, \pi] \tag{2.3.2}$$



(a) 12 bins, $3 \times 3 \times 3$ resolution

(b) 12 bins, $3 \times 3 \times 3$ resolution (uniform)

Figure 2.6: Distribution of the standard deviation $\sqrt{\mathrm{Var}(\mathcal{O}_a)}$ of the observables $\mathcal{O}_a$ (colored lines, denoted as $\mathcal{O}_a$ in the legend below) and of the average $\frac{1}{12}\sum_a\sqrt{\mathrm{Var}(\mathcal{O}_a)}$ (black line) for binnings in 12 subsets of the phase space. The plot markers of the black line denote the midpoints of the bins in the standard deviation. In (a), the edge points of the cubes were randomly sampled as described in the text. In (b), the edge points were fixed to divide the phase space into uniform $2\pi/3 \times \pi/3 \times \pi/3$ cubes ("uniform resolution"). For the binnings with non-uniform $3 \times 3 \times 3$ resolution, many were discarded because the matrix $\overline{B}_a(i)$ was (close to) singular, resulting in significantly lower statistics.



(a) 12 bins, $5 \times 5 \times 5$ resolution

(b) 27 bins, $5 \times 5 \times 5$ resolution

Figure 2.7: Distribution of the standard deviation $\sqrt{\mathrm{Var}(\mathcal{O}_a)}$ of the observables $\mathcal{O}_a$ for binnings in 12 and 27 arbitrary subsets of the phase space with a $5 \times 5 \times 5$ resolution. See figure 2.6 for more information on the interpretation of the plots.

for $1 \leq i \leq N_\chi$, $1 \leq j \leq N_{\theta_\ell}$, $1 \leq k \leq N_{\theta_{D^*}}$. If we only have access to e.g. the differential cross section in one variable, we can treat this scenario with the same formalism by simply assuming to only have one bin in the two other variables.

For later use, we also define

$$I = \{1, \ldots, N_\chi\} \times \{1, \ldots, N_{\theta_\ell}\} \times \{1, \ldots, N_{\theta_{D^*}}\}, \tag{2.3.3}$$

so that we can quickly write $(ijk) \in I$.

The angle dependencies B_a can be factored as

$$B_a(\chi, \theta_\ell, \theta_{D^*}) = f_{q_a}(\chi)\, g_{r_a}(\theta_\ell)\, h_{s_a}(\theta_{D^*}), \tag{2.3.4}$$

where $1 \leq q_a \leq n_\chi := 5$, $1 \leq r_a \leq n_{\theta_\ell} := 5$ and $1 \leq s_a \leq n_{\theta_{D^*}} := 3$. Explicitly, we have the functions

$$\begin{aligned}
&f_1(\chi) := 1, && g_1(\theta_\ell) := \sin(\theta_l), && \\
&f_2(\chi) := \sin(\chi), && g_2(\theta_\ell) := \sin^2(\theta_l), && h_1(\theta_{D^*}) := \sin^3(\theta_{D^*}), \\
&f_3(\chi) := \cos(\chi), && g_3(\theta_\ell) := \sin^3(\theta_l), && h_2(\theta_{D^*}) := \sin(\theta_{D^*})\sin(2\theta_{D^*}), \\
&f_4(\chi) := \sin(2\chi), && g_4(\theta_\ell) := \sin(\theta_l)\cos(\theta_l), && h_3(\theta_{D^*}) := \sin(\theta_{D^*})\cos^2(\theta_{D^*}). \\
&f_5(\chi) := \cos(2\chi), && g_5(\theta_\ell) := \sin(\theta_l)\sin(2\theta_l), &&
\end{aligned} \tag{2.3.5}$$

All three sets of functions are linearly independent (as can be checked via Lemma B.0.6).

Note that the amplitudes W_5^T, W_3^{0T} and W_4^{0T} vanish for CP averaged cross sections and that in this case f_2 and f_4 are no longer needed, reducing the number of functions in χ to three. However, f_2 and f_4 are the only odd functions in χ, and (in most cases) merely have the effect of forcing the weighting for the extraction of amplitudes proportional to f_1, f_3 or f_5 to be "symmetric". As all other functions in χ are even, this requirement is unlikely to have much impact on the sensitivity obtained for $W_a \neq W_5^T, W_3^{0T}, W_4^{0T}$ (but further studies should investigate this more carefully).

Comparing with table 2.1, the indizes q_a, r_a, s_a are then given by the matrix

a	10	20	30	$1T$	$2T$	$3T$	$4T$	$5T$	$10T$	$20T$	$30T$	$40T$
q_a	1	1	1	1	1	1	5	4	3	3	2	2
r_a	1	3	4	1	3	4	3	3	5	2	2	5
s_a	3	3	3	1	1	1	1	1	2	2	2	2

, (2.3.6)

which is also illustrated in figure 2.8. Because of the bijection $A \ni a \longleftrightarrow (q_a, r_a, s_a)$, we also use triplets (q, r, s) as indizes in place of a, e.g. $B_{113} := B_1^0$, $\overline{W}_{111} := \overline{W}_1^T$. We define the corresponding set of indizes as $A' := \{(q_a, r_a, s_a) \mid a \in A\}$.

The bins of the differential cross section are given by

$$\Gamma(i,j,k) := \int_{U(i,j,k)} \mathrm{d}\chi\, \mathrm{d}\theta_\ell\, \mathrm{d}\theta_{D^*} \int_{q^2_{\min}}^{q^2_{\max}} \mathrm{d}q^2 \frac{\mathrm{d}^4\Gamma}{\mathrm{d}q^2\, \mathrm{d}\chi\, \mathrm{d}\theta_\ell\, \mathrm{d}\theta_{D^*}}. \tag{2.3.7}$$

Defining "binned" functions in χ, θ_ℓ and θ_{D^*},

$$\begin{aligned}
F_q(i) &:= \int_{U_\chi(i)} \mathrm{d}\chi\, f_q(\chi) && 1 \leq i \leq N_\chi, \\
G_r(j) &:= \int_{U_{\theta_\ell}(j)} \mathrm{d}\theta_\ell\, g_r(\theta_\ell) && 1 \leq j \leq N_{\theta_\ell}, \\
H_s(k) &:= \int_{U_{\theta_{D^*}}(k)} \mathrm{d}\theta_{D^*}\, h_s(\theta_{D^*}) && 1 \leq k \leq N_{\theta_{D^*}},
\end{aligned} \tag{2.3.8}$$

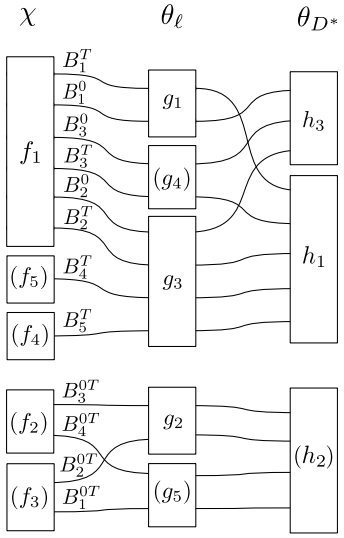


Figure 2.8: Illustration of the triples $(q_a r_a s_a) \in A'$, corresponding to $B_a(\chi, \theta_\ell, \theta_{D^*}) = f_{q_a}(\chi) g_{r_a}(\theta_\ell) g_{s_a}(\theta_{D^*})$. Note how the mixed terms B_1^{0T}, B_2^{0T}, B_3^{0T} and B_4^{0T} form their own connected component and how the longitudinal, transverse and mixed terms (designated with upper index 0, T or $0T$) each couple to a different function in θ_{D^*}. Functions in parenthesis vanish when integrated over the whole domain of the respective angle.

we have $\overline{B}_{qrs}(i,j,k) = F_q(i) G_r(j) H_s(k)$ and thus arrive at the pendant to equation (2.2.4):

$$\Gamma(i,j,k) = \sum_{(qrs)\in A'} F_q(i) G_r(j) H_s(k) \cdot \overline{W}_{qrs}. \tag{2.3.9}$$

This again has the form of a matrix equation with solutions if and only if the $12 \times N_\chi N_{\theta_\ell} N_{\theta_{D^*}}$ matrix

$$B = \left(F_q(i) G_r(j) H_s(k)\right)_{\substack{(qrs)\in A \\ (ijk)\in I}} \tag{2.3.10}$$

has full rank 12. The solutions will be of the form

$$\overline{W}_{qrs} = \sum_{i=1}^{N_\chi} \sum_{j=1}^{N_{\theta_\ell}} \sum_{k=1}^{N_{\theta_{D^*}}} \omega_{qrs}(i,j,k)\, \Gamma(i,j,k), \tag{2.3.11}$$

with (not necessarily unique) weights $\omega_a(i,j,k)$, which have to obey the equation (obtained by combining (2.3.9) with (2.3.11)):

$$\delta_{qq'}\delta_{rr'}\delta_{ss'} = \sum_{i=1}^{N_\chi} \sum_{j=1}^{N_{\theta_\ell}} \sum_{k=1}^{N_{\theta_{D^*}}} \omega_{qrs}(i,j,k) F_{q'}(i) G_{r'}(j) H_{s'}(k) \qquad \forall\, (q'r's') \in A', \tag{2.3.12}$$

i.e. are right-inverses of B.

Note that "general" bins of section 2.3.1 with corresponding weights can also be expressed in this formalism (by setting some weights to be equal), but the number of minimally required bins will of course be equal or higher. In fact, looking at matrix (2.3.6) (or figure 2.8), we can easily set a lower bound on the number of bins in each variable by counting columns which only differ in one entry: We have the columns $113, 111$ $(10, 2T)$, so we need at least two bins in θ_{D^*}, or else the $2 \times N_\chi N_{\theta_\ell} N_{\theta_{D^*}}$ sub-matrix of Q

$$\left(F_1(i) G_1(j) H_s(k)\right)_{\substack{s\in\{1,2\} \\ (ijk)\in I}} \tag{2.3.13}$$

has non-maximal (row-)rank and rank $B < 12$ as a result. Similarly, matrix (2.3.6) includes $113, 133, 143$ ($a = 10, 20, 30$), so we need at least three bins in θ_ℓ and $131, 531, 431$ ($a = 2T, 4T, 5T$), so we need at least three bins in θ_{D^*}. In summary:

$$N_\chi \geq 3,\ N_{\theta_\ell} \geq 3,\ N_{\theta_{D^*}} \geq 2 \quad \Longrightarrow \quad N_\chi N_{\theta_\ell} N_{\theta_{D^*}} \geq 18. \tag{2.3.14}$$

A simple $3 \times 3 \times 2$ binning is given by

$$U(i,j,k) := \left[-\pi + \frac{2\pi}{3}(i-1), -\pi + \frac{2\pi}{3} i\right] \times \left[\frac{\pi}{3}(j-1), \frac{\pi}{3} j\right] \times \begin{cases} [0, 2\pi/3] & k = 1, \\ [2\pi/3, \pi] & k = 2. \end{cases} \tag{2.3.15}$$

which can be confirmed by verifying that the 12×18 matrix $\overline{B}_a(i) = \overline{B}_{(qrs)(ijk)}$ has full rank. In fact, this also seems to be close to the best bin setup for $3 \times 3 \times 2$ bins (together with a second minimum with $\{0, \pi/3, \pi\}$ as bin edges for θ_{D^*}) as demonstrated by figure 2.9.

Similar studies for $3 \times 3 \times 3$ bins show that an approximately equidistant spacing of the bins yields the best results, as shown in figure 2.11.

2.3.3 Product bins, non-independent product form weights

The product form of the B_a and the bins prompts us to look for weights that come in product form as well:

$$\omega_{qrs}(i,j,k) = \omega^\chi_{qrs}(i)\omega^{\theta_\ell}_{qrs}(j)\omega^{\theta_{D^*}}_{qrs}(k). \tag{2.3.16}$$

While this places additional restrictions on the weights (and we thus cannot expect an equally good performance of the observables constructed in this way), there are a couple of advantages:

- As one dimensional functions, the weights are easy to visualize, e.g. via plots of step functions such as
$$\tilde{f}_i(\chi) = \sum_i \omega^\chi_{qrs}(i) \mathbb{1}_{U_\chi(i)}(\chi), \tag{2.3.17}$$
where $\mathbb{1}_A(\chi) = 1$ for $\chi \in A$ and zero otherwise.

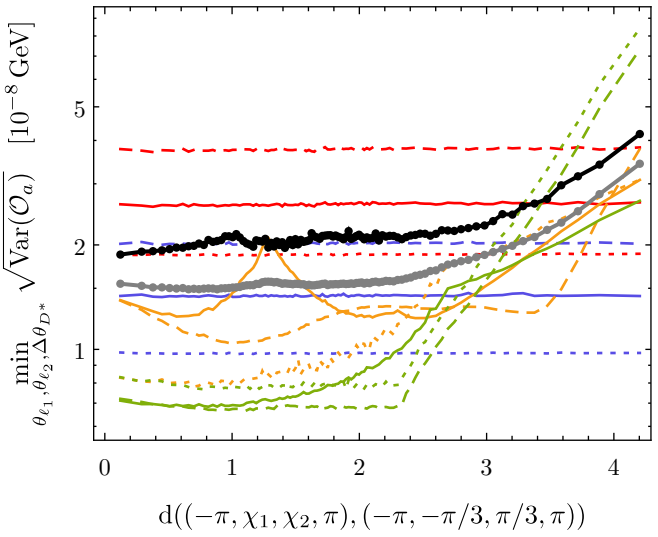
- One can consider the problem of finding weights in several steps, as the defining property of the weights will be factored as
$$\delta_{qq'}\delta_{rr'}\delta_{ss'} = \underbrace{\left(\sum_{i=1}^{N_\chi} \omega^\chi_{qrs}(i) F_{q'}(i)\right)}_{=:a_{qrs}(q')} \underbrace{\left(\sum_{j=1}^{N_{\theta_\ell}} \omega^{\theta_\ell}_{qrs}(j) G_{r'}(j)\right)}_{=:b_{qrs}(r')} \underbrace{\left(\sum_{k=1}^{N_{\theta_{D^*}}} \omega^{\theta_{D^*}}_{qrs}(k) H_{s'}(k)\right)}_{=:c_{qrs}(s')} \tag{2.3.18}$$
for all $(q'r's') \in A'$. This also allows finding simple examples with closed expressions for the weights "by hand".

- In the previous two cases we could so far only construct either all observables $\mathcal{O}_a$ at once or none at all.[6] Here, we will be able to construct the observables $\mathcal{O}_a$ separately and, by limiting ourselves to only a selection of $\mathcal{O}_a$, also work with lower numbers of bins.

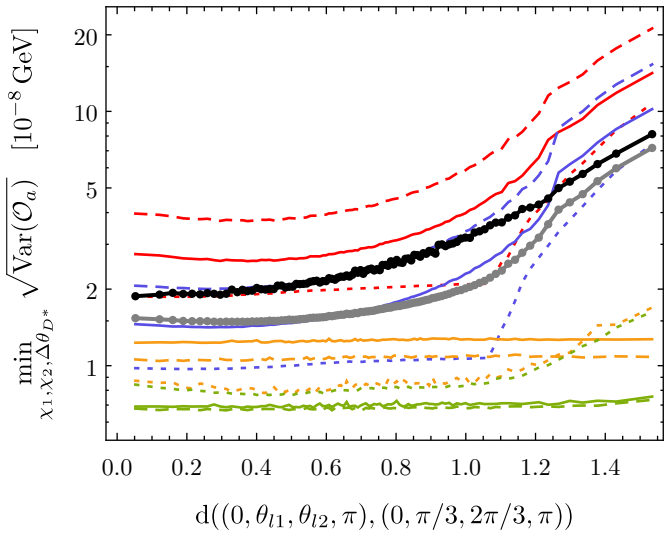
Equation (2.3.18) makes it tempting to directly draw the conclusion that $a_{qrs}(q') \overset{!}{=} \delta_{qq'}$ etc. Unfortunately, that is not entirely correct, but rather only conclusions such as

$$a_{qrs}(q') = \begin{cases} \delta_{qq'} & q'rs \in A' \\ ? & \text{else} \end{cases} \tag{2.3.19}$$

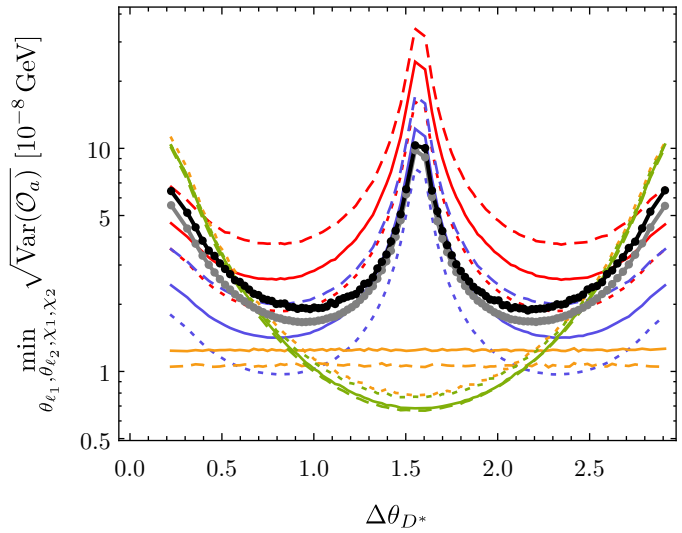
[6] Note however, that the choice of the degrees of freedom in the weights is always performed independently per observable.



(a) Optimal sensitivity vs. fixed binning in χ



(b) Optimal sensitivity vs. fixed binning in θ_ℓ



(c) Optimal sensitivity vs. fixed binning in θ_{D^*}

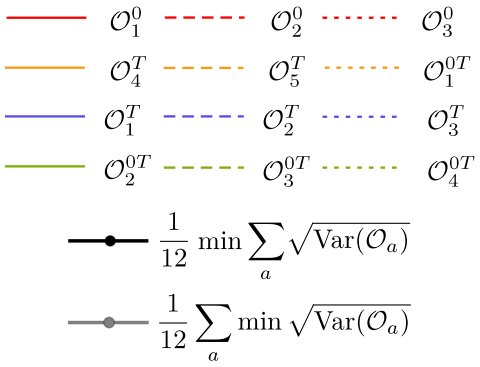


Figure 2.9: Probing for the best bin spacing for $3 \times 3 \times 2$ bins. The three plots show the dependency of the optimal sensitivity on the binning in χ, θ_ℓ and θ_{D^*}, respectively. For this, a set of 252 000 toy binnings, each with edges $\{-\pi, \chi_1, \chi_2, \pi\} \times (0, \theta_{\ell 1}, \theta_{\ell 2}, \pi\} \times \{0, \Delta\theta_{D^*}, \pi\}$, was generated as explained in section 2.3.2 (however, the numerical construction of all observables $\mathcal{O}_a$ only succeeded for 52% of the toys). First considering figure (c), the toy binnings are divided into 100 subsets (bins) $B_i^{\theta_{D^*}}$ based on their $\Delta\theta_{D^*}$ value. In order to not introduce statistical effects in the following, this is done in such a way that the number of toy binnings per bin, $\#B_i^{\theta_{D^*}}$, is constant. The colored lines represent the optimal value of $\sqrt{\mathrm{Var}(\mathcal{O}_a)}$ for each observable in each bin, i.e. $\min_{B_i^{\theta_{D^*}}} \sqrt{\mathrm{Var}(\mathcal{O}_a)}$ (note that this in particular minimizes over all values of the remaining parameters $\theta_{\ell 1}$, $\theta_{\ell 2}$, and $\Delta\theta_{D^*}$). To visualize the general trend of the colored lines, the thick gray line shows the average of these 12 values, i.e. $\frac{1}{12}\sum_a \min_{B_i} \sqrt{\mathrm{Var}(\mathcal{O}_a)}$. Note however, that the optimum for each $\mathcal{O}_a$ might be obtained for different elements of $B_i^{\theta_{D^*}}$, i.e. for different values of the remaining variables χ_1, $\chi_2\theta_{\ell 1}$ and $\theta_{\ell 2}$, so this number might in fact be misleading. To address this, the black line shows $\min_{B_i} \frac{1}{12}\sum_a \sqrt{\mathrm{Var}(\mathcal{O}_a)}$, i.e. the optimal average. The plot markers for the black and gray line show the midpoints of the bins in $\Delta\theta_{D^*}$. Figures (a) and (c) are analog. However, because the binnings for χ and θ_ℓ depend on two parameters, the $\mathbb{R}^4$ distance d of the vector of the binning edge points to the equidistant binning is taken as variable on the x axis. Due to the construction of the toy binnings, the distributions of the distances as shown on the x-axis in (a) and (c) peak near the center and thin out towards the low and high values of the distance spectrum. In contrast, the distribution in $\Delta\theta_{D^*}$ is flat, however the number of failed observable constructions increases drastically near the three peaks at low high and mid $\Delta\theta_{D^*}$.

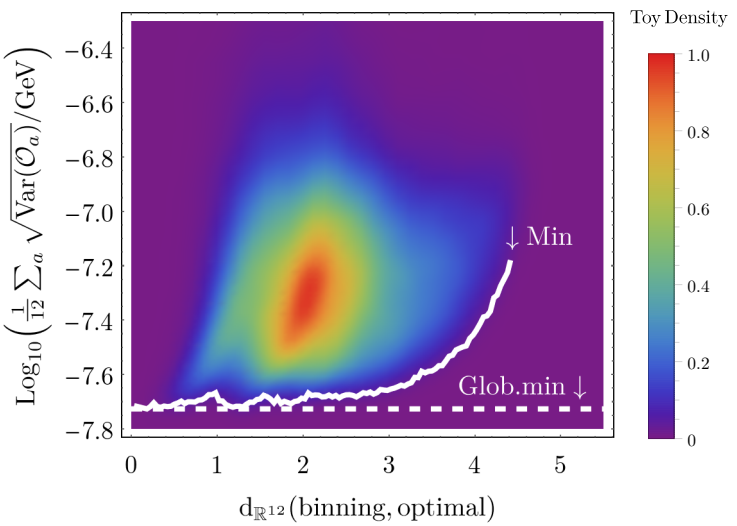


Figure 2.10: The average of the sensitivity of the observables vs. the distance from the optimal binning conjecture for $3 \times 3 \times 2$ bins with the same dataset as in figure 2.9. The x axis of this plot shows the minimal $\mathbb{R}^{12}$ distance of the concatenated vectors of the bin edges, $(-\pi, \chi_1, \chi_2, \pi, 0, \theta_{\ell 1}, \theta_{\ell 2}, \pi, 0, \Delta\theta_{D^*}, \pi)$, from our two optimal binning conjectures, $(-\pi, -\pi/3, \pi/3, \pi, 0, \pi/3, 2\pi/3, \pi, 0, \pi/3, \pi)$ and $(-\pi, -\pi/3, \pi/3, \pi, 0, \pi/3, 2\pi/3, \pi, 0, 2\pi/3, \pi)$. The y axis shows the obtained average sensitivity for the 12 observables for *one* toy binning. Binning all toy binnings in these two variables then results in the (smoothed out) density plot shown above, with the color value corresponding to the toy density. The thick white line shows the lower edge of the distribution (i.e. the best average sensitivity obtainable at a fixed distance from the optimal binning, similar to the thick black line in figure 2.9). The thick white dashed line shows the global minimum $1.88 \times 10^{-8}\,\mathrm{GeV}$ of the average sensitivity. A comparison of these two lines legitimates our conjecture about the optimal binnings.

can be drawn. Therefore we first consider a further simplification, before returning to the general problem in section 2.3.3.

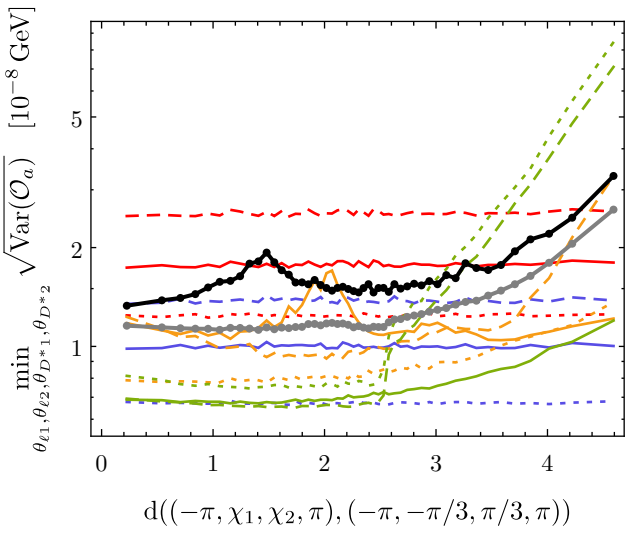
If we strengthen (2.3.16) by assuming

$$\omega_{qrs}(i,j,k) = \omega_q^{\chi}(i)\omega_r^{\theta_\ell}(j)\omega_s^{\theta_{D^*}}(k), \tag{2.3.20}$$

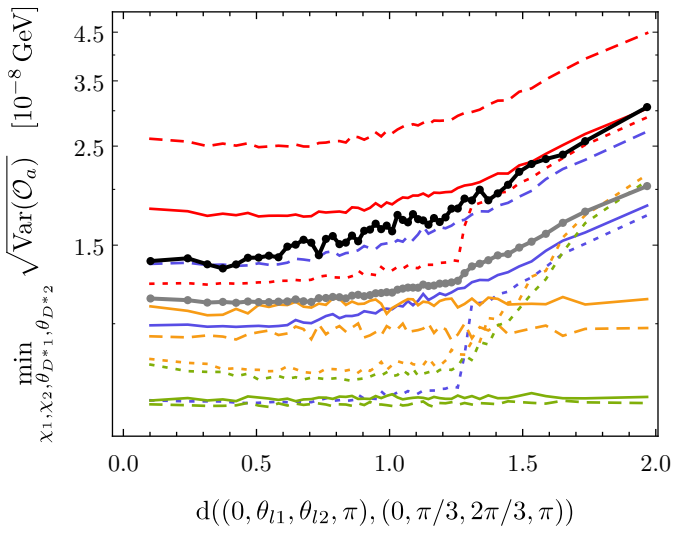
(i.e. assume that a given weight does not depend on all function indizes q, r, s, but only on one of them), the problem (2.3.18) becomes equivalent to the three independent equations

$$\begin{aligned}
\delta_{qq'} &\overset{!}{=} a_q(q') := \sum_{i=1}^{N_\chi} \omega_q^{\chi}(i) F_{q'}(i) && \forall\, 1 \le q' \le n_\chi = 5, \\
b\delta_{rr'} &\overset{!}{=} b_r(r') := \sum_{j=1}^{N_{\theta_\ell}} \omega_r^{\theta_\ell}(j) G_{r'}(j) && \forall\, 1 \le r' \le n_{\theta_\ell} = 5, \\
\delta_{ss'} &\overset{!}{=} c_s(s') := \sum_{k=1}^{N_{\theta_{D^*}}} \omega_s^{\theta_{D^*}}(k) H_{s'}(k) && \forall\, 1 \le s' \le n_{\theta_{D^*}} = 3,
\end{aligned} \tag{2.3.21}$$

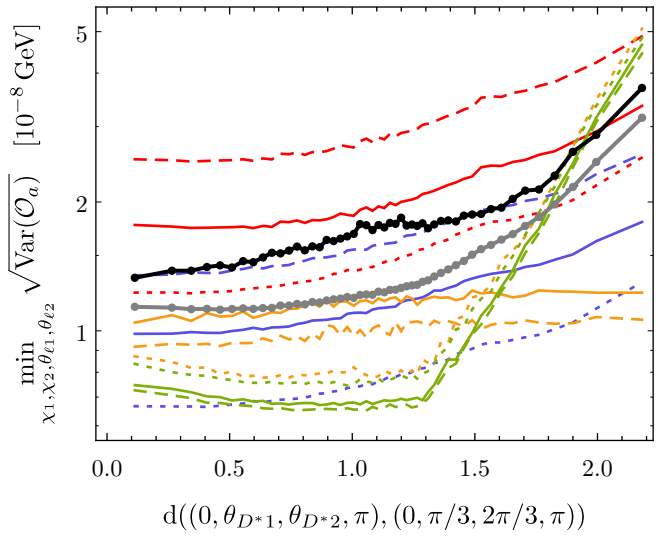
each having solutions fully characterized in appendix B. The minimal number of bins required to construct all $\mathcal{O}_a$ is then $n_\chi \times n_{\theta_\ell} \times n_{\theta_{D^*}} = N_\chi \times N_{\theta_\ell} \times N_{\theta_{D^*}} = 5 \times 5 \times 3$. For the minimal number of bins, the solutions for the weights are unique (only depending on the size of the bins), i.e. there are no degrees of freedom left in the weights. The weights for $5 \times 5 \times 3$ equidistant bins are given in figures 2.13, 2.14 and 2.15.



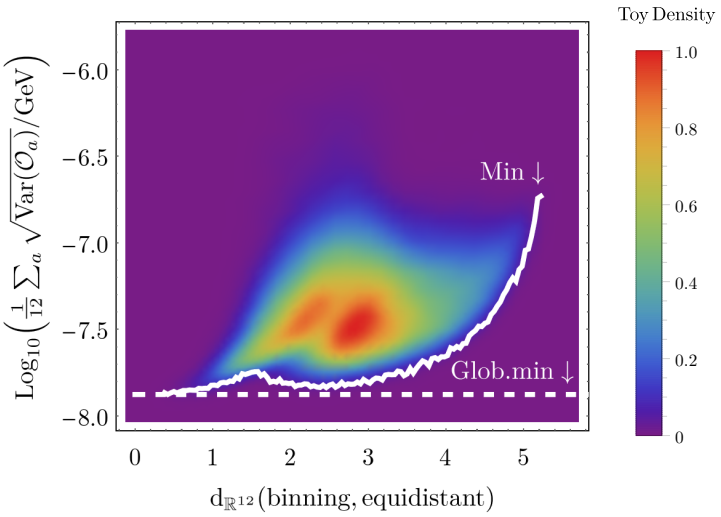
(a) Optimal sensitivity vs. fixed binning in χ



(b) Optimal sensitivity vs. fixed binning in θ_ℓ



(c) Optimal sensitivity vs. fixed binning in θ_{D^*}



(d) Optimal sensitivity vs. distance from equidistant binning

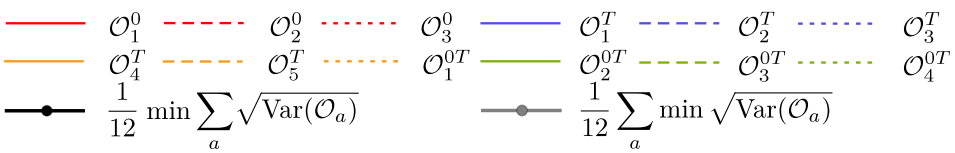


Figure 2.11: Similar to figure 2.9 and 2.10 but for $3 \times 3 \times 3$ bins. A total of 200 000 toy binnings were generated, out of which the construction of all observables succeeded for 68%.

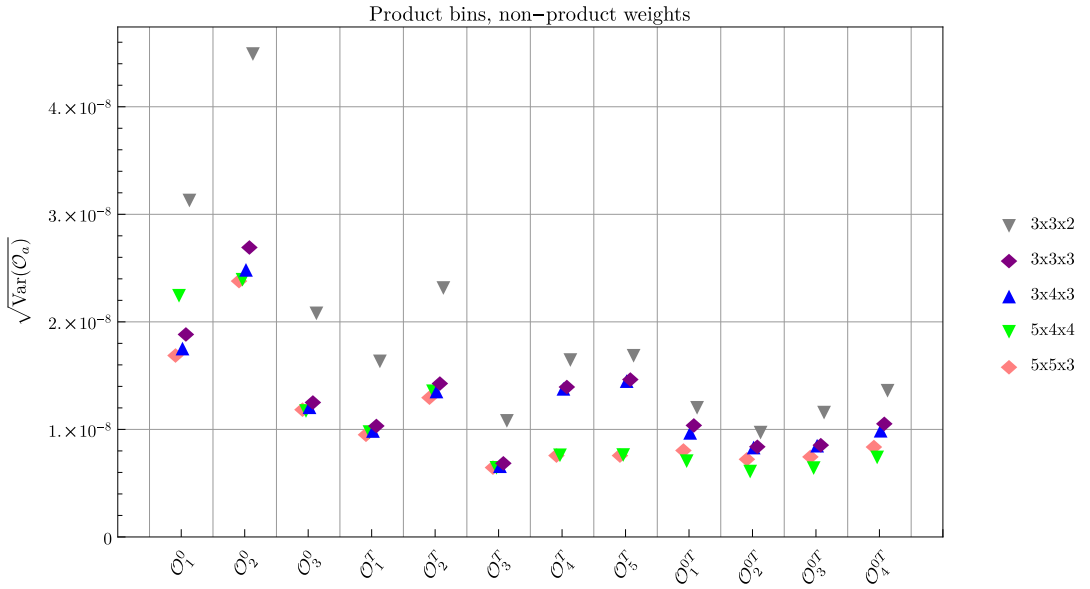
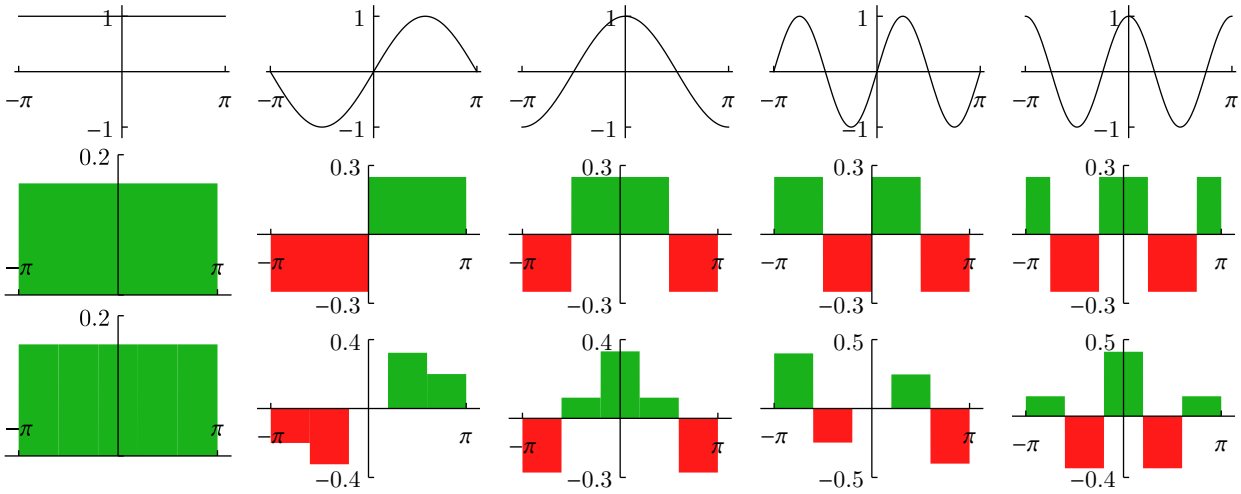


Figure 2.12: Comparison of the estimated error $\sqrt{\mathrm{Var}\,\mathcal{O}_a}$ for product bins, non-product weights. The binning for $3 \times 3 \times 2$ has been chosen as in (2.3.15), all other binnings are equidistantly spaced.



(a) $\mathbb{1}(\chi)$ (b) $\sin(\chi)$ (c) $\cos(\chi)$ (d) $\sin(2\chi)$ (e) $\cos(2\chi)$

Figure 2.13: Functions in χ and weights to separate them. First row of weights: Focus on simple expressions (given in table 2.2), second row: unique weights for equidistant bins of width $\pi/5$.

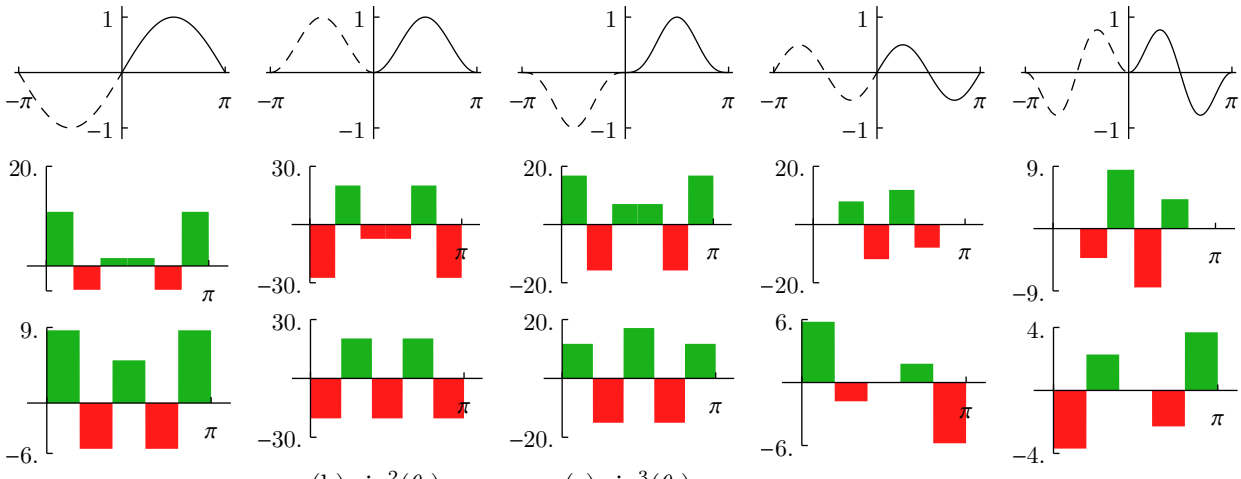
In general we have $(N_\chi - 5)(N_{\theta_\ell} - 5)(N_{\theta_{D^*}} - 3)$ degrees of freedom, which can be chosen independently for every $(qrs) \in A'$.

Oftentimes symmetry arguments make it easy to find weights by hand or at least derive constraints for them. This is especially the case for the functions in χ. Figures 2.13 and 2.15 also show an example of such weights, which are also given as expressions in tables 2.2 and 2.3.

2.3.4 Product bins, independent product form weights

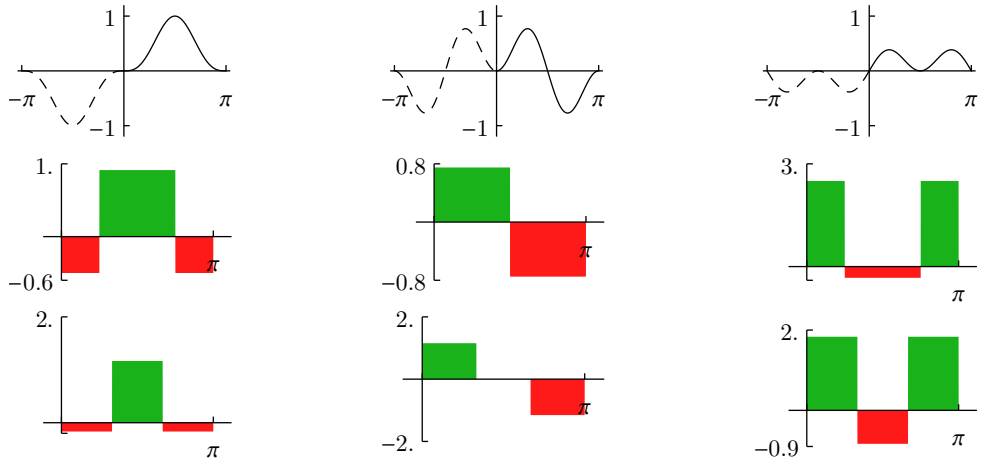
Back to the more general form of the weights as given in (2.3.16):

$$\omega_{qrs}(i,j,k) = \omega^{\chi}_{qrs}(i)\omega^{\theta_\ell}_{qrs}(j)\omega^{\theta_{D^*}}_{qrs}(k). \tag{2.3.22}$$



(a) $\sin(\theta_l)$ (b) $\sin^2(\theta_l)$ (c) $\sin^3(\theta_l)$ (d) $\sin(\theta_l)\cos(\theta_l)$ (e) $\sin(\theta_l)\sin(2\theta_l)$

Figure 2.14: Functions in θ_ℓ and weights to separate them. Note that we have only access to the domain $[0, \pi]$ of the functions. First row of weights: focus on simple expressions (equidistant bins of width $\pi/6$, with first and last weight constrained to be equal), second row: equidistant bins of width $\pi/5$.



(a) $\sin^3(\theta_{D^*})$ (b) $\sin(\theta_{D^*})\sin(2\theta_{D^*})$ (c) $\sin(\theta_{D^*})\cos^2(\theta_{D^*})$

Figure 2.15: Functions in θ_{D^*} and weights to separate them. First row of weights: bin edges $\{0, \pi/4, 3\pi/4, \pi\}$, second row: equidistant bins of width $\pi/3$. weights are given in table 2.3.

The basic idea that leads to improvements over section 2.3.3 is that, since not all combinations of functions $\{f_i\}$, $\{g_j\}$ and $\{h_k\}$ appear, it is not always necessary to distinguish between the functions from *all* three sets.

Looking at figure 2.8 it is for example very easy to see that, to construct $\mathcal{O}_4^T$ and $\mathcal{O}_5^T$, it is enough to only distinguish between the 5 functions in χ, as B_4^T and B_5^T are the only terms that have the χ dependency $f_4(\chi)$ and $f_5(\chi)$. Thus, $\mathcal{O}_4^T$ and $\mathcal{O}_5^T$ can be constructed even for single differential cross sections $\mathrm{d}\Gamma/\,\mathrm{d}\chi$ (or equivalently a $5 \times 1 \times 1$ binning; note however that O_5^T vanishes regardless of NP when CP averaging).

Besides reducing the number of bins necessary, this also allows for sensitivity improvement as there are more degrees of freedom in the weights that can be used for performance optimization.

A more general strategy to isolate B_{qrs} will therefore consist of a set of functions in χ that we have to distinguish, a set of functions in θ_ℓ that we have to distinguish and a set of functions

$$\tilde{f}_1 = \frac{1}{2\pi}[-\pi,\pi] \qquad \tilde{f}_3 = -\frac{1}{4}[-\pi,\pi] + \frac{1}{2}[-\pi/2,\pi/2]$$
$$\tilde{f}_2 = \frac{1}{4}\Big([0,\pi] - [-\pi,0]\Big) \qquad \tilde{f}_4 = \frac{1}{4}\Big([-\pi,-\pi/2] - [-\pi/2,0] + [0,\pi/2] - [\pi/2,\pi]\Big)$$
$$\tilde{f}_5 = \frac{1}{4}\Big([-\pi,-3\pi/4] - [-3\pi/4,-\pi/4] + [-\pi/4,\pi/4] - [\pi/4,3\pi/4] + [3\pi/4,\pi]\Big)$$

Table 2.2: Expressions for the weights $\omega_q^\chi(i)$ shown in the first row of weights in figure 2.13. In our short hand notation, we write intervals $[a,b]$ instead of $\mathbb{1}_{[a,b]}(\chi)$ and also drop the arguments χ.

$$\tilde{g}_1 = -\frac{1}{2}[0,\pi] + \sqrt{2}[\pi/4,3\pi/4] \qquad \tilde{g}_1 = \frac{1}{6}\Big(-[0,\pi] + 8[\pi/3,2\pi/3]\Big)$$
$$\tilde{g}_2 = \frac{3}{4}\Big([0,\pi/2] - [\pi/2,\pi]\Big) \qquad \tilde{g}_2 = \frac{2\sqrt{3}}{3}\Big([0,\pi/3] - [\pi/3,2\pi/3]\Big)$$
$$\tilde{g}_3 = \frac{5}{2}[0,\pi] - 2\sqrt{2}[\pi/4,3\pi/4] \qquad \tilde{g}_3 = \frac{1}{6}\Big(11[0,\pi] - 16[\pi/3,2\pi/3]\Big)$$

Table 2.3: Expressions for the weights shown in figure 2.15. Notation as in table 2.2.

in θ_{D^*} that we have to distinguish. That is, we consider a triple (Q,R,S) of subsets of the corresponding function indizes (referred to as *strategy*):

$$(Q,R,S) \quad \text{with} \quad Q \subseteq \{1,\dots,5\},\ R \subseteq \{1,\dots,5\},\ S \subseteq \{1,2,3\}. \tag{2.3.23}$$

The set of all strategies (Q,R,S) that allow to isolate $(qrs) \in A'$ will be called $\mathcal{S}_{qrs}$, i.e.

$$\mathcal{S}_{qrs} = \{(Q,R,S) \mid (Q,R,S) \text{ strategy to extract } B_{qrs}\}. \tag{2.3.24}$$

We furthermore introduce a partial order "$\subseteq$" on $\mathcal{S}_{qrs}$ with

$$(Q,R,S) \subseteq (Q',R',S') \ :\Longleftrightarrow\ Q \subseteq Q' \text{ and } R \subseteq R' \text{ and } S \subseteq S' \tag{2.3.25}$$

with $\supseteq$ defined accordingly and also talk of "subsets" and "supersets" correspondingly. Thus, for each $qrs \in A'$, we are looking for "minimal" sets (Q,R,S) with respect to "$\subseteq$" because any superset of a strategy for B_{qrs} will be a valid strategy as well.

Given a strategy, the number of required bins is $\#Q \times \#R \times \#S$. For completeness, we therefore also define a partial order $\leq$ on the set of all triples of integers:

$$(a,b,c) \leq (a',b',c') \ :\Longleftrightarrow\ a \leq a' \text{ and } b \leq b' \text{ and } c \leq c', \tag{2.3.26}$$

with $\geq$ defined as the opposite relation and talk about "finer" or "coarser" binnings accordingly. Note that $(Q,R,S) \subseteq (Q',R',S')$ also implies $(\#Q,\#R,\#S) \leq (\#Q',\#R',\#S')$.

Concerning our three key questions, we then give the following answers:

1. A minimal number of bins required to isolate B_{qrs} will be given by (non-unique) minima

$$\min\Big\{(\#Q,\#R,\#S) \;\Big|\; (Q,R,S) \in \mathcal{S}_{qrs}\Big\}. \tag{2.3.27}$$

A minimal number of bins required to isolate *all* B_{qrs} is given by a (non-unique) minimum

$$\min\Big\{(\max_{qrs}\#Q_{qrs}, \max_{qrs}\#R_{qrs}, \max_{qrs}\#S_{qrs}) \;\Big|\; (Q_{qrs},R_{qrs},S_{qrs}) \in \mathcal{S}_{qrs},\ qrs \in A'\Big\}. \tag{2.3.28}$$

2. Finding the most sensitive observables: Consider a given binning with $N_\chi \times N_{\theta_\ell} \times N_{\theta_{D^*}}$ bins. Let $(qrs) \in A'$ be fixed. For each corresponding strategy (Q, R, S), calculate weights $\omega_q^\chi(1), \dots, \omega_q^\chi(N_\chi)$ to discriminate between the functions $\{f_{q'} \mid q' \in Q\}$, parametrizing all degrees of freedom $\alpha_q^\chi(1), \dots, \alpha_q^\chi(N\chi - \#Q)$. Proceed similarly for R and S.

   Finally having parametrized weights $\omega_{qrs}(i,j,k)$ for all bins i,j,k, calculate the expected variance (see section 2.2.1) and minimize it with respect to the degrees of freedom, i.e. calculate

$$\underset{\substack{\alpha_q^\chi(1),\dots,\alpha_q^\chi(\#Q-N_\chi)\\ \alpha_r^{\theta_\ell}(1),\dots,\alpha_q^\chi(\#R-N_{\theta_\ell})\\ \alpha_s^{\theta_{D^*}}(1),\dots,\alpha_q^\chi(\#S-N_{\theta_{D^*}})}}{\arg\min} \operatorname{Var}\Big(\textstyle\sum_{(ijk)\in I} \omega_{qrs}(i,j,k)\,\Gamma(i,j,k)\Big). \tag{2.3.29}$$

   We thus get weights $\omega_{qrs}(i,j,k)$ for each strategy (Q,R,S). Finally select the strategy/weights with the least expected variance. It is sufficient to only consider minimal strategies, as those will have the maximum amount of degrees of freedom, which allow for the best optimization.

3. Finding "simple symbolic results": Let $(qrs) \in A'$ be fixed. Find simple step functions to distinguish between the functions with indizes in Q, R and S. Usually there will be more than one option, i.e. we have a selection $\tilde{f}_q^{(i)}$, $\tilde{g}_r^{(j)}$ and $\tilde{h}_s^{(k)}$, out of which we select the nicest triple $\tilde{f}_q^{(i)}\tilde{g}_r^{(j)}\tilde{h}_s^{(k)}$.

To find strategies $\mathcal{S}_{qrs}$ two different methods were considered:

2.3.4.1 Method 1

This is a straightforward way to get improved results and can easily be performed "by hand".

Let $(qrs) \in A'$. Define the following subsets of the triples from A':

$$A'|_{q**} = \{(qr's') \in A\}, \qquad A'|_{qr*} = \{(qrs') \in A\}, \quad \text{etc.} \tag{2.3.30}$$

Furthermore for $B' \subseteq A'$ define the projections

$$\begin{aligned} q(B') &= \{q' \mid (q'r's') \in B'\},\\ r(B') &= \{r' \mid (q'r's') \in B'\},\\ s(B') &= \{s' \mid (q'r's') \in B'\}, \end{aligned} \tag{2.3.31}$$

We proceed in three steps, basically considering figure 2.8 as a decision tree:

1. Set $Q = q(A')$, i.e. distinguish between all functions in χ,
2. Set $R = r(A'|_{q**})$, i.e. distinguish between all functions in θ_ℓ which appear together with $f_q(\chi)$
3. Set $S = s(A'|_{qr*})$, i.e. distinguish between all functions in θ_{D^*} which appear together with $f_q(\chi)g_r(\theta_{D^*})$

It is clear that by doing so, we can construct $\mathcal{O}_{qrs}$. Note that this procedure depends on the order of the variables we consider, thus it has to be repeated six times for all different orders of $\chi, \theta_\ell, \theta_{D^*}$, giving us six triples $(Q,R,S)_{qrs}$. From these, we pick the minimal one(s).

The minimal choices for $(Q,R,S)_{qrs}$ and the corresponding minimal number of required bins are shown in table 2.4. Note how the strategy necessarily has either $\#Q \geq 5$, $\#R \geq 5$ or $\#S \geq 3$ (depending on which order of variables is considered).

Minimal numbers of bins for the construction of all $\mathcal{O}_a$ are

$$3 \times 3 \times 3, \qquad 5 \times 3 \times 2 \quad \text{and} \quad 3 \times 5 \times 2, \tag{2.3.32}$$

dominated by the construction of $\mathcal{O}_2^T$.

2.3.4.2 Method 2

Method 2 basically brute-forces the problem. Let $(qrs) \in A'$ be fixed.

Consider arbitrary subsets of the function indizes ($2^5 \cdot 2^5 \cdot 2^3 = 8192$ possibilities in total):

$$\begin{aligned} Q \subseteq q(A') = \{1, 2, \ldots, 5\} \qquad & \text{with} \quad q \in Q, \\ R \subseteq r(A') = \{1, 2, \ldots, 5\} \qquad & \text{with} \quad r \in R, \\ S \subseteq s(A') = \{1, 2, 3\} \qquad & \text{with} \quad s \in S. \end{aligned} \tag{2.3.33}$$

As before, we suppose that our procedure can distinguish between the members of these subsets, but does not gain any insight on their complements $Q^c = q(A')\backslash Q$ etc. We now check whether we can isolate (qrs) by only distinguishing between those functions. If, for any other $(qrs) \neq (q'r's') \in A'$, we have that

$$\begin{aligned} & (q' = q \text{ or } q' \in Q^c) \\ \text{and } & (r' = r \text{ or } r' \in R^c) \\ \text{and } & (s' = s \text{ or } s' \in S^c), \end{aligned} \tag{2.3.34}$$

then (qrs) and $(q'r's')$ are assumed to be not distinguishable. For our fixed $(qrs) \in A'$, we collect all Q, R, S where this case does not happen.

The minimal choices for $(Q, R, S)_{qrs}$ and the corresponding minimal number of required bins are shown in table 2.5. As the ways to separate the "bottleneck" B_2^T are unchanged, the minimally required number of bins to separate all B_2^T remains unchanged as well.

All results shown in the following are based on the strategies obtained using method 2.

2.3.4.3 Simple results

One possibility of weights with as simple of an expression as possible is given in table 2.7 and illustrated in table 2.6. Note that we give the results in the basis $(\chi, \cos\theta_\ell, \cos\theta_{D^*})$ rather than $(\chi, \theta_\ell, \theta_{D^*})$ for a better comparison with literature (but this "change of coordinates" does not affect the weights themselves). The correlation matrix of the observables $\mathcal{O}_a$ is shown in figure 2.17a. Comparing with figure 2.5, we verify that high correlations of the angle dependencies $\mathrm{Corr}(B_a, B_{a'})$ go along with high correlations of the corresponding observables $\mathrm{Corr}(\mathcal{O}_a, \mathcal{O}_{a'})$.

The sensitivities of these observables are shown in figure 2.18. Compared to results that are optimized for sensitivity, the expected errors on each observable are increased by a factor of 5 to 50 (cf. e.g. figure 2.21a).

2.3.4.4 Results with optimized variance

As a first test, $\mathrm{Var}(\mathcal{O}_a)$ is calculated for a range of equidistantly spaced binnings (i.e. given bin numbers n_χ, n_{θ_ℓ} and $n_{\theta_{D^*}}$, the bin width is $2\pi/n_\chi$, π/n_{θ_ℓ}, and $\pi/n_{\theta_{D^*}}$). The results are shown in figure 2.16. The number of bins has a surprisingly small influence on the variance expectation: For solutions with $\mathcal{O}(10)$ bins to $\mathcal{O}(1000)$ bins, the calculated variance values are at most a factor of five apart.

The binning $3 \times 3 \times 3$ (the smallest number of bins to construct all $\mathcal{O}_a$ under our assumptions) is a promising candidate. Only $\mathcal{O}_4^T$ and $\mathcal{O}_5^T$ show significantly reduced performance for $3 \times 3 \times 3$ bins than for slightly finer binnings such as $5 \times 5 \times 3$.

$\mathcal{O}_1^0$ $\mathcal{O}_2^0$ $\mathcal{O}_3^0$ $\mathcal{O}_1^T$	$\mathcal{O}_2^T$	$\mathcal{O}_3^T$	$\mathcal{O}_4^T$ $\mathcal{O}_5^T$	$\mathcal{O}_1^{0T}$ $\mathcal{O}_2^{0T}$ $\mathcal{O}_3^{0T}$ $\mathcal{O}_4^{0T}$
$\{1,2,3,4,5\},\{1,3,4\},\{1,3\}$ $\{\mathbf{1}\},\{\mathbf{1,2,3,4,5}\},\{\mathbf{1,3}\}\}$ $\{\mathbf{1}\},\{\mathbf{1,3,4}\},\{\mathbf{1,2,3}\}$	$\{\mathbf{1,2,3,4,5}\},\{\mathbf{1,3,4}\},\{\mathbf{1,3}\}$ $\{\mathbf{1,4,5}\},\{\mathbf{1,2,3,4,5}\},\{\mathbf{1,3}\}$ $\{\mathbf{1,4,5}\},\{\mathbf{1,3,4}\},\{\mathbf{1,2,3}\}$	$\{1,2,3,4,5\},\{1,3,4\},\{1,3\}$ $\{\mathbf{1}\},\{\mathbf{1,2,3,4,5}\},\{\mathbf{1,3}\}$ $\{\mathbf{1}\},\{\mathbf{1,3,4}\},\{\mathbf{1,2,3}\}$	$\{\mathbf{1,2,3,4,5}\},\{\mathbf{3}\},\{\mathbf{1}\}$ $\{1,4,5\},\{1,2,3,4,5\},\{1\}$ $\{1,4,5\},\{3\},\{1,2,3\}$	$\{\mathbf{1,2,3,4,5}\},\{\mathbf{2,5}\},\{\mathbf{2}\}$ $\{\mathbf{2,3}\},\{\mathbf{1,2,3,4,5}\},\{\mathbf{2}\}$ $\{\mathbf{2,3}\},\{\mathbf{2,5}\},\{\mathbf{1,2,3}\}$
$5\times3\times2=30$ $\mathbf{1\times5\times2=10}$ $\mathbf{1\times3\times3=9}$	$\mathbf{5\times3\times2=30}$ $\mathbf{3\times5\times2=30}$ $\mathbf{3\times3\times3=27}$	$5\times3\times2=30$ $\mathbf{1\times5\times2=10}$ $\mathbf{1\times3\times3=9}$	$\mathbf{5\times1\times1=5}$ $3\times5\times1=15$ $3\times1\times3=9$	$\mathbf{5\times2\times1=10}$ $\mathbf{2\times5\times1=10}$ $\mathbf{2\times2\times3=12}$

Table 2.4: Results of method 1 for product bins with non-independent product weights: Sets Q_a, R_a and S_a (given in that order in each row of the upper half table) of the indizes of distinguished functions in χ, θ_ℓ and θ_{D^*} respectively, that allow to construct $\mathcal{O}_a$ (columns). Lower half: Corresponding minimal number of bins in χ, θ_ℓ and θ_{D^*} and total number of bins required to construct $\mathcal{O}_a$. Options in bold face cannot be improved by method 2.

$\mathcal{O}_1^0$ $\mathcal{O}_2^0$ $\mathcal{O}_3^0$ $\mathcal{O}_1^T$	$\mathcal{O}_2^T$	$\mathcal{O}_3^T$	$\mathcal{O}_4^T$ $\mathcal{O}_5^T$	$\mathcal{O}_1^{0T}$ $\mathcal{O}_2^{0T}$ $\mathcal{O}_3^{0T}$ $\mathcal{O}_4^{0T}$
$\{1\},\{1,3,4\},\{1,2,3\}$ $\{1\},\{1,2,3,4,5\},\{1,3\}$ $\{\mathbf{1,2,3}\},\{\mathbf{1,3,4}\},\{\mathbf{1,3}\}$	$\{1,4,5\},\{1,3,4\},\{1,2,3\}$ $\{1,4,5\},\{1,2,3,4,5\},\{1,3\}$ $\{1,2,3,4,5\},\{1,3,4\},\{1,3\}$	$\{1\},\{1,3,4\},\{1,2,3\}$ $\{1\},\{1,2,3,4,5\},\{1,3\}$ $\{\mathbf{1,2,3}\},\{\mathbf{1,3,4}\},\{\mathbf{1,3}\}$	$\{\mathbf{1,4,5}\},\{\mathbf{3}\},\{\mathbf{1,2}\}$ $\{\mathbf{1,4,5}\},\{\mathbf{2,3,5}\},\{\mathbf{1}\}$ $\{1,2,3,4,5\},\{3\},\{1\}$	$\{2,3\},\{2,5\},\{1,2,3\}$ $\{2,3\},\{1,2,3,4,5\},\{1\}$ $\{\mathbf{1,2,3}\},\{\mathbf{2,5}\},\{\mathbf{1,2}\}$ $\{\mathbf{1,2,3}\},\{\mathbf{2,3,5}\},\{\mathbf{2}\}$ $\{1,2,3,4,5\},\{2,5\},\{2\}$
$1\times3\times3=9$ $1\times5\times2=10$ $\mathbf{3\times3\times2=18}$	$3\times3\times3=27$ $3\times5\times2=30$ $5\times3\times2=30$	$1\times3\times3=9$ $1\times5\times2=10$ $\mathbf{3\times3\times2=18}$	$\mathbf{3\times1\times2=6}$ $\mathbf{3\times3\times1=9}$ $5\times1\times1=5$	$2\times2\times3=12$ $2\times5\times1=10$ $\mathbf{3\times2\times2=12}$ $\mathbf{3\times3\times1=9}$ $5\times2\times1=10$

Table 2.5: Results of method 2 for product bins with non-independent product weights: Sets Q_a, R_a and S_a (given in that order in each row of the upper half table) of indizes of distinguished functions in χ, θ_ℓ and θ_{D^*} respectively, that allow to construct $\mathcal{O}_a$ (columns). Lower half: Corresponding minimal number of bins in χ, θ_ℓ and θ_{D^*} and total number of bins required to construct $\mathcal{O}_a$. Options in bold face are an improvement in comparison to method 1.

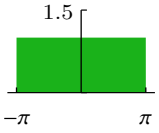
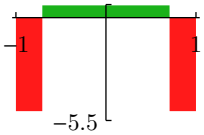
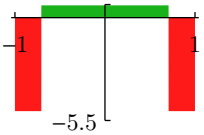
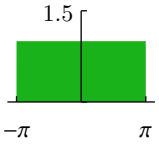
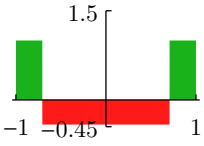
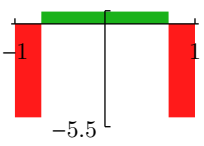
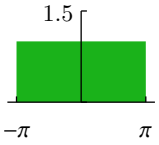
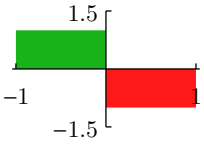
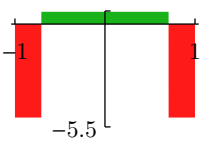
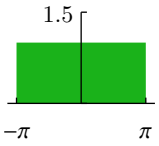
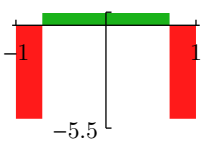
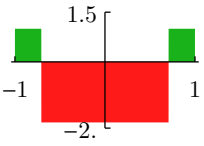
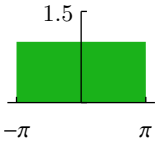
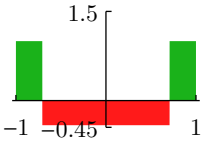
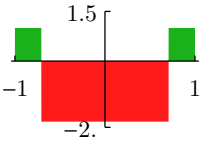
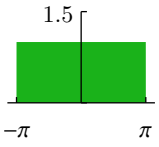
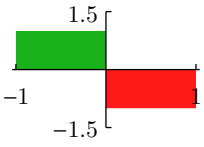
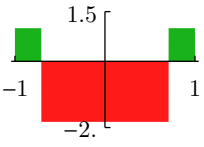
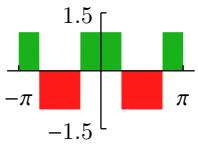
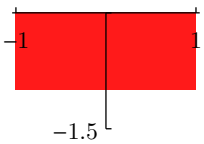
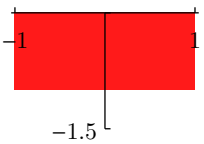
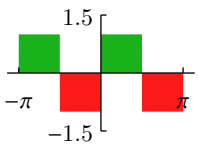
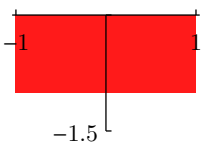
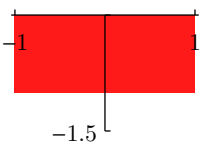
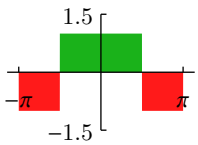
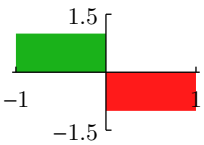
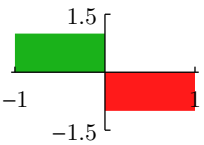
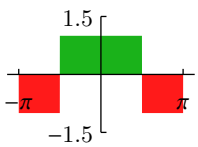
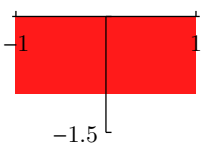
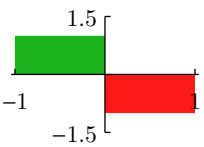
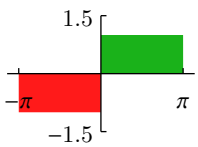
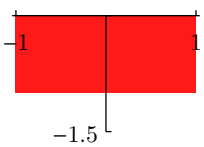
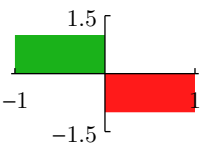
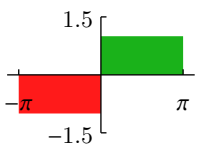
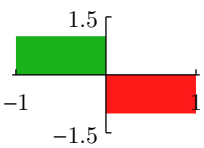
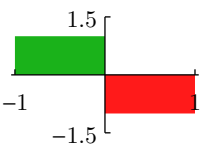
	Norm	$\tilde{f}_i(\chi)$	$\tilde{g}_j(\cos\theta_\ell)$	$\tilde{g}_k(\cos\theta_{D^*})$
$\mathcal{O}_1^0$	$\frac{1}{8\pi}$			
$\mathcal{O}_2^0$	$\frac{3}{4\pi}$			
$\mathcal{O}_3^0$	$\frac{1}{4\pi}$			
$\mathcal{O}_1^T$	$\frac{1}{8\pi}$			
$\mathcal{O}_2^T$	$\frac{3}{4\pi}$			
$\mathcal{O}_3^T$	$\frac{1}{4\pi}$			
$\mathcal{O}_4^T$	$\frac{9}{64}$			
$\mathcal{O}_5^T$	$\frac{9}{64}$			
$\mathcal{O}_1^{0T}$	$\frac{9}{64}$			
$\mathcal{O}_2^{0T}$	$\frac{3}{8\pi}$			
$\mathcal{O}_3^{0T}$	$\frac{3}{8\pi}$			
$\mathcal{O}_4^{0T}$	$\frac{9}{64}$			

Table 2.6: Extracting every W_a with simple step functions.

	Normalization	$\tilde{f}_i(\chi)$	$\tilde{g}_j(\theta_\ell)$	$\tilde{g}_k(\theta_{D^*})$
$\mathcal{O}_1^0$	$\frac{1}{8\pi}$	$[-\pi,\pi]$	$-5[-1,1]+4\sqrt{2}[-\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}]$	$-5[-1,1]+4\sqrt{2}[-\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}]$
$\mathcal{O}_2^0$	$\frac{3}{4\pi}$	$[-\pi,\pi]$	$[-1,1]-\sqrt{2}[-\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}]$	$-5[-1,1]+4\sqrt{2}[-\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}]$
$\mathcal{O}_3^0$	$\frac{1}{4\pi}$	$[-\pi,\pi]$	$[-1,0]-[0,1]$	$-5[-1,1]+4\sqrt{2}[-\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}]$
$\mathcal{O}_1^T$	$\frac{1}{8\pi}$	$[-\pi,\pi]$	$-5[-1,1]+4\sqrt{2}[-\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}]$	$[-1,1]-2\sqrt{2}[-\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}]$
$\mathcal{O}_2^T$	$\frac{3}{4\pi}$	$[-\pi,\pi]$	$[-1,1]-\sqrt{2}[-\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}]$	$[-1,1]-2\sqrt{2}[-\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}]$
$\mathcal{O}_3^T$	$\frac{1}{4\pi}$	$[-\pi,\pi]$	$[-1,0]-1[0,1]$	$[-1,1]-2\sqrt{2}[-\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}]$
$\mathcal{O}_4^T$	$\frac{9}{64}$	$[-\pi,-\frac{3\pi}{4}]-[-\frac{3\pi}{4},-\frac{\pi}{4}]+[-\frac{\pi}{4},\frac{\pi}{4}]-[\frac{\pi}{4},\frac{3\pi}{4}]+[\frac{3\pi}{4},\pi]$	$-[-1,1]$	$-[-1,1]$
$\mathcal{O}_5^T$	$\frac{9}{64}$	$[0,\frac{\pi}{2}]+[-\pi,-\frac{\pi}{2}]-[-\frac{\pi}{2},0]-[\frac{\pi}{2},\pi]$	$-[-1,1]$	$-[-1,1]$
$\mathcal{O}_1^{0T}$	$\frac{9}{64}$	$-[-\pi,\pi]+2[-\frac{\pi}{2},\frac{\pi}{2}]$	$[-1,0]-[0,1]$	$[-1,0]-[0,1]$
$\mathcal{O}_2^{0T}$	$\frac{3}{8\pi}$	$-[-\pi,\pi]+2[-\frac{\pi}{2},\frac{\pi}{2}]$	$-[-1,1]$	$[-1,0]-[0,1]$
$\mathcal{O}_3^{0T}$	$\frac{3}{8\pi}$	$[0,\pi]-[-\pi,0]$	$-[-1,1]$	$[-1,0]-[0,1]$
$\mathcal{O}_4^{0T}$	$\frac{9}{64}$	$[0,\pi]-[-\pi,0]$	$[-1,0]-[0,1]$	$[-1,0]-[0,1]$

Table 2.7: Expressions for the weights shown in figure 2.6. Notation as in 2.2.

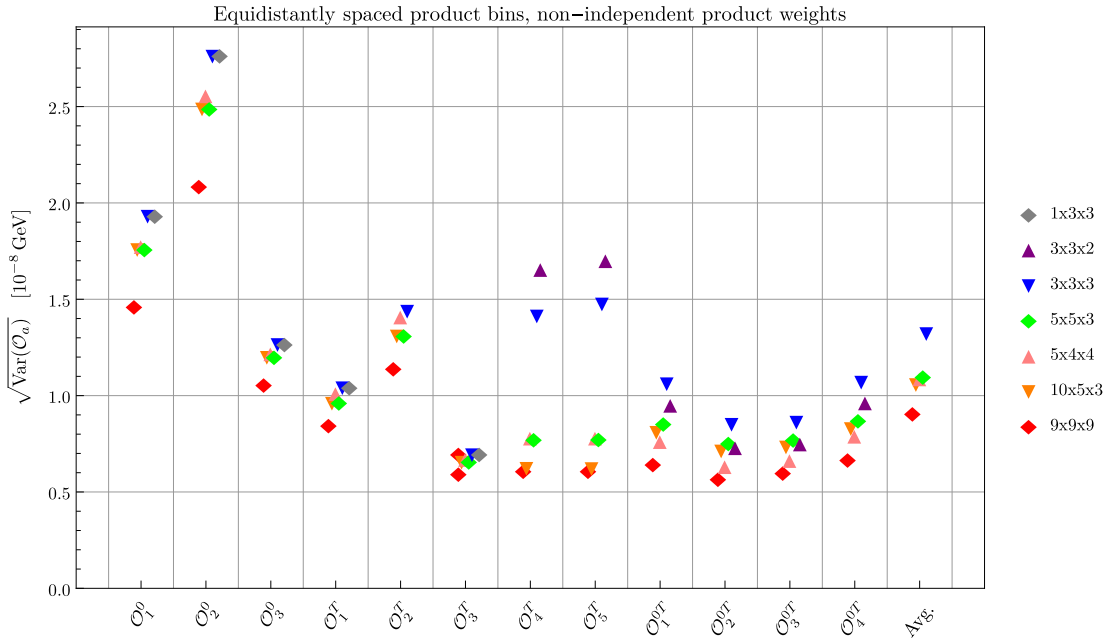


Figure 2.16: Comparison of uncertainties associated with observables $\mathcal{O}_a$ for equidistantly spaced product bins with non-independent product weights.

The influence of the bin edges was probed by calculating the variances for $\mathcal{O}_a$ using randomly generated binnings, in particular for $3 \times 3 \times 3$ bins (as the influence of the bin spacing is expected to play less and less influence for increasingly high bin numbers). As all functions are either even or odd around the middle of their domain (0 or $\pi/2$), first "symmetric" binnings were considered, i.e.

$$[-\pi, -\Delta\chi, \Delta\chi, \pi] \times \left[0, \frac{\pi}{2} - \Delta\theta_\ell, \frac{\pi}{2} + \Delta\theta_\ell, \pi\right] \times \left[0, \frac{\pi}{2} - \Delta\theta_{D^*}, \frac{\pi}{2} + \Delta\theta_{D^*}, \pi\right], \tag{2.3.35}$$

which have the advantage that the parameter space to sample only has 3 dimensions $\Delta\chi$, $\Delta\theta_\ell$ and $\Delta\theta_{D^*}$. The results are shown in figure 2.19 and seem to indicate, that in fact spacing all bins equidistantly ($\Delta\chi = \Delta\theta_\ell = \Delta\theta_{D^*} \approx \pi/3$) is an optimal solution (with a second solution $\Delta\chi \approx 2\pi/3$).

This result was sought to be confirmed without the symmetry requirement, with a six-dimensional parameter space to sample. With the higher dimension to sample, not the same resolution can be obtained, but figure 2.20 clearly confirms the trend towards an equidistant binning for high sensitivities.

The correlation matrix of the equidistant $3 \times 3 \times 3$ solution is shown in figure 2.17b and generally follows the trend of the "beautiful solution".

2.3.4.5 Further improvements

The central assumption above was that we can characterize a strategy for the isolation of qrs with the three subsets Q, R, S of functions between which we can distinguish. However, being even more general, what we need is not a way to distinguish all functions of, say, Q, but we only need to distinguish the function with index q from all others in Q, but do not have to distinguish *between* these themselves. In other words, we only have to exclude all $q' \in Q$ with $q' \neq q$ and this is a simpler task.

For example, given one symmetric function together with $N - 1 \geq 2$ antisymmetric ones, we can always isolate the symmetric function with only one bin with trivial weight 1 (instead of at least N), simply because all others vanish when integrated over the whole interval. Furthermore, even

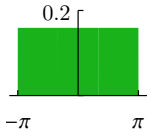
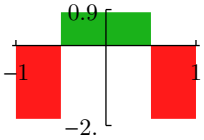
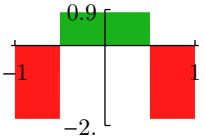
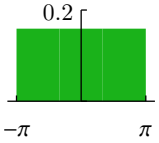
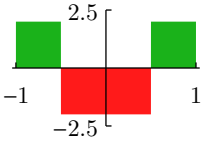
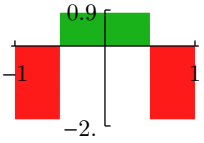
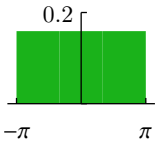
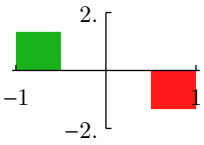
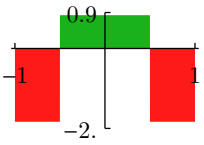
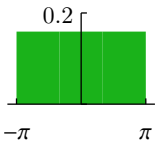
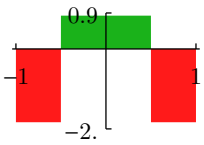
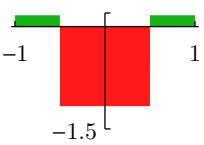
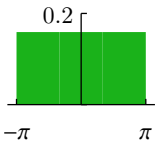
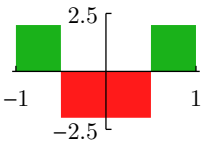
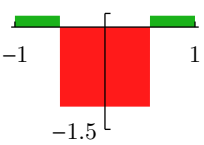
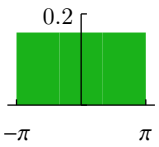
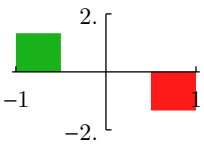
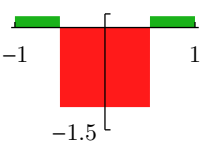
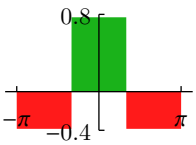
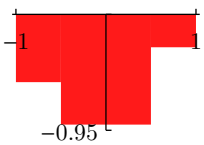
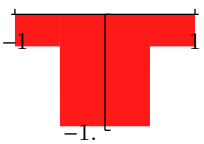
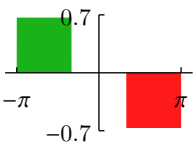
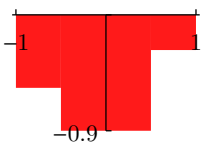
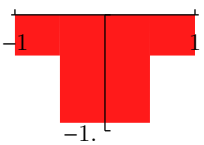
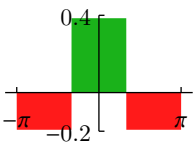
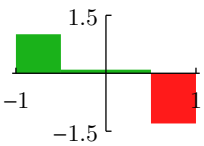
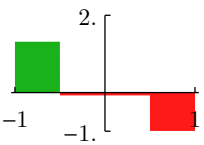
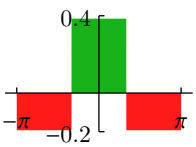
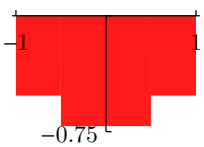
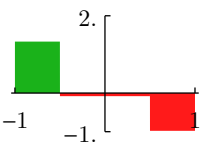
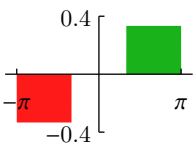
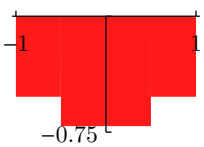
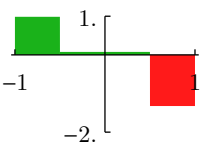
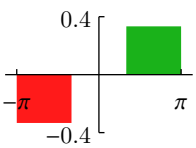
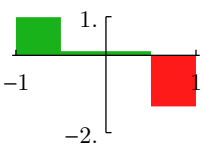
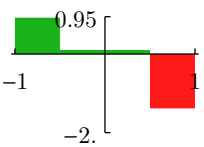
	$\sqrt{\mathrm{Var}(\mathcal{O}_a)}/\mathrm{GeV}$	$\tilde{f}_i(\chi)$	$\tilde{g}_j(\cos\theta_\ell)$	$\tilde{g}_k(\cos\theta_{D^*})$
$\mathcal{O}_1^0$	1.94×10^{-8}			
$\mathcal{O}_2^0$	2.77×10^{-8}			
$\mathcal{O}_3^0$	1.27×10^{-8}			
$\mathcal{O}_1^T$	1.05×10^{-8}			
$\mathcal{O}_2^T$	1.44×10^{-8}			
$\mathcal{O}_3^T$	6.99×10^{-9}			
$\mathcal{O}_4^T$	1.42×10^{-8}			
$\mathcal{O}_5^T$	1.49×10^{-8}			
$\mathcal{O}_1^{0T}$	1.07×10^{-8}			
$\mathcal{O}_2^{0T}$	8.57×10^{-9}			
$\mathcal{O}_3^{0T}$	8.69×10^{-9}			
$\mathcal{O}_4^{0T}$	1.08×10^{-8}			

Table 2.8: Optimal results for $3 \times 3 \times 3$ bins with non-independent product weights.

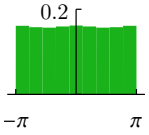
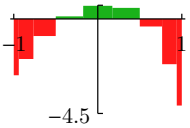
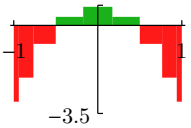
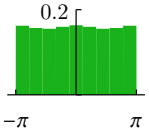
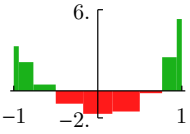
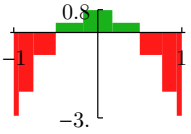
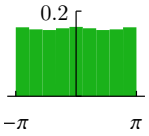
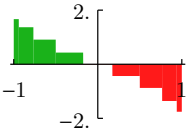
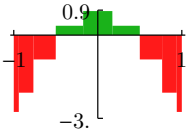
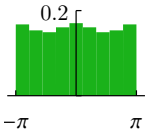
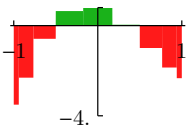
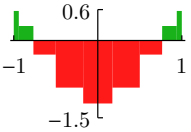
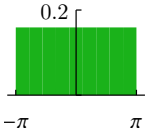
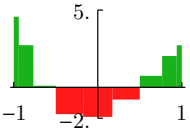
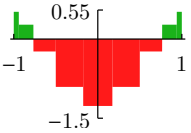
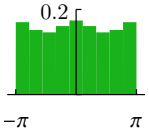
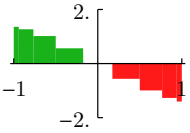
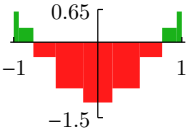
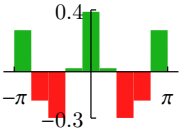
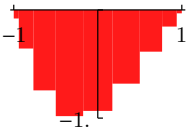
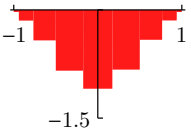
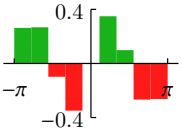
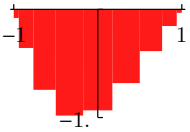
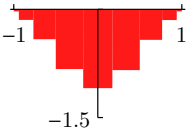
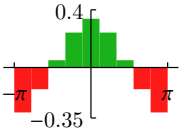
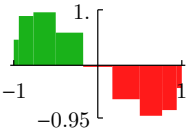
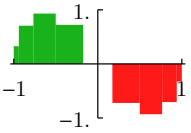
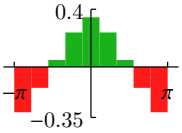
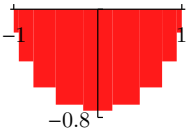
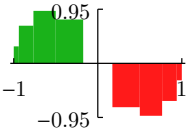
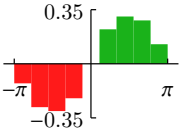
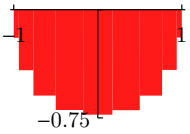
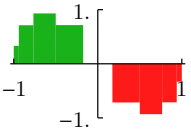
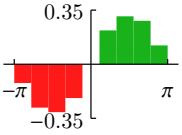
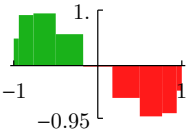
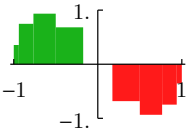
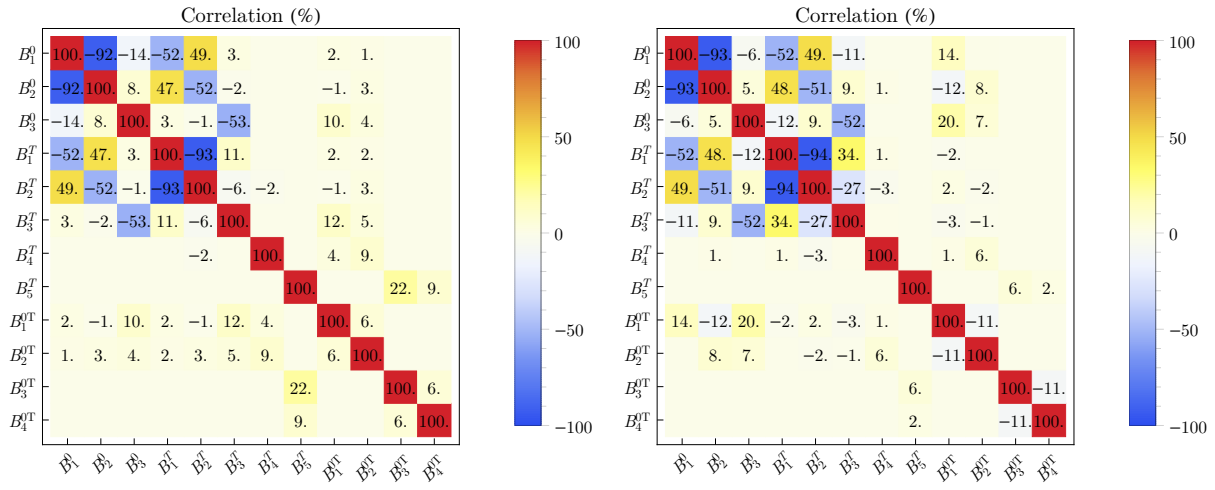
	$\sqrt{\mathrm{Var}(\mathcal{O}_a)}$/GeV	$\tilde{f}_i(\chi)$	$\tilde{g}_j(\cos\theta_\ell)$	$\tilde{g}_k(\cos\theta_{D^*})$
$\mathcal{O}_1^0$	1.47×10^{-8}			
$\mathcal{O}_2^0$	2.09×10^{-8}			
$\mathcal{O}_3^0$	1.06×10^{-8}			
$\mathcal{O}_1^T$	8.50×10^{-9}			
$\mathcal{O}_2^T$	1.15×10^{-8}			
$\mathcal{O}_3^T$	5.98×10^{-9}			
$\mathcal{O}_4^T$	6.13×10^{-9}			
$\mathcal{O}_5^T$	6.13×10^{-9}			
$\mathcal{O}_1^{0\mathrm{T}}$	$6,47 \times 10^{-9}$			
$\mathcal{O}_2^{0\mathrm{T}}$	5.72×10^{-9}			
$\mathcal{O}_3^{0\mathrm{T}}$	6.03×10^{-9}			
$\mathcal{O}_4^{0\mathrm{T}}$	6.72×10^{-9}			

Table 2.9: Optimal results for $9 \times 9 \times 9$ bins with non-independent product weights.



(a) "Beautiful solution"

(b) $3 \times 3 \times 3$ equidistant bins

Figure 2.17: Correlation matrices.

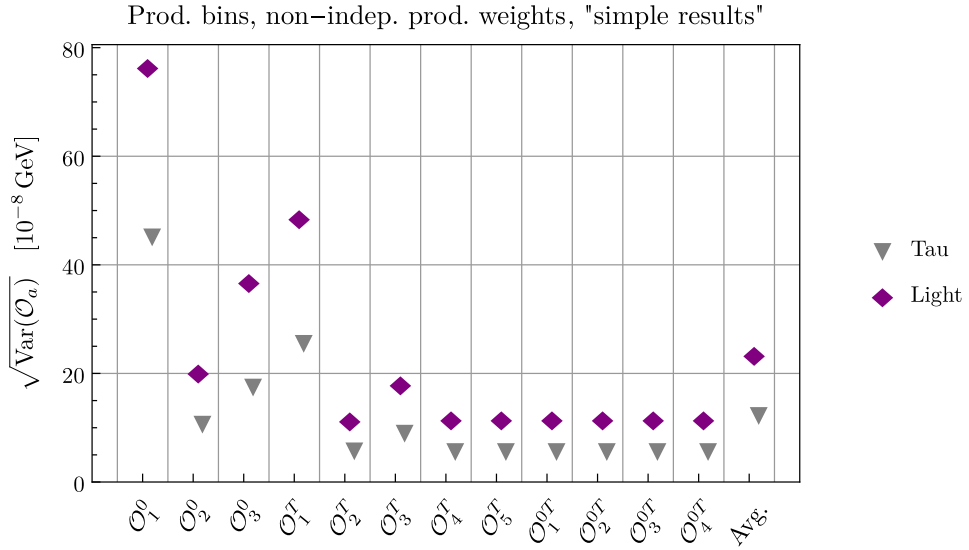
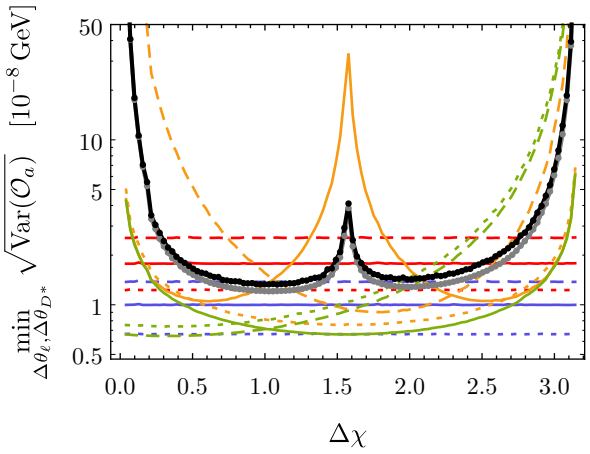


Figure 2.18: Expected errors $\sqrt{\mathrm{Var}(\mathcal{O}_a)}$ on the observables $\mathcal{O}_a$ given in table 2.7 and illustrated in figure 2.6.

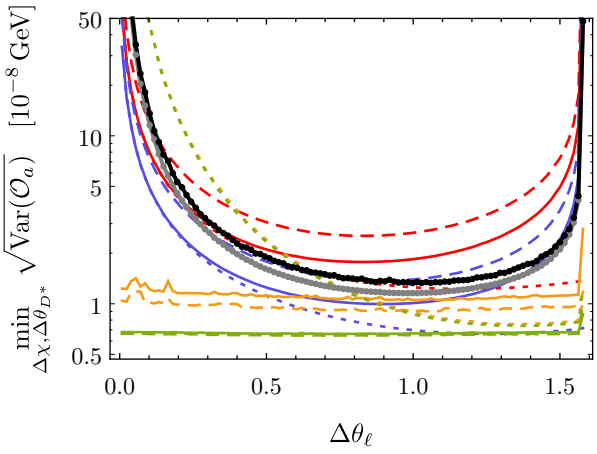
when being given N bins, if they are symmetric, we have $\lceil N/2 \rceil - 1$ degrees of freedom! Note that the approach from corollary B.0.4 will not give any solution here at all (as can be seen e.g. from lemma B.0.8, because $N - 1 > \lfloor N/2 \rfloor$).

The problem with this argument is that it depends very much on the exact placement of the bins. Considering the same example, but being given two bins which are not symmetric, weights to extract the symmetric function might no longer exist at all. While the strategy of corollary B.0.4 will almost always[7] work when being given bins with random edges, a procedure with fewer bins than functions will almost never work and the expected improvements might no longer be justified by the effort to investigate this further.

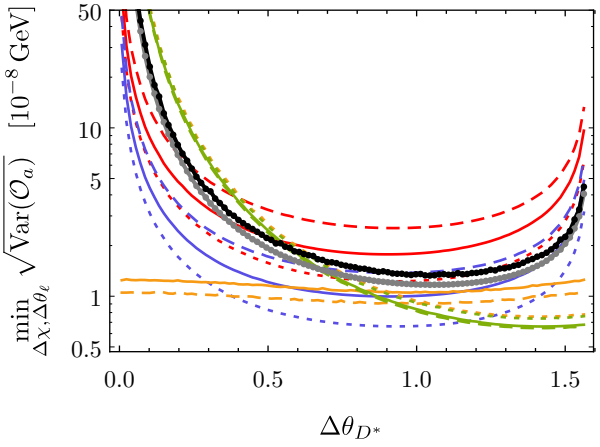
[7]The necessary requirement for the construction of the observables is that a matrix of function values has a non-zero determinant. However, the set of invertible matrices lies dense in the space of all matrices, thus hitting a zero-determinant matrix is very unlucky (and the fact that it happens a lot when trying around to find good bins is due to our aesthetic preference for symmetric bins with "nice" edge points, which often happen to fall together with symmetry properties of the considered functions).



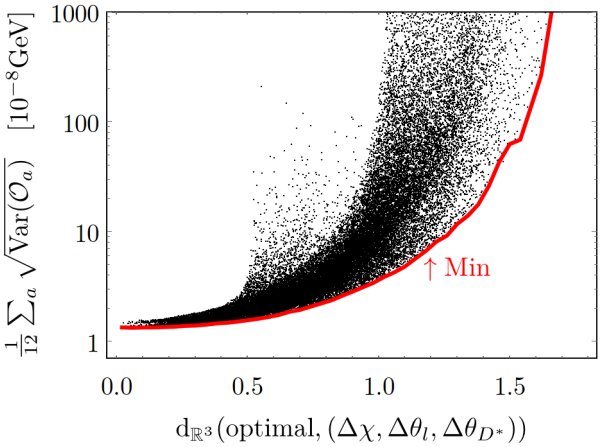
(a) Optimal sensitivity vs. fixed binning in χ



(b) Optimal sensitivity vs. fixed binning in θ_ℓ



(c) Optimal sensitivity vs. fixed binning in θ_{D^*}



(d) Sensitivity vs. distance from optimum conjectures

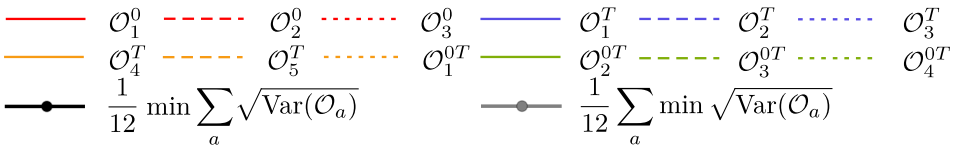


Figure 2.19: Expected uncertainty for $3 \times 3 \times 3$ binnings with randomly sampled symmetric bin edges $\{-\pi, -\Delta\chi, \Delta\chi, \pi\} \times \{0, \pi/2 - \Delta\theta_\ell, \pi/2 + \Delta\theta_\ell, \pi\} \times \{0, \pi/2 - \Delta\theta_{D^*}, \pi/2 + \Delta\theta_{D^*}, \pi\}$. For a detailed description of the plotting principles of (a)–(c), see figure 2.9. Since $\Delta\chi$, $\Delta\theta_\ell$ and $\Delta\theta_{D^*}$ are sampled from uniform distributions, equally spaced bins in $\Delta\chi$, $\Delta\theta_\ell$ and $\Delta\theta_{D^*}$ are used for plots (a)–(c). For a detailed description of the plotting principles of (d), see figure 2.9, though we use a scatter plot rather than a density plot here.

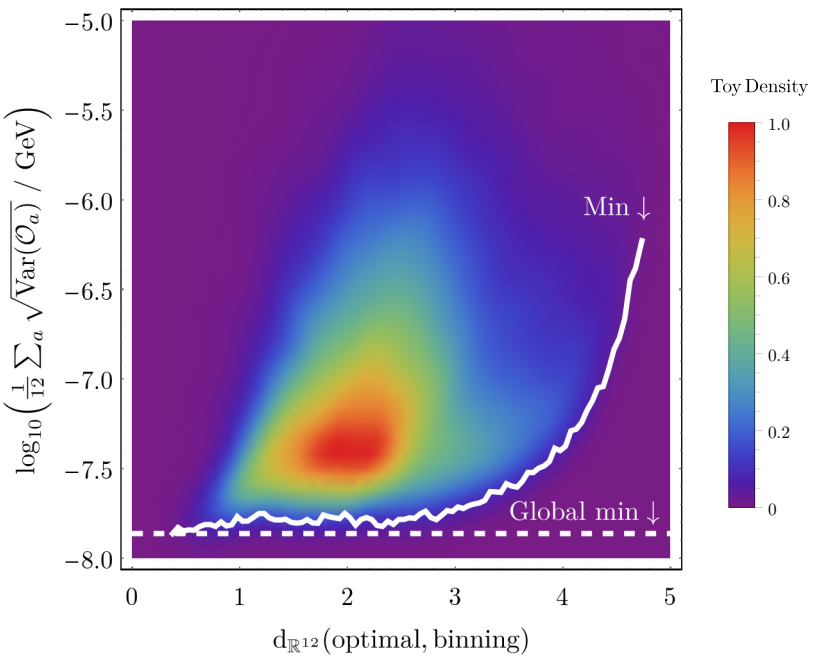


Figure 2.20: Probing for the optimal $3 \times 3 \times 3$ bin edges without the symmetry requirement. See figure 2.10 for a detailed description of the plotting principles. Here $\mathrm{d}_{\mathbb{R}^{12}}(\mathrm{optimal}, \mathrm{binning})$ is the minimal $\mathbb{R}^{12}$ distance between the vector of the joint bin edges, $(-\pi, \chi_1, \chi_2, \pi, 0, \theta_{\ell 1}, \theta_{\ell 2}, \pi, 0, \theta_{D^*1}, \theta_{D^*2}, \pi)$ to the corresponding vectors given by the symmetric binnings $\Delta\theta_\ell = \Delta\theta_{D^*} = \pi/3$ and $\Delta\chi = \pi/3$ or $\Delta\chi = 2\pi/3$.

Note however, that the "combinatorial information" from table 2.5 is still the basis of this method, only our strategy to find the corresponding binning and weights differs.

2.3.5 Comparison of different strategies

A comparison of selected results of the previous sections is shown in figure 2.21.

While we are mostly interested in binnings with a small number of bins, the binning $9 \times 9 \times 9$ (with non-independent product form weights) is given for comparison: The corresponding optimal observables showed the lowest error expectations we have encountered so far.

Technically, the case of 12 general bins shows the lowest number of bins required to construct all $\mathcal{O}_a$. It is surprising, that even our rough sampling of the high number of degrees of freedom gave rise to an observable whose sensitivity is in the same order of magnitude as others with significantly more bins (and it is not clear to what extent this result might still be improved by more thorough studies or by simply increasing the statistics). Nonetheless, as mentioned before, this result is so far more of mathematical interest than of practical use.

The lowest number of product bins to construct all $\mathcal{O}_a$ is $3 \times 3 \times 2$ (non-prod. weights). The performance is particularly good for $\mathcal{O}_3^T$ to $\mathcal{O}_4^{0T}$, but can be significantly improved for $\mathcal{O}_1^0$ to $\mathcal{O}_2^T$ by adding another bin in θ_{D^*} and considering a $3 \times 3 \times 3$ binning. The influence of the bin edges has been investigated and the optimal binning is given in equation (2.3.15).

When considering product weights (section 2.3.4), the minimal number of required bins to construct *some* (but not all) $\mathcal{O}_a$ can be even lower. This is shown in table 2.10. The performance of the solution for $1 \times 3 \times 3$ bins is shown in figure 2.21. This binning allows to extract W_1^0, W_2^0, W_3^0, W_1^T and W_3^T with surprising precision, which is only bested by the $9 \times 9 \times 9$ binning (and only by at most 30%).

$\mathcal{O}_a$	$\mathcal{O}_1^0$, $\mathcal{O}_2^0$, $\mathcal{O}_3^0$, $\mathcal{O}_1^T$, $\mathcal{O}_3^T$	$\mathcal{O}_2^T$	$\mathcal{O}_4^T$, $\mathcal{O}_5^T$	$\mathcal{O}_1^{0T}$, $\mathcal{O}_2^{0T}$, $\mathcal{O}_3^{0T}$, $\mathcal{O}_4^{0T}$
Minimal # bins	$1 \times 3 \times 3$	$3 \times 3 \times 2^*$	$3 \times 1 \times 2$	$2 \times 2 \times 3$
	$1 \times 5 \times 2$	$3 \times 3 \times 3$	$3 \times 3 \times 1$	$2 \times 5 \times 1$
	$3 \times 3 \times 2$	$3 \times 5 \times 2$	$5 \times 1 \times 1$	$3 \times 2 \times 2$
		$5 \times 3 \times 2$		$3 \times 3 \times 1$
				$5 \times 2 \times 1$

Table 2.10: Minimal number of product bins (with product weights) required to construct each observable. For the observable $\mathcal{O}_2^T$, solutions with product weights require more bins than the general solution for product bins with non-product weights, $3 \times 3 \times 2$ (also included as reference and denoted with an asterisk).

For $3\times3\times3$ bins it is notable that the solutions for non-product weights and for (non-independent) product weights have comparable sensitivity. Non-product weights are conceptually less complex to calculate, but cannot be visualized as in figure 2.8, thus the former case is slightly more interesting. For both cases, an equidistant spacing of the bins in χ, θ_ℓ and θ_{D^*} was found to minimize the expected errors. Generally, the sensitivity seems to be close to optimal for all W_a but for W_4^T and W_5^T, where the expected uncertainty can still be improved by about a factor of two e.g. by $5 \times 5 \times 3$ bins (non-indep. prod. weights).

The expected relative uncertainty is compared with the relative uncertainty on the total integrated cross section (as calculated in section 2.2.1) in figure 2.21b. The relative uncertainty for our observables compared to the uncertainty on the total cross section is found to be about 5 to 15 times (5 to 25 times) higher for a $3 \times 3 \times 3$ ($9 \times 9 \times 9$) binning.

2.4 Dependency of the Amplitudes on NP Coupling Constants

Having constructed observables $\mathcal{O}_a$, which map experimental measurements of the cross sections to measurements of the q^2-integrated coefficient $\overline{W}_a$, we have extracted all information contained in the angle dependency of the differential cross section. In order to use this information to constrain NP contributions, the next step is to determine how the $\overline{W}_a$ depend on the NP coupling constants. W_a can be parametrized as

$$W_a(q^2) = \sum_{i=1}^{13} c_i \, W_a^{(i)}(q^2) \quad \Longrightarrow \quad \overline{W}_a = \sum_{i=1}^{13} c_i \, \overline{W}_a^{(i)}, \tag{2.4.1}$$

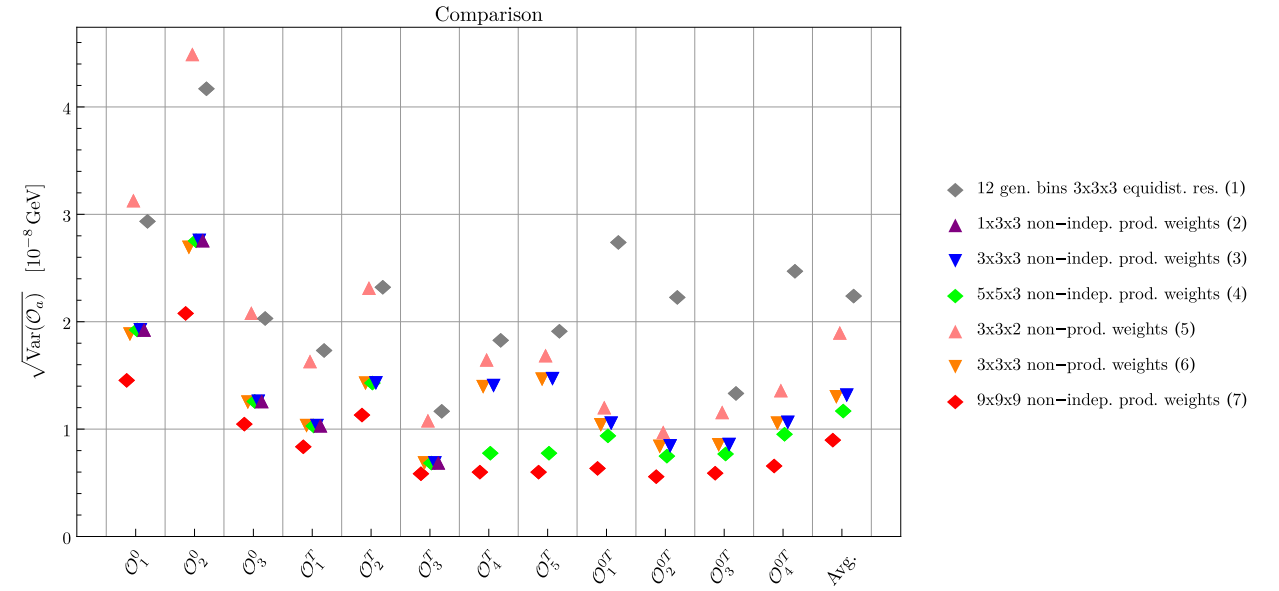
where the 13 terms c_i are linearly independent quadratic combinations of the NP coupling constants g_V, g_A, g_S, g_P and T_L ($g_V = g_A = g_S = g_P = T_L = 0$ for SM) as given in table 2.11 and defined via the Lagrangian below. Table 2.12 shows the numeric values of $\overline{W}_a^{(i)}$.

Written in terms of $\mathcal{C}_{V_L}$, $\mathcal{C}_{V_R}$, $\mathcal{C}_{S_L}$, $\mathcal{C}_{S_R}$, $\mathcal{C}_T$ via

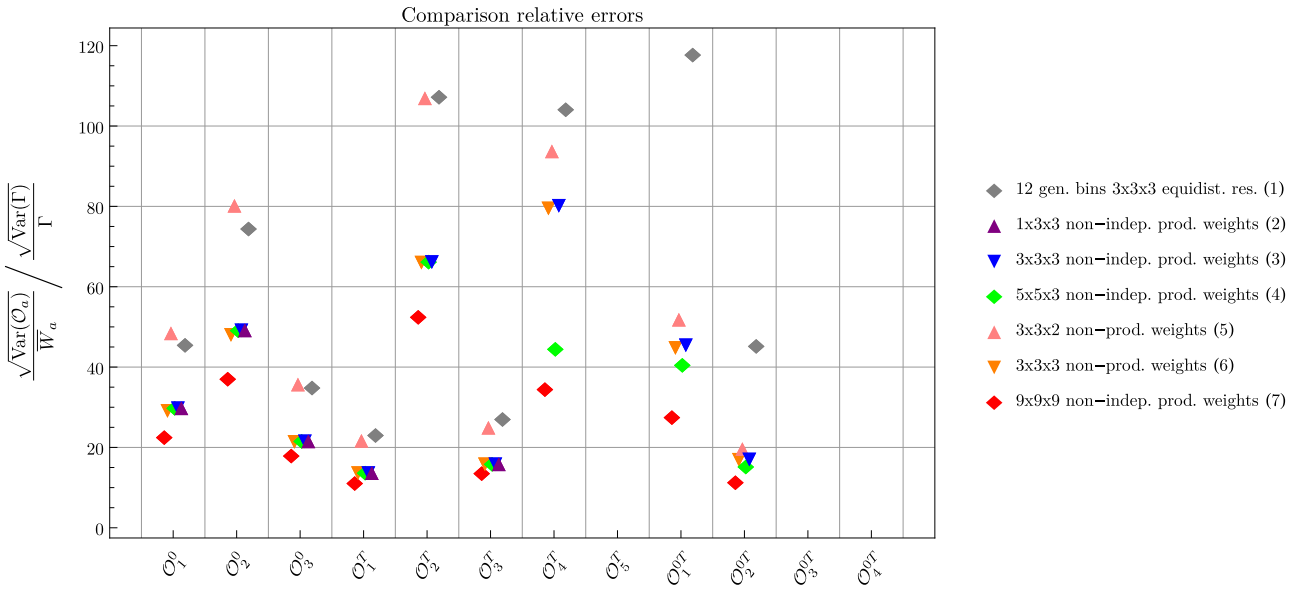
$$g_{V,A} = \mathcal{C}_{V_R} \pm \mathcal{C}_{V_L}, \qquad g_{S,P} = \mathcal{C}_{S_R} \pm \mathcal{C}_{S_L}, \qquad T_L = \mathcal{C}_T, \tag{2.4.2}$$

($\mathcal{C}_{V_{L/R}} = \mathcal{C}_{S_{L/R}} = \mathcal{C}_T = 0$ for the SM), the considered effective Lagrangian is given by [116]:

$$\begin{aligned} \mathcal{H}_{\text{eff}} = \frac{4G_F V_{cb}}{\sqrt{2}} \Big[& (1 + \mathcal{C}_{V_L}[\bar{c}\gamma_\mu P_L b][\bar{l}\gamma^\mu P_L \nu_l] + \mathcal{C}_{V_R}[\bar{c}\gamma^\mu P_R b][\bar{l}\gamma_\mu P_L \nu_l] + \\ & + \mathcal{C}_{S_L}[\bar{c} P_L b][\bar{l} P_L \nu_l] + \mathcal{C}_{S_R}[\bar{c} P_R b][\bar{l} P_L \nu_l] + T_L[\bar{c}\sigma^{\mu\nu} P_L b][\bar{l}\sigma_{\mu\nu} P_L \nu_l] \Big], \end{aligned} \tag{2.4.3}$$



(a) Absolute errors



(b) Relative errors

Figure 2.21: Comparing the sensitivities of observables derived based on different assumptions. Result (1) is calculated as described in section 2.3.1 (our best result for 12 bins, found with $3 \times 3 \times 3$ cubes of fixed size $2\pi/3 \times \pi/3 \times \pi/3$), results (5) and (6) as in section 2.3.2 (equidistant binning for (6), binning (2.3.15) for (5)) and results (2), (3), (4) and (7) as in section 2.3.3 (non-independent product bins, equidistant binnings). Figure 2.21a shows absolute errors, figure 2.21b relative errors in comparison to the estimated statistical error of the total integrated cross section as calculated in section 2.2.1 (for all observables with non-vanishing SM expectation value).

Convention 3	Convention 2	Convention 1	SM
c_1	$\lvert 1-g_A\rvert^2$	$\lvert\mathcal{C}_{V_L}-\mathcal{C}_{V_R}+1\rvert^2$	**1**
c_2	$\lvert g_V+1\rvert^2$	$\lvert\mathcal{C}_{V_L}+\mathcal{C}_{V_R}+1\rvert^2$	**1**
c_3	$\lvert g_P\rvert^2$	$\lvert\mathcal{C}_{S_L}-\mathcal{C}_{S_R}\rvert^2$	0
c_4	$\lvert T_L\rvert^2$	$\lvert\mathcal{C}_T\rvert^2$	0
c_5	$\mathrm{Re}((1-g_A)((g_V)^*+1))$	$\mathrm{Re}((\mathcal{C}_{V_L}-\mathcal{C}_{V_R}+1)((\mathcal{C}_{V_L}+\mathcal{C}_{V_R})^*+1))$	**1**
c_6	$\mathrm{Re}((1-g_A)(g_P)^*)$	$\mathrm{Re}((\mathcal{C}_{V_L}-\mathcal{C}_{V_R}+1)((\mathcal{C}_{S_R})^*-(\mathcal{C}_{S_L})^*))$	0
c_7	$\mathrm{Re}((1-g_A)(T_L)^*)$	$\mathrm{Re}((\mathcal{C}_T)^*(\mathcal{C}_{V_L}-\mathcal{C}_{V_R}+1))$	0
c_8	$\mathrm{Re}((g_V+1)(T_L)^*)$	$\mathrm{Re}((\mathcal{C}_T)^*(\mathcal{C}_{V_L}+\mathcal{C}_{V_R}+1))$	0
c_9	$\mathrm{Re}(g_P(T_L)^*)$	$\mathrm{Re}((\mathcal{C}_T)^*(\mathcal{C}_{S_R}-\mathcal{C}_{S_L}))$	0
c_{10}	$\mathrm{Im}((1-g_A)((g_V)^*+1))$	$\mathrm{Im}((\mathcal{C}_{V_L}-\mathcal{C}_{V_R}+1)((\mathcal{C}_{V_L}+\mathcal{C}_{V_R})^*+1))$	0
c_{11}	$\mathrm{Im}((1-g_A)(T_L)^*)$	$\mathrm{Im}((\mathcal{C}_T)^*(\mathcal{C}_{V_L}-\mathcal{C}_{V_R}+1))$	0
c_{12}	$\mathrm{Im}((g_V+1)(g_P)^*)$	$\mathrm{Im}((\mathcal{C}_{V_L}+\mathcal{C}_{V_R}+1)((\mathcal{C}_{S_R})^*-(\mathcal{C}_{S_L})^*))$	0
c_{13}	$\mathrm{Im}(g_P(T_L)^*)$	$\mathrm{Im}((\mathcal{C}_T)^*(\mathcal{C}_{S_R}-\mathcal{C}_{S_L}))$	0

Table 2.11: Expressing the c_i in terms of coupling constants

where G_F is the Fermi coupling constant, V_{cb} the relevant element of the Cabibbo-Kobayashi-Maskawa (CKM) matrix, $P_{L/R} := (1 \mp \gamma_5)/2$ the projectors on left/right-handed eigenstates and $\sigma_{\mu\nu} := \mathrm{i}(\gamma_\mu\gamma_\nu - \gamma_\nu\gamma_\mu)/2$. Interaction terms for right-handed neutrinos are not included.

Note that, while the c_i are *linearly* independent, there are at most eight real degrees of freedom in total for $\bar{B}\longrightarrow D^* l^- \bar{\nu}_l$ (real and imaginary parts of g_V, g_A, g_P and T_L, as g_S does not contribute), thus they *do* form an overcomplete set of parameters. In fact, the absolute values c_1, c_2, c_3 and c_4 together with the relative phases $\angle(1-g_A, g_V^*+1)$, $\angle(g_V+1, g_P^*)$ and $\angle(g_P, T_L^*)$ are enough to determine all 13 c_i, thus there are even only seven real degrees of freedom. Similarly one could also consider relations such as $c_7^2 + c_{11}^2 = c_1 c_4$ etc. to come to a similar conclusion.

	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9	c_{10}	c_{11}	c_{12}	c_{13}
$\overline{\mathbf{W}}_\mathbf{1}^\mathbf{0}$	1.028(53)			−12.86(69)									
$\overline{\mathbf{W}}_\mathbf{2}^\mathbf{0}$	1.028(53)			−12.93(66)									
$\overline{W}_3^0$	−1.060(64)					−0.739(45)	5.34(35)		3.87(25)				
$\overline{\mathbf{W}}_\mathbf{1}^\mathbf{T}$	1.234(63)	0.1399(97)		14.39(77)			−6.68(34)	1.60(11)					
$\overline{\mathbf{W}}_\mathbf{2}^\mathbf{T}$	−0.361(18)	−0.0374(26)		8.64(46)									
$\overline{W}_3^T$				−13.13(74)	0.794(44)		4.30(24)	−2.29(13)					
$\overline{W}_4^T$	−0.361(18)	0.0374(26)		3.64(22)									
$\overline{W}_5^T$										−0.222(12)			
$\overline{W}_1^{0T}$	0.428(22)			−6.28(33)									
$\overline{W}_2^{0T}$	0.403(24)			−6.62(38)	0.498(28)	0.290(17)	0.144(95)	−1.163(67)	−1.90(11)				
$\overline{W}_3^{0T}$										0.1537(97)	−1.12(11)	−0.1049(66)	−1.289(81)
$\overline{W}_4^{0T}$										0.1357(76)			

(a) $\ell = \tau$

	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9	c_{10}	c_{11}	c_{12}	c_{13}
$\overline{\mathbf{W}}_\mathbf{1}^\mathbf{0}$			1.010(72)	111.0(6.6)									
$\overline{\mathbf{W}}_\mathbf{2}^\mathbf{0}$	12.61(70)			−111.0(6.6)									
$\overline{W}_3^0$									20.9(1.3)				
$\overline{\mathbf{W}}_\mathbf{1}^\mathbf{T}$	4.96(26)	0.840(59)											
$\overline{\mathbf{W}}_\mathbf{2}^\mathbf{T}$	−2.48(13)	−0.420(30)		129.0(7.4)									
$\overline{W}_3^T$					3.91(22)								
$\overline{W}_4^T$	−2.48(13)	0.420(30)		14.1(2.8)									
$\overline{W}_5^T$										−1.96(11)			
$\overline{W}_1^{0T}$	3.68(19)			−60.9(3.3)									
$\overline{W}_2^{0T}$					3.14(18)				−11.62(67)				
$\overline{W}_3^{0T}$													−9.95(61)
$\overline{W}_4^{0T}$									1.570(91)				

(b) $m_l = 0$ ($\ell = e, \mu$)

Table 2.12: Values $\overline{W}_a^{(i)}$. All numbers are understood times 10^{-16} GeV. Bold $\overline{W}_a$ contribute to the total integrated cross section.

		$(1,3)_0$	$(1,2)_{\frac{1}{2}}$	$(3,3)_{\frac{2}{3}}$	$(3,1)_{\frac{2}{3}}$	$(3,2)_{\frac{7}{6}}$	$(\bar{3},2)_{\frac{5}{6}}$	$(\bar{3},3)_{\frac{1}{3}}$	$(\bar{3},1)_{\frac{1}{3}}$
$\frac{1}{2}(g_A+g_V)$	$\mathcal{C}_{V_R}$								
$\frac{1}{2}(g_V-g_A)$	$\mathcal{C}_{V_L}$	$\mathcal{C}_{V_L}$		$\mathcal{C}_{V_L}$	$\mathcal{C}_{V_L}$			$\mathcal{C}_{V_L}$	$\mathcal{C}_{V_L}$
$\frac{1}{2}(g_P+g_S)$	$\mathcal{C}_{S_R}$		$\mathcal{C}_{S_R}$		$\mathcal{C}_{S_R}$		$\mathcal{C}_{S_R}$		
$\frac{1}{2}(g_S-g_P)$	$\mathcal{C}_{S_L}$		$\mathcal{C}_{S_L}$			$\mathcal{C}_{S_L}$			$\mathcal{C}_{S_L}$
T_L	$\mathcal{C}_T$					$\frac{\mathcal{C}_{S_L}}{4}$			$-\frac{\mathcal{C}_{S_L}}{4}$
$\mathcal{C}_{V_L}+\mathcal{C}_{V_R}$	g_V	g_V		g_V	g_V			g_V	g_V
$\mathcal{C}_{V_R}-\mathcal{C}_{V_L}$	g_A	$-g_V$		$-g_V$	$-g_V$			$-g_V$	$-g_V$
$\mathcal{C}_{S_L}+\mathcal{C}_{S_R}$	g_S		g_S		g_S	g_S	g_S		g_S
$\mathcal{C}_{S_R}-\mathcal{C}_{S_L}$	g_P		g_P		g_S	$-g_S$	g_S		$-g_S$
$\mathcal{C}_T$	T_L					$\frac{g_S}{4}$			$-\frac{g_S}{4}$
$\lvert 1-g_A\rvert^2$	c_1	c_1	1	c_1	c_1	1	1	c_1	c_1
$\lvert g_V+1\rvert^2$	c_2	c_1	1	c_1	c_1	1	1	c_1	c_1
$\lvert g_P\rvert^2$	c_3		c_3		c_3	c_3	c_3		c_3
$\lvert T_L\rvert^2$	c_4					$\frac{c_3}{16}$			$\frac{c_3}{16}$
$\mathrm{Re}((1-g_A)((g_V)^*+1))$	c_5	c_1	1	c_1	c_1	1	1	c_1	c_1
$\mathrm{Re}((1-g_A)(g_P)^*)$	c_6		c_6		c_6	c_6	c_6		c_6
$\mathrm{Re}((1-g_A)(T_L)^*)$	c_7					$-\frac{c_6}{4}$			$\frac{c_6}{4}$
$\mathrm{Re}((g_V+1)(T_L)^*)$	c_8					$-\frac{c_6}{4}$			$\frac{c_6}{4}$
$\mathrm{Re}(g_P(T_L)^*)$	c_9					$-\frac{c_3}{4}$			$\frac{c_3}{4}$
$\mathrm{Im}((1-g_A)((g_V)^*+1))$	c_{10}								
$\mathrm{Im}((1-g_A)(T_L)^*)$	c_{11}					c_{11}			c_{11}
$\mathrm{Im}((g_V+1)(g_P)^*)$	c_{12}		c_{12}		c_{12}	$-4c_{11}$	c_{12}		$4\ c_{11}$
$\mathrm{Im}(g_P(T_L)^*)$	c_{13}								

Table 2.13: Couplings of the considered operators expressed in different conventions. Whenever two coupling constants (or c_i) are constrained to be equal for a particular operator, we simply use the name of the first to describe both. For example, we write c_1 for c_2 and c_5 when considering the operator $(\mathbf{1},\mathbf{3})_0$.

	SM	All NP	$(\mathbf{1},\mathbf{3})_0$, $(\mathbf{3},\mathbf{3})_{\frac{2}{3}}$, $(\bar{\mathbf{3}},\mathbf{3})_{\frac{1}{3}}$	$(\mathbf{1},\mathbf{2})_{\frac{1}{2}}$, $(\bar{\mathbf{3}},\mathbf{2})_{\frac{5}{6}}$	$(\mathbf{3},\mathbf{1})_{\frac{2}{3}}$	$(\mathbf{3},\mathbf{2})_{\frac{7}{6}}$	$(\bar{\mathbf{3}},\mathbf{1})_{\frac{1}{3}}$
$\mathbf{W_1^0}$		c_1, c_3, c_4, c_6, c_7	c_1	c_3, c_6	c_1, c_3, c_6	c_3, c_6	c_1, c_3, c_6
$\mathbf{W_2^0}$		c_1, c_4	c_1		c_1	c_3	c_1, c_3
W_3^0		c_1, c_6, c_7, c_9	c_1	c_6	c_1, c_6	c_6, c_3	c_1, c_6, c_3
$\mathbf{W_1^T}$		c_1, c_2, c_4, c_7, c_8	c_1		c_1	c_3, c_6	c_1, c_3, c_6
$\mathbf{W_2^T}$		c_1, c_2, c_4	c_1		c_1	c_3	c_1, c_3
W_3^T		c_4, c_5, c_7, c_8	c_1		c_1	c_3, c_6	c_3, c_1, c_6
W_4^T		c_1, c_2, c_4	c_1		c_1	c_3	c_1, c_3
W_5^T		c_{10}					
W_1^{0T}		c_1, c_4	c_1		c_1	c_3	c_1, c_3
W_2^{0T}		$c_1, c_4, c_5, c_6, c_7, c_8, c_9$	c_1	c_6	c_1, c_6	c_3, c_6	c_1, c_3, c_6
W_3^{0T}		$c_{10}, c_{11}, c_{12}, c_{13}$		c_{12}	c_{12}	c_{11}	c_{11}
W_4^{0T}		c_{10}					

(a) $\ell = \tau$

	SM	All NP	$(\mathbf{1},\mathbf{3})_0$, $(\mathbf{3},\mathbf{3})_{\frac{2}{3}}$, $(\bar{\mathbf{3}},\mathbf{3})_{\frac{1}{3}}$	$(\mathbf{1},\mathbf{2})_{\frac{1}{2}}$, $(\bar{\mathbf{3}},\mathbf{2})_{\frac{5}{6}}$	$(\mathbf{3},\mathbf{1})_{\frac{2}{3}}$	$(\mathbf{3},\mathbf{2})_{\frac{7}{6}}$	$(\bar{\mathbf{3}},\mathbf{1})_{\frac{1}{3}}$
$\mathbf{W_1^0}$		c_3, c_4		c_3	c_3	c_3	c_3
$\mathbf{W_2^0}$		c_1, c_4	c_1		c_1	c_3	c_1, c_3
W_3^0		c_9				c_3	c_3
$\mathbf{W_1^T}$		c_1, c_2	c_1		c_1		c_1
$\mathbf{W_2^T}$		c_1, c_2, c_4	c_1		c_1	c_3	c_1, c_3
W_3^T		c_5	c_1		c_1		c_1
W_4^T		c_1, c_2, c_4	c_1		c_1	c_3	c_1, c_3
W_5^T		c_{10}					
W_1^{0T}		c_1, c_4	c_1		c_1	c_3	c_1, c_3
W_2^{0T}		c_5, c_9	c_1		c_1	c_3	c_1, c_3
W_3^{0T}		c_{13}					
W_4^{0T}		c_{10}					

(b) Assuming vanishing lepton mass ($\ell = e, \mu$)

Table 2.14: Coefficients c_i relevant for W_a when considering a particular operator. As in table 2.13, only parameters linearly independent for the considered operator are distinguished, i.e. whenever $c_a \propto c_b$, only c_a is mentioned. The cells of the table are colored white, if $W_a = 0$ for a particular operator, gray, if $W_a \neq 0$ but W_a does not depend on NP couplings, green if $W_a \propto c_i$ (exactly one c_i involved), yellow if $W_a \propto \text{const} + c_i$ (affine linearly related, but not directly proportional to exactly one c_i) and orange for a dependency on more than one c_i. Bold W_a contribute to the total integrated cross section.

	$(\mathbf{1},\mathbf{3})_0$, $(\mathbf{3},\mathbf{3})_{\frac{2}{3}}$, $(\bar{\mathbf{3}},\mathbf{3})_{\frac{1}{3}}$	$(\mathbf{1},\mathbf{2})_{\frac{1}{2}}$, $(\bar{\mathbf{3}},\mathbf{2})_{\frac{5}{6}}$	$(\mathbf{3},\mathbf{1})_{\frac{2}{3}}$	$(\mathbf{3},\mathbf{2})_{\frac{7}{6}}$	$(\bar{\mathbf{3}},\mathbf{1})_{\frac{1}{3}}$
$\overline{\mathbf{W_1^0}}$	$1.18\,c_1$	$0.172\,c_3 + 0.489\,c_6 + 1.18$	$1.18\,c_1 + 0.172\,c_3 + 0.489\,c_6$	$1.59\,c_3 + 2.61\,c_6 + 1.18$	$1.18\,c_1 + 1.59\,c_3 - 1.63\,c_6$
$\overline{\mathbf{W_2^0}}$	$1.03\,c_1$	1.03	$1.03\,c_1$	$1.03 - 0.804\,c_3$	$1.03\,c_1 - 0.804\,c_3$
$\overline{W}_3^0$	$-1.06\,c_1$	$-0.739\,c_6 - 1.06$	$-1.06\,c_1 - 0.739\,c_6$	$-0.967\,c_3 - 2.07\,c_6 - 1.06$	$-1.06\,c_1 + 0.967\,c_3 + 0.597\,c_6$
$\overline{\mathbf{W_1^T}}$	$1.37\,c_1$	1.37	$1.37\,c_1$	$0.899\,c_3 + 1.27\,c_6 + 1.37$	$1.37\,c_1 + 0.899\,c_3 - 1.27\,c_6$
$\overline{\mathbf{W_2^T}}$	$-0.399\,c_1$	-0.399	$-0.399\,c_1$	$0.54\,c_3 - 0.399$	$0.54\,c_3 - 0.399\,c_1$
$\overline{W}_3^T$	$0.794\,c_1$	0.794	$0.794\,c_1$	$-0.821\,c_3 - 0.503\,c_6 + 0.794$	$0.794\,c_1 - 0.821\,c_3 + 0.503\,c_6$
$\overline{W}_4^T$	$-0.324\,c_1$	-0.324	$-0.324\,c_1$	$0.228\,c_3 - 0.324$	$0.228\,c_3 - 0.324\,c_1$
$\overline{W}_5^T$					
$\overline{W}_1^{0T}$	$0.428\,c_1$	0.428	$0.428\,c_1$	$0.428 - 0.393\,c_3$	$0.428\,c_1 - 0.393\,c_3$
$\overline{W}_2^{0T}$	$0.902\,c_1$	$0.29\,c_6 + 0.902$	$0.902\,c_1 + 0.29\,c_6$	$0.0607\,c_3 + 0.545\,c_6 + 0.902$	$0.902\,c_1 - 0.889\,c_3 + 0.0356\,c_6$
$\overline{W}_3^{0T}$		$-0.105\,c_{12}$	$-0.105\,c_{12}$	$-0.703\,c_{11}$	$-1.54\,c_{11}$
$\overline{W}_4^{0T}$					

(a) $\ell = \tau$

	$(\mathbf{1},\mathbf{3})_0$, $(\mathbf{3},\mathbf{3})_{\frac{2}{3}}$, $(\bar{\mathbf{3}},\mathbf{3})_{\frac{1}{3}}$	$(\mathbf{1},\mathbf{2})_{\frac{1}{2}}$, $(\bar{\mathbf{3}},\mathbf{2})_{\frac{5}{6}}$	$(\mathbf{3},\mathbf{1})_{\frac{2}{3}}$	$(\mathbf{3},\mathbf{2})_{\frac{7}{6}}$	$(\bar{\mathbf{3}},\mathbf{1})_{\frac{1}{3}}$
$\overline{\mathbf{W_1^0}}$		$1.01\,c_3$	$1.01\,c_3$	$7.93\,c_3$	$7.93\,c_3$
$\overline{\mathbf{W_2^0}}$	$12.6\,c_1$	12.6	$12.6\,c_1$	$12.6 - 6.92\,c_3$	$12.6\,c_1 - 6.92\,c_3$
$\overline{W}_3^0$				$-5.23\,c_3$	$5.23\,c_3$
$\overline{\mathbf{W_1^T}}$	$5.8\,c_1$	5.8	$5.8\,c_1$	5.8	$5.8\,c_1$
$\overline{\mathbf{W_2^T}}$	$-2.9\,c_1$	-2.9	$-2.9\,c_1$	$8.07\,c_3 - 2.9$	$8.07\,c_3 - 2.9\,c_1$
$\overline{W}_3^T$	$3.91\,c_1$	3.91	$3.91\,c_1$	3.91	$3.91\,c_1$
$\overline{W}_4^T$	$-2.06\,c_1$	-2.06	$-2.06\,c_1$	$0.879\,c_3 - 2.06$	$0.879\,c_3 - 2.06\,c_1$
$\overline{W}_5^T$					
$\overline{W}_1^{0T}$	$3.68\,c_1$	3.68	$3.68\,c_1$	$3.68 - 3.82\,c_3$	$3.68\,c_1 - 3.82\,c_3$
$\overline{W}_2^{0T}$	$3.14\,c_1$	3.14	$3.14\,c_1$	$2.91\,c_3 + 3.14$	$3.14\,c_1 - 2.91\,c_3$
$\overline{W}_3^{0T}$					
$\overline{W}_4^{0T}$					

(b) Assuming vanishing lepton mass ($\ell = e, \mu$)

Table 2.15: Expressions for $W_a = \sum_a c_i \overline{W}_a^{(i)}$. The color scheme is the same as in 2.14. All values are given in units of 10^{-16} GeV. Bold $\overline{W}_a$ contribute to the total integrated cross section

2.5 Operator Base

The effect of the seven degrees of freedom is still hard to investigate in full generality. Prior studies such as [117] have therefore oftentimes considered the effects of one coupling constant g_A, g_V, g_P or T_L at a time, while assuming all others to vanish.

However, a better-motivated approach is to consider a base of operators that correspond to actually possible mediator particles in the decay. Therefore, we consider all operators of dimension six with non-flavor universal gauge-invariant couplings relevant to $\bar{B}\longrightarrow D^* l^- \bar{\nu}_l$ as presented in [22].[8] Note that, in particular, the right-handed vector coupling is flavor-universal and thus not included. The coupling constants relevant for each operator are shown in table 2.13, the sums $\sum_a c_i \overline{W}_a^{(i)}$ for each operator in table 2.14 and 2.15.

Several of the operators are indistinguishable in our analysis, namely $(\mathbf{1},\mathbf{3})_0$, $(\mathbf{3},\mathbf{3})_{2/3}$, and $(\bar{\mathbf{3}},\mathbf{2})_{5/3}$, as well as $(\mathbf{1},\mathbf{2})_{1/2}$ and $(\bar{\mathbf{3}},\mathbf{2})_{5/3}$ (see sections 2.5.2 and 2.5.3).

It should also be noted that the coupling constants can generally be different for different lepton flavors, in particular between tauons and light leptons. To distinguish them, coupling constants to light leptons will be primed, whenever a distinction is necessary. However, it is usually assumed that NP contributions to the light modes are negligible.

Furthermore, the coupling constants only obey the relations shown in table 2.13 at the scale of the integrated out mediator particles and (because the coupling constants run differently) start to diverge from these relations for other scales. Since no quantitative calculations are presented here, this does not need to be taken into account at this point.

Special attention is paid to $R(D^*)$ as it is the most prominent measured observable so far and can be easily described in the framework of the observables as considered in this chapter. Other measurements such as $\mathrm{d}\Gamma/\,\mathrm{d}q^2$ or measurements of the τ asymmetry are not discussed in detail due to our focus on angular measurements.

The best way to distinguish between the operators is of course to construct a complete set of observables $\mathcal{O}_a$ as described in sections 2.2 and 2.3, fit the degrees of freedom in each operator and finally compare the goodness of the fit to determine the most promising operator.

Nonetheless, we also mention some (combinations of) observables for each operator that have high discriminatory power without performing complete fits. This is not only relevant for cases where not all observables can be constructed, but can also be important in order to prioritize the construction or performance of some observables over the others.

In the following sections, we focus on the $\ell = \tau$ case, but section 2.5.7 quickly summarizes the (less complicated) $\ell = e, \mu$ case.

2.5.1 General NP

Without any further assumptions, the value of $R(D^*)$ is given by

$$\begin{aligned} R(D^*) &= 0.253 \times \frac{0.944\,c_1 + 0.056\,c_2 + 0.043\,c_3 + 13.347\,c_4 + 0.121\,c_6 - 5.353\,c_7 + 0.783\,c_8}{0.931\,c_1' + 0.0694\,c_2' + 0.063\,c_3' + 12.980\,c_4'} = \\ &= 0.253 \times \Big(0.944|1-g_A|^2 + 0.056|g_V+1|^2 + 0.043|g_P|^2 + 13.347|T_L|^2 + \\ &\quad + 0.121\,\mathrm{Re}((1-g_A)(g_P)^*) - 5.353\,\mathrm{Re}((1-g_A)(T_L)^*) + 0.783\,\mathrm{Re}(g_V+1)(T_L)^*\Big)/ \\ &\quad \Big(0.930621|1-g_A'|^2 + 0.063|g_P'|^2 + 0.069|g_V'+1|^2 + 12.980|T_L'|^2\Big). \end{aligned} \tag{2.5.1}$$

Note that c_{10} to c_{13} do not appear in $R(D^*)$, thus $R(D^*)$ is not influenced by CP averages in the measurements.

[8] Note that the operator $(\bar{\mathbf{3}},\mathbf{2})_{5/6}$ seems to be mistakenly denoted as $(\bar{\mathbf{3}},\mathbf{2})_{5/3}$ in [22], which is corrected here.

Looking at table 2.14, it is striking that without assuming any specific model, $W_5^T, W_4^{0T} \propto c_{10} = \mathrm{Im}((1-g_A)(g_V^*+1))$. However, for all operators that we consider, the couplings c_{10} vanishes. This is because $g_V = -g_A$, i.e. due to our choice of omitting the flavor-universal right-handed vector currents.

Thus, $\mathcal{O}_5^T$ and $\mathcal{O}_4^{0T}$ would be interesting as checks for CP violating right-handed vector currents, or as cross checks when focusing on the operators discussed here. Both $\mathcal{O}_5^T$ and $\mathcal{O}_4^{0T}$ are constructible with similar error expectations for $3\times3\times2$ or $3\times3\times3$ bins and also show little correlation to other observables. Note however, that c_{10} drops out in CP averages, thus both observables will vanish (regardless of NP) for CP averaged data.

Furthermore, note that c_2, c_4, c_5, c_7, c_8, c_9 do not appear for any of the operators, because they are either 1 or related to another c_i under the relations of each operator.

2.5.2 $(\mathbf{1},\mathbf{3})_0$, $(\mathbf{3},\mathbf{3})_{2/3}$ and $(\bar{\mathbf{3}},\mathbf{3})_{1/3}$

The operator $(\mathbf{1},\mathbf{3})_0$ can arise from a W' mediated interaction (section 1.3.3), $(\mathbf{3},\mathbf{2})_{2/3}$ and $(\bar{\mathbf{3}},\mathbf{3})_{1/3}$ from leptoquark interactions (section 1.3.4). For all three operators, $g_A = -g_V$ with all other couplings vanishing, leaving only the coupling coefficient $c_1 = c_2 = c_5 = |1-g_V|^2$. As c_1, c_2 and c_5 are also exactly the coupling constants that are non-vanishing for the SM, this means that these operators correspond to a global rescaling of the differential cross section: $\Gamma^{\mathrm{NP}} = \Gamma^{\mathrm{SM}} c_1$. Accordingly, the value of $R(D^*)$ will scale as

$$R(D^*) = 0.253 \cdot \frac{c_1}{c_1'} = 0.253 \cdot \frac{|1-g_V|^2}{|1-g_V'|^2}, \tag{2.5.2}$$

and (assuming $g_V' = 0$), the contour lines corresponding to fixed values of $R(D^*)$ will be circles around $(-1,0)$ in the $(\mathrm{Re}\, g_V, \mathrm{Im}\, g_V)$ plane (figure 2.22).

In fact, the same is true for all observables linear in the differential cross section: $\mathcal{O}(\Gamma^{\mathrm{NP}}) = \mathcal{O}(\Gamma^{\mathrm{SM}})c_1$. This also means that we can easily construct observables that stay *invariant* under contributions from these operators, by considering any ratio $\mathcal{O}/\mathcal{O}'$ of two such observables. In particular a ratio of $\mathcal{O}_1^0$, $\mathcal{O}_3^0$ or $\mathcal{O}_2^{0T}$ to any other observable could be used to single out this set of operators, as it is the only set for which the NP coupling constants will cancel (unless there is some fine tuning between c_i in the other models).

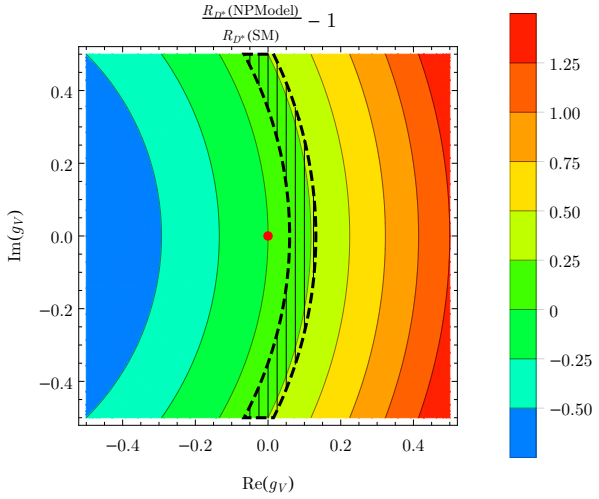
An interesting test for these three operators is the observable $\mathcal{O}_3^{0T}$, which is zero for these three operators but receives NP contributions for *all other* operators. Unfortunately these NP contributions are c_{11} and c_{12}, thus drop out in CP averages.

2.5.3 $(\mathbf{1},\mathbf{2})_{1/2}$ and $(\bar{\mathbf{3}},\mathbf{2})_{5/6}$

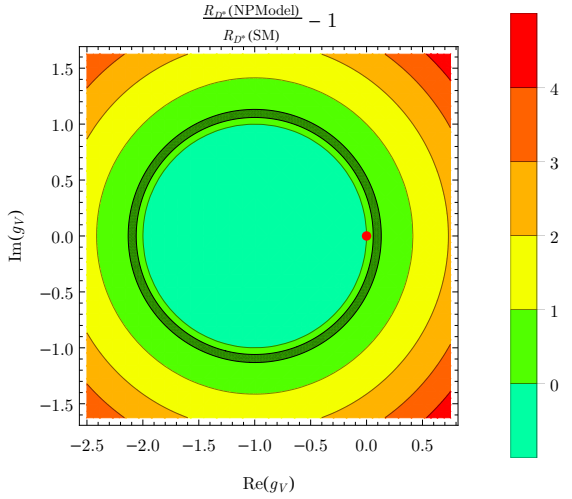
The operator $(\mathbf{1},\mathbf{2})_{1/2}$ can arise from interactions mediated by a charged Higgs boson (section 1.3.2), the operator $(\bar{\mathbf{3}},\mathbf{2})_{5/6}$ from leptoquark interactions. For $(\mathbf{1},\mathbf{2})_{1/2}$, $g_V = g_A = T_L = 0$. Consequently $c_1 = c_2 = c_5 = 1$, $c_3 = |g_P|^2$, $c_6 = \mathrm{Re}\, g_P$ and $c_{12} = \mathrm{Im}\, g_P$ and in particular $c_{12} = c_3^2 + c_6^2$. Note that g_S does not contribute to $\bar{B}\longrightarrow D^* l^- \bar{\nu}_l$. Thus, the operator $(\bar{\mathbf{3}},\mathbf{2})_{5/6}$ with the additional constraint $g_S = g_P$ can be treated exactly in the same way for $\bar{B}\longrightarrow D^*\ell^-\bar{\nu}_\ell$ (note however, that g_S *does* contribute to $\bar{B}\longrightarrow D l^-\bar{\nu}_l$, thus the operator models are distinguishable there).

The value of $R(D^*)$ is:

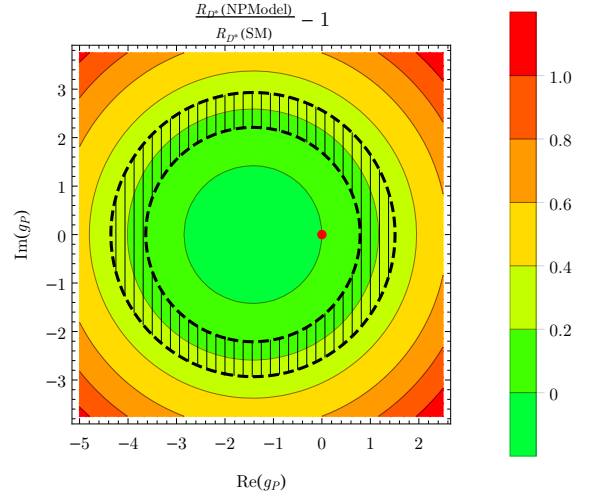
$$\begin{aligned} R(D^*) &= 0.253 \times \frac{1+0.043\,c_3+0.121\,c_6}{1+0.063\,c_3'} = 0.253 \times \frac{1+0.043\,|g_P|^2+0.121\,\mathrm{Re}\,g_P}{1+0.063|g'_P|^2} = \\ &= 0.253 \times \frac{0.289+0.043\,|1.423+g_P|^2}{1+0.063\,|\tilde{g}_P|^2}, \end{aligned} \tag{2.5.3}$$



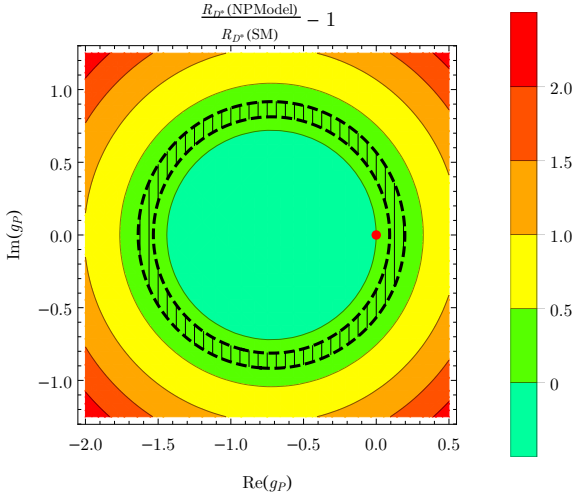
(a) $(\mathbf{1},\mathbf{3})_0$, $(\mathbf{3},\mathbf{3})_{2/3}$ and $(\bar{\mathbf{3}},\mathbf{3})_{1/3}$ (zoomed)



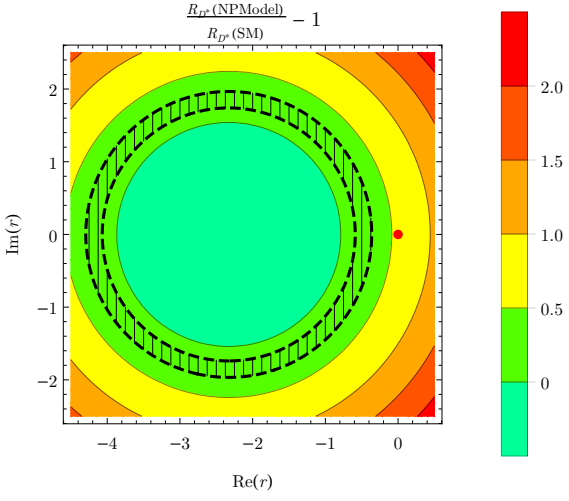
(b) $(\mathbf{1},\mathbf{3})_0$, $(\mathbf{3},\mathbf{3})_{2/3}$ and $(\bar{\mathbf{3}},\mathbf{3})_{1/3}$



(c) $(\mathbf{1},\mathbf{2})_{1/2}$, $(\mathbf{3},\mathbf{2})_{5/3}$ and upon replacing $g_P \longrightarrow \frac{g_P^*}{1+g_V}$ also for $(\mathbf{3},\mathbf{1})_{2/3}$ with $|1+g_V| = 1$.



(d) $(\mathbf{3},\mathbf{2})_{7/6}$



(e) $(\bar{\mathbf{3}},\mathbf{1})_{1/3}$ with $\tau = \frac{g_S^*}{1+g_V}$ at $|1+g_V| = 1$

Figure 2.22: $R(D^*)^{\text{NP}}/R(D^*)^{\text{SM}} - 1$ for different operators. The red dot marks the SM (corresponding to $g_A = g_V = g_S = g_P = T_L = 0$), the dashed or gray rings represent the current experimental value (with 1σ errors).

thus (for negligible light NP contributions) resulting in concentric circles around -1.423 as level lines in the $(\mathrm{Re}\, g_P, \mathrm{Im}\, g_P)$ plane.

The value of $c_{12} = \mathrm{Im}\, g_P$ is directly measurable via $\mathcal{O}_3^{0T}$ (which however drops out in CP averages).

In contrast to all other operators considered, $\mathcal{O}_2^0$, $\mathcal{O}_1^T$, $\mathcal{O}_2^T$, $\mathcal{O}_3^T$, $\mathcal{O}_4^5$ and $\mathcal{O}_1^{0T}$ strike out by being unmodified by NP contributions. Thus, confirming that any one of these is SM-like would point towards this operator,[9] while finding significant NP contributions would exclude this operator.

2.5.4 $(\mathbf{3},\mathbf{1})_{2/3}$

The operator $(\mathbf{3},\mathbf{1})_{2/3}$ can arise in leptoquark mediated interactions. It can be treated similar to the operator $(\mathbf{1},\mathbf{2})_{1/2}$ but introduces the additional couplings of $g_A = -g_V$ (which had been set to zero for $(\mathbf{1},\mathbf{2})_{1/2}$). This means $c_3 = |g_P|^2$, $c_6 = \mathrm{Re}((1+g_V)g_P^*)$, $c_{12} = \mathrm{Im}((g_V+1)g_P^*)$ and $c_5 = c_2 = c_1 = |1+g_V|^2$ (all others zero), reduced to three degrees of freedom via the relation $c_1 c_3 = c_6^2 + c_{12}^2$. As c_{12} does not appear in the fully integrated cross section, $R(D^*)$ takes a similar form to the $(\mathbf{1},\mathbf{2})_{1/2}$ result in terms of the c_i:

$$\begin{aligned} R(D^*) &= 0.253 \times \frac{c_1 + 0.043\, c_3 + 0.121\, c_6}{1 + 0.063\, c_3'} = \\ &= 0.253 \times \frac{|1+g_V|^2 + 0.043\, |g_P|^2 + 0.121\, \mathrm{Re}((1+g_V)g_P^*)}{|g'_V + 1|^2 + 0.063\, |g'_P|^2} \\ &= 0.253 \times \frac{0.289 + 0.043\, |1+g_V|^2 \left|1.423 + \frac{g_P^*}{1+g_V}\right|^2}{1 + 0.063\, |1+g'_V|^2 \left|\frac{g_P'^*}{1+g'_V}\right|^2}. \end{aligned} \tag{2.5.4}$$

Identifying $g_P \longrightarrow \frac{g_P'^*}{1+g'_V}$ and fixing $|1+g_V|^2 = 1$, this is the same relationship as for $(\mathbf{1},\mathbf{2})_{1/2}$ and $(\mathbf{3},\mathbf{2})_{5/3}$ and will have the same concentric circles around -1.423 as level lines. For $|1+g_V|^2 \neq 1$, only the value corresponding to each level line changes (or put differently, the level lines will increase or decrease in diameter for fixed function values).

To distinguish this operator from $(\mathbf{3},\mathbf{2})_{7/6}$ and $(\bar{\mathbf{3}},\mathbf{1})_{1/3}$, one could consider ratios between $\mathcal{O}_2^0$, $\mathcal{O}_1^T$, $\mathcal{O}_2^T$, $\mathcal{O}_3^T$, $\mathcal{O}_4^T$ and $\mathcal{O}_1^{0T}$ and verify that they are free of NP contributions. Finding that any of these observables by themselves have a NP contribution would distinguish this operator from $(\mathbf{1},\mathbf{2})_{1/2}$ and $(\bar{\mathbf{3}},\mathbf{2})_{5/3}$, Finally, to distinguish from this operator from $(\mathbf{1},\mathbf{3})_0$, $(\mathbf{3},\mathbf{3})_{2/3}$, and $(\bar{\mathbf{3}},\mathbf{2})_{5/3}$, directly measure a non-zero value of $c_{12} = -\,\mathrm{Im}\, g_P$ via $\mathcal{O}_3^{0T}$ (not possible for CP averaged data).

Finally, note that extracting the coupling constant $c_1 = |1+g_V|^2$ is very easy in this model.

2.5.5 $(\mathbf{3},\mathbf{2})_{7/6}$

Here, we have $g_V = g_A = 0$, $g_P = -g_S$, $T_L = -g_P/4$. Thus, $c_1 = c_2 = c_5 = 1$, $c_3 = |g_P|^2$, $c_4 = |g_P|^2/16$, $c_6 = \mathrm{Re}\, g_P$, $c_7 = c_8 = -g_P/4$, $c_9 = -|g_P|^2/4$, $c_{11} = g_P/4$ and $c_{12} = -4\,\mathrm{Im}\, g_P/4$. $R(D^*)$ is given by

$$\begin{aligned} R(D^*) &= 0.253 \times \frac{1 + 0.877\, c_3 + 1.264\, c_6}{1 + 0.875\, c_3'} = 0.253 \times \frac{1 + 0.877\, |g_P|^2 + 1.264\, \mathrm{Re}\, g_P}{1 + 0.875\, |g'_P|^2} = \\ &= 0.253 \times \frac{0.545 + 0.877\, |g_P + 0.721|^2}{1 + 0.875\, |g'_P|^2}. \end{aligned} \tag{2.5.5}$$

[9] Although in some cases contributions of different c_i to some W_a might cancel for $(\mathbf{3},\mathbf{2})_{7/6}$ and $(\bar{\mathbf{3}},\mathbf{1})_{1/3}$ if fine-tuned accordingly.

This operator can be distinguished from the first three classes of operators by considering ratios of observables and from the first class of operators by measuring a non-zero value of $c_{11} = \mathrm{Im}\, T_L = -g_P/4$. Distinguishing this operator from $(\bar{\mathbf{3}}, \mathbf{1})_{1/3}$ is discussed in the next section.

2.5.6 $(\bar{\mathbf{3}}, \mathbf{1})_{1/3}$

This operator (relevant for leptoquark models) can for the most part be treated similar to $(\mathbf{3}, \mathbf{2})_{7/6}$, but adds $g_A = -g_V$ and thereby notably introduces c_1 again. The condition $g_P = -g_S$ is the same as for $(\mathbf{3}, \mathbf{2})_{7/6}$, but T_L changes its sign to $T_L = g_P/4$. Thus, $c_1 = c_2 = c_5 = |1+g_V|^2$, $c_3 = |g_P|^2$, $c_4 = |g_P|^2/16$, $c_6 = \mathrm{Re}\, g_P$, $c_7 = c_8 = g_P/4$, $c_9 = |g_P|^2/4$, $c_{11} = -\mathrm{Im}\, g_P/4$ and $c_{12} = 4\,\mathrm{Im}\, g_P$. The value of $R(D^*)$ is:

$$\begin{aligned} R(D^*) &= 0.253 \times \frac{c_1 + 0.877\, c_3 + 1.021\, c_6}{c_1 + 0.875\, c_3'} = \\ &= 0.253 \times \frac{|1+g_V|^2 + 0.877\,|g_P|^2 + 1.021\,\mathrm{Re}((1+g_V)g_P^*)}{|1+g_V'|^2 + 0.875\,|g_P'|^2} = \\ &= 0.253 \times \frac{0.552 + 0.877\,|1+g_V|^2 \left|0.429 - \frac{g_P^*}{1+g_V}\right|^2}{1 + 0.875\,|1+g_V'|^2 \left|\frac{g_P'^*}{1+g_V'}\right|^2}, \end{aligned} \tag{2.5.6}$$

with (for fixed $|1+g_V|^2$) concentric circles as level lines in the $(\mathrm{Re}(g_P^*/(1+g_V)), \mathrm{Im}(g_P^*/(1+g_V)))$ plane. Different values of $|1+g_V|^2$ change the value associated with each level line, but not their shape.

Distinguishing this operator from the first three groups of operators can be done in the same way as described for $(\mathbf{3}, \mathbf{2})_{7/6}$.

Distinguishing $(\mathbf{3}, \mathbf{2})_{7/6}$ from $(\bar{\mathbf{3}}, \mathbf{1})_{1/3}$ is difficult depending on the specific values of the coupling constants. In particular for $c_1 \approx 1$, the only difference is the coefficient of c_6 in $\mathcal{O}_1^0$, $\mathcal{O}_3^0$, $\mathcal{O}_1^T$, $\mathcal{O}_4^T$, and $\mathcal{O}_2^{0T}$, which either flips signs or at least shows some sizable difference (caused by the different sign of T_L relative to g_P). Thus, there is no clear "smoking gun" observable. Instead, several observables have to be combined in order to determine best-fit parameter values and to select the model with the best goodness of fit.

2.5.7 NP in the light modes

Most studies have so far assumed any NP contribution in the light modes to be negligible. In fact, the opposite would also imply that we could not use the light mode as normalization mode for observables of $\ell = \tau$ without more careful analysis of possible NP contributions. In this case, we would likely use the total cross section as normalization for any observable $\mathcal{O}_a$ for $\ell = \tau$, i.e. consider $\mathcal{O}_a/\Gamma$.

Here, we only give a short walk through on how to use the angular observables for possible studies to discriminate between the NP contributions of different operators in the light modes.

Finding a non-vanishing value for $\mathcal{O}_1^0$ would rule out $(\mathbf{1}, \mathbf{3})_0$, $(\mathbf{3}, \mathbf{3})_{2/3}$, and $(\bar{\mathbf{3}}, \mathbf{2})_{5/3}$ (similar to the role of the observable $\mathcal{O}_3^{0T}$ in the heavy case, which however now vanishes). As before, $(\mathbf{1}, \mathbf{2})_{1/2}$ and $(\bar{\mathbf{3}}, \mathbf{2})_{5/3}$ are the only operators for which most observables are SM-like. In fact, only $\mathcal{O}_1^0$ shows a NP contribution for light leptons. $(\mathbf{3}, \mathbf{1})_{2/3}$ can be distinguished from the first class of operators via $\mathcal{O}_1^0$ and from the second class by finding NP contributions in any observable but $\mathcal{O}_1^0$. Ratios of different observables distinguish $(\mathbf{3}, \mathbf{2})_{7/6}$ and $(\bar{\mathbf{3}}, \mathbf{1})_{1/3}$ from the first three classes of observables. This time both operators are also easier to distinguish, either by finding NP contributions to $\mathcal{O}_1^T$ or $\mathcal{O}_3^T$ or by considering the ratio $\mathcal{O}_3^0/\mathcal{O}_1^0$.

2.6 Comparison with Literature and Summary

As we have focused on *angular* observables, we will concentrate on a comparison with similar approaches in the literature and do not comment explicitly on observables that build on the polarization of the lepton or the D^* or employ the q^2 dependency.[10]. Furthermore, we do not include the discussion of *unbinned* observables (that require knowledge of the angles *per event*).[11] So far, the most prominent such observable built from a binned angular distribution has been the opening angle asymmetry (also called forward-backward asymmetry), considered e.g. in [59, 95, 117, 118, 122–127]:

$$A_\theta = \left[\int_{-1}^{0} - \int_{0}^{1}\right] \mathrm{d}\cos\theta_\ell \frac{\mathrm{d}\Gamma}{\mathrm{d}\cos\theta_\ell} = -\frac{4}{3}\pi(\overline{W}_3^0 + 2\overline{W}_3^T), \tag{2.6.1}$$

sometimes also considered with q^2 dependency (by integrating over $\mathrm{d}\Gamma/\,\mathrm{d}q^2\,\mathrm{d}\cos\theta_\ell$). Employing the same short hand notation as in table 2.2, we write $A_\theta = [-\pi,\pi] \times ([0,\pi/2] - [\pi/2,\pi]) \times [0,\pi]$.

Reference [124] went one step further and also considered asymmetries in χ, corresponding to

$$\begin{aligned}
A_C^{(1)} &= ([-\pi,0] - [0,\pi]) \times [0,\pi] \times [0,\pi], \\
A_C^{(2)} &= ([-\pi,0] - [0,\pi]) \times [0,\pi] \times ([0,\pi/2] - [\pi/2,\pi]) \dashrightarrow \propto \mathcal{O}_3^{0T}, \\
A_C^{(3)} &= ([-\pi,0] - [0,\pi]) \times ([0,\pi/2] - [\pi/2,\pi]) \times ([0,\pi/2] - [\pi/2,\pi]) \dashrightarrow \propto \mathcal{O}_4^{0T},
\end{aligned}$$

which were used together with an effective Lagrangian (that did however not include the tensor operator) to study CP violating triple products (which vanish in the SM without hadronic complications). Furthermore, the influence of g_A, g_V and g_P was considered (always assuming only one non-vanishing contribution) qualitatively by considering plots of the q^2 dependency of these observables for several benchmark coupling constants. The observables $A_C^{(2)}$ and $A_C^{(3)}$ match our expressions for $\mathcal{O}_3^{0T}$ and $\mathcal{O}_4^{0T}$ in section 2.3.4.3.

The (to our knowledge) most notable approach to a complete consideration of angle dependencies can be found in [117]. Though written down in a different convention, the considered Lagrangian and differential cross section underly the same assumptions as our approach. A total of eleven observables for $\bar{B} \longrightarrow D^* \tau^- \bar{\nu}_\tau$ are presented, out of which two split the differential decay rate along the lepton and D^* polarization and are not considered for comparison here. In addition to A_θ and $R(D^*)$, this leaves seven more observables, all of which have counterparts in section 2.3.4.3.[12]

The sensitivity of the observables to NP in [117] seems to be solely judged based on the visual impression of their q^2 dependency, separately considering several non-vanishing values for one coupling constant g_A, g_V, g_S, g_P and T_L at a time.

In comparison to [117], we have improved on several aspects:

- The construction of such angular observables has been formalized and their degrees of freedom for a given binning have been discussed under several different assumptions.

  In particular, the minimal number of bins required to construct all observables $\mathcal{O}_a$ has been determined to be $3 \times 3 \times 2$ (product bins). Considering easier to visualize weights that factor over the binning variables (product weights), a limit of $3 \times 3 \times 3$ was found.

[10] Some examples for such observables are found in [59, 94, 95, 118–120].

[11] An interesting and very recent approach that constructs unbinned observables for lepton universality tests that do not depend on the form factor parametrizations can be found in [121]. Note however, that their observables (that require χ, θ_ℓ, θ_{D^*}, and q^2 *per event*) are experimentally much more challenging.

[12] Specifically, comparing their notation to our results from 2.3.4.3, we have $A_5 \approx \mathcal{O}_3^T$ (slightly different, but similar expression), $A_8 = \mathcal{O}_3^{0T}$, $A_9 = \mathcal{O}_2^{0T}$, $A_{10} = \mathcal{O}_4^{0T}$ and $A_{11} = \mathcal{O}_1^{0T}$. For the observables C_χ and S_χ, no explicit integration formulas were given in the paper, but $C_\chi = \mathcal{O}_4^T$ and $S_\chi = \mathcal{O}_5^T$ is likely.

Furthermore, an even lower number of bins was found to be necessary for the construction of individual observables $\mathcal{O}_a$ (as shown in table 2.10).

- The constructed observables were compared based on their associated expected uncertainty and this figure of merit was used to maximize the sensitivity of the observables over their degrees of freedom. In comparison to "hand calculated" angular observables as presented in [117], this improves the sensitivity by a factor of 5 to 50, while using less or equally many bins.

  A comparison of the performance of the observables constructed under different assumptions was given in section 2.3.5. Already with $1 \times 3 \times 3$ bins, five observables with sensitivities close to the optimum can be constructed. Adding two more bins in χ, all $\mathcal{O}_a$ can be constructed. Even with significantly more bins, further sensitivity gains seems to be rather small. The expected statistical error of each observable for the $3 \times 3 \times 3$ binning is about 5 to 25 times higher than the expected statistical error on the total integrated cross section.

- The dependency of the obtainable sensitivity on the spacing of the bins for a fixed number of bins has been studied based on different assumptions. For $3 \times 3 \times 3$ bins, an equidistant spacing of bins in $\chi \times \theta_\ell \times \theta_{D^*}$ proved to be optimal.

- The dependency on the NP couplings has been made explicit and the discriminatory power of the observables was discussed. Rather than only considering one of the coupling constants g_A, g_V, g_S, g_P or T_L at a time, a basis of operators corresponding to possible mediator particles relevant to $\bar{B} \longrightarrow D^* l^- \bar{\nu}_l$ was considered.

2.7 Outlook

As Belle I already studied single differential rates for $\bar{B} \longrightarrow D^* \ell^- \bar{\nu}_\ell$ in ten bins in each kinematic variable for light leptons [61] and studies measuring the tau polarization in $\bar{B} \longrightarrow D^* \tau^- \bar{\nu}_\tau$ [18] demonstrated recent improvements in measurements of tau leptons, the requirement of $3 \times 3 \times 2$ bins in $\chi \times \theta_\ell \times \theta_{D^*}$ for the construction of the observables seems to be within reach, especially for higher statistics soon to be delivered by Belle II.

In light of that, a more thorough study of the obtainable sensitivity of the observables is warranted. In particular, the so far neglected effect of systematic uncertainties and of bin migration should be considered and the calculations should be scaled up to luminosity benchmarks. The effect of most systematic uncertainties is likely to be rather flat across all bins and of similar relative size as for the total differential cross section. As there are only a few bins in each variable, bin migration is also expected to be manageable, but a more careful analysis is needed.

Furthermore, the discriminatory power of the operators should be considered in a more quantitative way by considering benchmark points for potential NP contributions. For this, the coupling constants should be constrained with a combination of experimental results (and not just $R(D^*)$) in order to select a set of relevant benchmark parameters.

Chapter 3

Hadronically Tagged Analysis of $\bar{B} \longrightarrow D^* e^- \bar{\nu}_e$ vs $\bar{B} \longrightarrow D^* \mu^- \bar{\nu}_\mu$ with Data from Belle I

In this chapter, continuing work on an analysis of the $\bar{B} \longrightarrow D^* \ell^- \bar{\nu}_\ell$ decay channel is presented. The events are selected from 772 million $e^+e^- \to \Upsilon(4S) \to B\bar{B}$ events recorded with the Belle detector at KEKB, where one B meson is fully (hadronically) reconstructed. Unfolded differential decay rates in four kinematic variables are presented separately for $\ell = e$ and $\ell = \mu$ in the final state.

3.1 Introduction

The study of semileptonic[1] B decays into D^* mesons in context of the SM allows for a precise determination of the absolute value of the CKM matrix element V_{cb} as well as for the measurement of the form factor parameters involved in the decay. The measurement of $|V_{cb}|$ is of theoretical interest for tests of the unitarity of the CKM matrix in the SM, while precise measurements of the form factors also allow to reduce systematic uncertainties in other analyses in which the signal is a major background (for example $\bar{B} \longrightarrow D^* \ell^- \bar{\nu}_\ell$ is a background in $\bar{B} \longrightarrow \pi \ell^- \bar{\nu}$, $\bar{B} \longrightarrow D^* \tau^- \bar{\nu}$, and others [128, p. 2]). Furthermore, most theoretical predictions for $R(D^*)$ use experimental input for the form factors. In fact, the current average as calculated by the HFLAV collaboration [20] (see section 1.2) uses the unfolded kinematic distributions calculated by the analysis described in this chapter.[2]

The analysis presented here is *exclusive*, i.e. studies an explicit decay mode, $\bar{B} \longrightarrow D^* \ell^- \bar{\nu}_\ell$, of the B meson (in contrast to analyses such as $\bar{B} \longrightarrow X_c \ell^- \bar{\nu}$, called *inclusive*), where the full decay chain of the corresponding B meson (*signal* B) from the $B\bar{B}$ pair production is reconstructed. Furthermore, the analysis is *tagged*, i.e. the second B meson (*tag* B) is fully reconstructed. This analysis reconstructs the tag B decay in hadronic modes and thus is also said to be *hadronically tagged*. The full reconstruction of the tag side ensures a high purity of the considered data sample (and thus a more precise reconstruction of the decay kinematics) but leads to a lower reconstruction efficiency. A complementary untagged analysis of $\bar{B} \longrightarrow D^* \ell^- \bar{\nu}_\ell$ using an orthogonal data set from Belle I has been described in the conference note [128].

Both $B^- \longrightarrow D^{*0} \ell^- \bar{\nu}_\ell$ and $\bar{B}^0 \longrightarrow D^{*+} \ell^- \bar{\nu}_\ell$ are considered. In both cases, the presented data is CP averaged. The decays of charged B mesons were previously determined to be less well understood (in particular some inconsistencies and large error margins appeared in the fit to V_{cb} and the form factor parameters). Therefore the analysis so far focused on neutral B mesons.

[1] Involving at least one lepton and at least one other particle.

[2] More precisely, the unfolded kinematic distributions as published in the conference note [61] are being used.

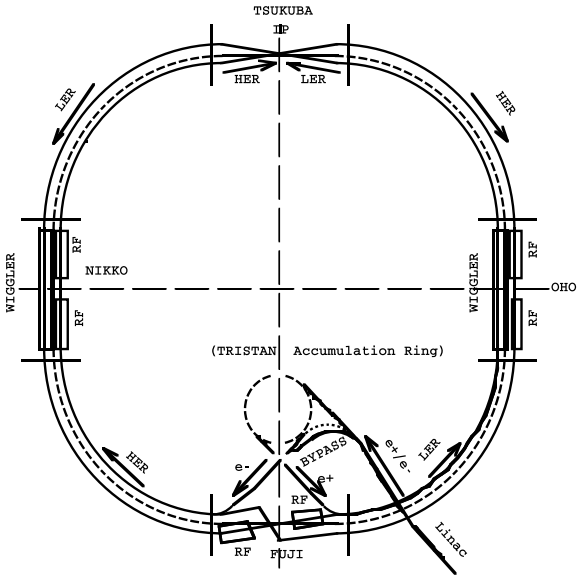


Figure 3.1: Schematic of the KEKB facility. Electrons circle clockwise in the HER, positrons counterclockwise in the LER. The four directions are denoted as Tsukuba, Oho, Fuji, and Nikko, corresponding to the direction of cities and landmarks. [130]

In addition to a determination of $|V_{cb}|$ and $B\longrightarrow D^*$ form factor parameters (not described in this thesis), the analysis also presents the unfolded (i.e. detector resolution corrected) differential decay rates of the four kinematic variables w, $\cos\theta_\ell$, $\cos\theta_{D^*}$ and χ.

The work on the analysis presented here was started in 2016 by a master's thesis of Saskia Falke (formerly Saskia Moenig) under the supervision of Florian Bernlocher and with additional input from Jan Hasenbusch, Max Rüger, and Jochen Dingfelder. Internally, the work is documented as Belle Note #1419 [129] with referees Ariane Frey (chair), Kay Kinoshita and Christian Pulvermacher. First results were shown at the conference CKM 2016 and a corresponding conference note [61] was published in February 2017, but further progress was stalled thereafter.

The aim of this work is therefore to prepare for a completion of the analysis, but also extend it in one aspect: So far, the unfolded differential cross sections and other results were presented for $\bar{B}\longrightarrow D^*\ell^-\bar{\nu}_\ell$ without distinguishing between $\ell = e$ and $\ell = \mu$ in the final state. However, publishing separate results for both lepton flavors would allow to weigh in on the debate of possible light lepton universality breaking features of potential NP.

As the Belle note [129] (only available for Belle members) and the conference note [61] are describing most parts of the analysis, they are considered as implicit references throughout this chapter. However, if no explicitly source is given, all numbers, plots and figures have been recalculated, regenerated or newly created by the author.

3.2 Dataset and Instrumentation

This analysis uses a data set of an integrated luminosity of 711 fb^{-1} at a center of mass energy $\sqrt{s} = 10.58\,\text{GeV}$ (corresponding to 772 million $e^+e^-\longrightarrow\Upsilon(4S)\longrightarrow B\bar{B}$) recorded between 1999 and 2010 with the Belle detector at the KEKB storage ring.

3.2.1 KEKB

KEKB is an electron-positron-collider located at the *High Energy Acceleration Research Organization* (KEK) in Tsukuba, Japan. It is described in detail in [11], which serves as primary source for this section.

KEKB features two storage rings, HER (High Energy Ring) and LER (Low Energy Ring) of 3 km circumference in which electrons and positrons circle at an energy of 8 GeV and 3.5 GeV, respectively. Both rings cross at a crossing angle of 11 mrad at a single point (Interaction Point, IP) at which electrons and positrons collide. The asymmetric energy setup translates to a "boosted" center of mass system of the electron positron pair, which moves with $\beta\gamma = (E(e^-) - E(e^+))/\sqrt{s} = 0.425$ along the beam axis as seen from the laboratory frame. The center of mass energy is tuned to the $\Upsilon(4S)$ resonance, which decays with at least 96% probability (at 95% confidence level [5]) to a pair of neutral or charged B mesons.

Both rings are fed from the KEKB injector linac [131], a 600 m long linear accelerator that directly reaches the target energies of 8 GeV (electrons) and 3.5 GeV (positrons). Radiofrequency cavities (RFs) and wigglers along the KEKB storage rings are used to counter energy loss from synchrotron radiation and to control beam properties.

Achieved on June 17, 2009, the peak luminosity of $21.083 \times 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ [132] is still the world record, although soon to be surpassed by the Super KEKB upgrade that aims to deliver a 40 fold higher target luminosity for the upcoming Belle II experiment [133].

3.2.2 Belle

The Belle detector is located at the interaction point of the KEKB storage rings and surrounds both beams. Cross sections of the detector are shown in figure 3.2 and 3.3. The z-axis is set to be parallel to the e^+ beam (but pointing in opposite, i.e. clockwise direction), the y axis lies in the horizontal plane (pointing toward the middle of the accelerator ring) and the x axis points upward such that x, y, z form a right-handed orthogonal coordinate system. The polar angle θ is measured with respect to the z axis, the azimuthal angle ϕ with respect to the x axis.

The Belle detector is described in the technical design report [10], which is the source of the information in this section, if not stated otherwise.

The Belle detector consists of the following components:

- Beam pipe: A double wall beryllium cylinder with a cooling liquid flowing between both walls. Originally with a diameter of 40 mm, it was replaced by a smaller version with 20 mm diameter in 2003 [134].
- SVD (Silicon Vertex Detector). The innermost detector, directly surrounding the beam pipe, offering precise vertex information and therefore important for the tracking of charged particles. The first generation SVD (SVD1) was in use until 2003, where it was replaced by a second and improved generation (SVD2) due to radiation damage [135]. The angle coverages were $23° < \theta < 139°$ (SVD1) and $17° < \theta < 150°$ (SVD2), respectively.
- EFC (Extreme Forward Calorimeter). A calorimeter on bismuth-germanate crystal base, complementing the ECL (see below) by covering $6.4° < \theta < 11.5°$ (forward) and $163.3° < \theta < 171.2°$.
- CDC (Central Drift Chamber). The CDC surrounds the SVD and parts of the EFC. It is filled with a $50\% - 50\%$ mixture of helium and ethane. Alternating field and sense wires detect avalanches of secondary charged particles following the path of energetic charged particles. This allows for tracking information (and due to the curved trajectories in the magnetic field also for momentum measurements), as well as for measurements of the energy loss $\mathrm{d}E/\,\mathrm{d}x$.
- ACC (Aerogel Cherenkov Counters). The ACC consist of several types of aerogel modules. Sufficiently fast particles traveling through the aerogel lead to Cherenkov radiation, which is read out with photomultiplier tubes and allows determining the particle velocity. Together

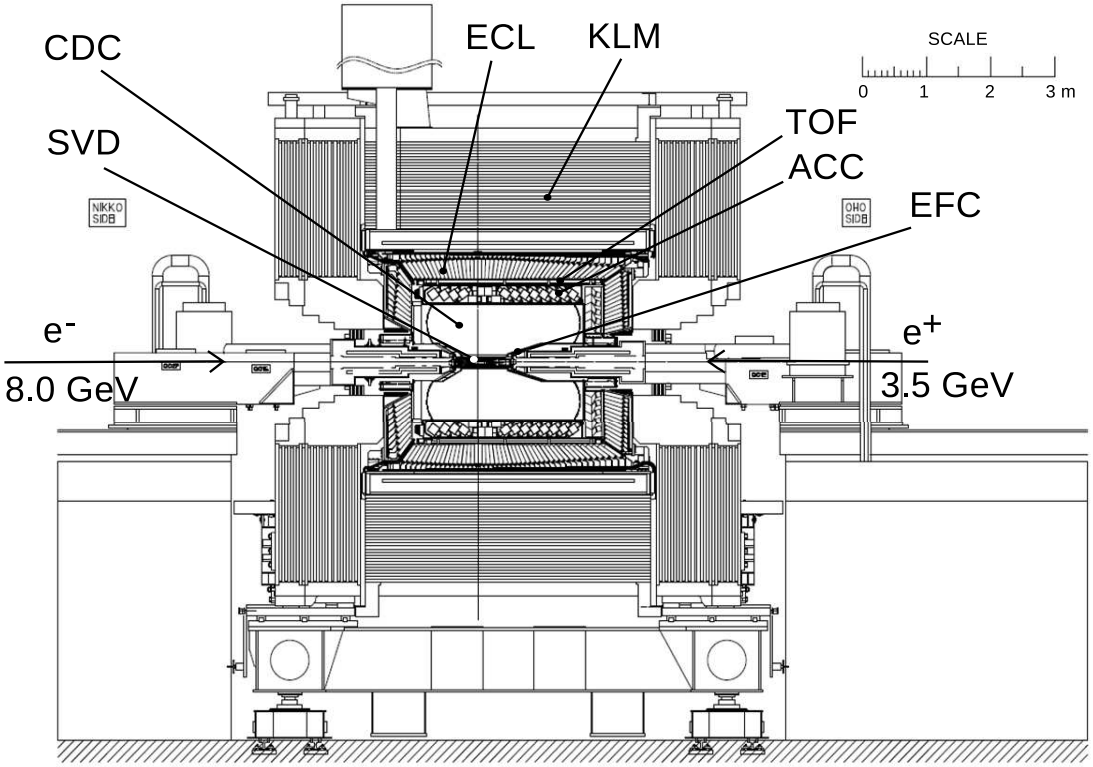


Figure 3.2: Longitudinal cross section of the Belle detector [130].

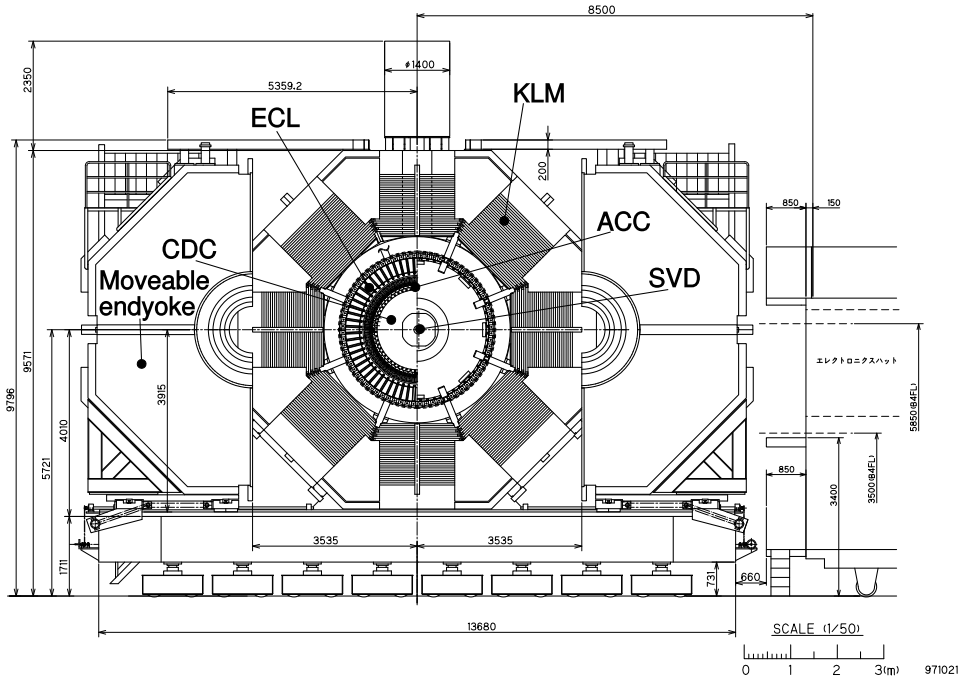


Figure 3.3: Transverse cross section of the Belle detector [130].

with the momentum measurement from the CDC, this allows calculating the mass of the particles. The ACC is particularly important to distinguish light mesons such as Kaons and Pions.

- TOF (Time Of Flight). The TOF uses plastic scintillation counters to measure the time of flight of particles over the 1.2 m distance from the interaction point. With a resolution of 100 ps, the TOF is effective for particles with momentum below 1.2 GeVc^{-1} (that is, for nearly 90% of the particles from the $\Upsilon(4S)$ resonance). The TOF covers $34° < \theta < 120°$.

- ECL (also abbreviated *CsI*, Electromagnetic Calorimeter). The ECL consists of CSI(TI) crystals placed around the z axis in a barrel structure with two end caps. The primary use is the measurement of energy and position of photons. The barrel covers $32.27° < \theta < 128.77°$, while the front and endcap cover $32.27° < \theta < 128.7°$ and $130.77° < \theta < 155.17°$, respectively.

- Magnet. The superconducting solenoid provides a magnetic field of 1.5 T. The magnetic system takes advantage of the iron layers of the KLM (see below), which are integrated into a magnetic return circuit.

- KLM (K_L and Muon Detector). The outermost part of the detector built from alternating layers of glass-resistive plate counters and iron plates, arranged as a barrel around the z axis with two endcaps. Covering $20° < \theta < 166°$, it allows the identification of muons and K_L mesons with high efficiency for $|p| > 600\,\mathrm{MeV}/c$.

3.3 Reconstruction and Event Selection

The event selection is performed with the objective of a sufficiently pure sample of $\bar{B} \longrightarrow D^* \ell^- \bar{\nu}_\ell$ decays that allows for a good reconstruction of the kinematic variables.

3.3.1 Kinematic variables

$|V_{cb}|$ and form factor parameters are extracted from measured distributions of the four kinematic variables w, $\cos\theta_\ell$, $\cos\theta_{D^*}$ and χ. The variable w is linearly related to the squared four-momentum transfer between the B and the D^*, $q^2 = (p_B - p_{D^*})^2$ via:

$$w = \frac{m_B^2 + m_{D^*}^2 - q^2}{2 m_B m_{D^*}}. \tag{3.3.1}$$

The helicity angles θ_ℓ, θ_{D^*} (also denoted θ_v) and the angle χ are defined as in figure 2.1.

In the reconstruction, w is calculated with $p_B = p_{e^+e^-} - p_{\mathrm{tag}}$ and the D^* four-momentum. To calculate the angles, the four-momenta of the D^* and lepton are boosted in the signal $\bar{B}$ meson frame. A comparison of the kinematic distributions of data vs MC is shown in figure 3.4.

As the resolution is similar in all four variables, the extraction of the differential branching ratios is performed in the same number of ten bins.

3.3.2 Tag side reconstruction

The candidate for a tagging side B (B_{tag} candidate) is reconstructed using the hadronic full reconstruction algorithm described in [136]. Based on the multivariate analysis package `NeuroBayes` [137] and employing more than 70 neural networks to reconstruct more than 1100 exclusive decay channels, it obtains an efficiency of 0.3% for $B^\pm$ and 0.18% for B^0.

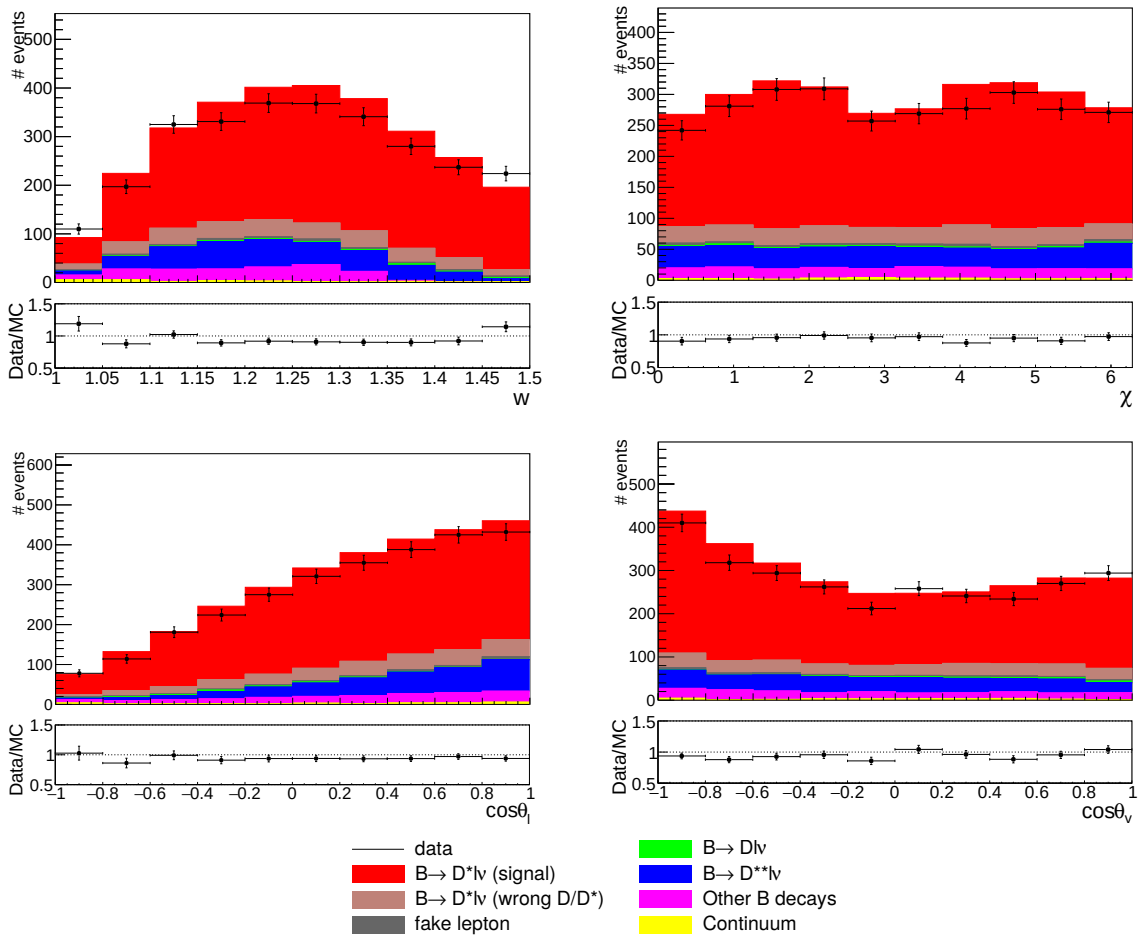


Figure 3.4: Distributions of w, χ, $\cos\theta_\ell$ and $\cos\theta_{D^*}$ in Data vs. MC (B^0, $\ell = e + \mu$; after the MC corrections described in section 3.3.5 have been applied) [61].

The beam constrained meson mass, missing energy, and the multivariate classifier (M_{bc}, ΔE and log(NB), respectively) have to fulfill the requirements

$$\begin{aligned} 5.265\,\mathrm{GeV}c^{-1} < M_{\mathrm{bc}} &= \sqrt{\frac{s}{4} - |\vec{p}_{\mathrm{tag}}|^2}, \\ -0.15\,\mathrm{GeV} < \Delta E &= E_{\mathrm{tag}} - \frac{\sqrt{s}}{2} < 0.1\,\mathrm{GeV}, \\ \log(\mathrm{NB}) &> -7, \end{aligned} \tag{3.3.2}$$

where $\sqrt{s}$ is the center of mass (CMS) energy of the e^+e^- collision and $\vec{p}_{\mathrm{tag}}$ (E_{tag}) is the 3-momentum (energy) of the B_{tag} candidate in the e^+e^- CMS.

The cuts on M_{bc} and log(NB) are chosen after considering the figure of merit $N_{\mathrm{s}}/\sqrt{N_{\mathrm{s}} + N_{\mathrm{b}}}$ on MC data (see section 3.3.5), where N_{s} is the total number of correctly reconstructed signal events and N_{b} is the total number of remaining (background) events.

In each event, a single B_{tag} candidate is chosen and all tracks and neutral clusters associated with it are removed from the event for the reconstruction of the signal side. If no candidate is found, the event is vetoed.

3.3.3 Signal side reconstruction

The $\bar{B}\longrightarrow D^*\ell^-\bar{\nu}_\ell$ signal side is reconstructed in three steps:

1. Selection of the charged lepton candidate (events with more than one well identified charged lepton are vetoed). The following cuts are applied:

$$\begin{aligned} \mathcal{L}_e(e) > 0.5 \quad &\text{or} \quad \mathcal{L}_\mu(\mu) > 0.9, \\ |\vec{p}_e| > 0.3\,\text{GeV} \quad &\text{or} \quad |\vec{p}_\mu| > 0.6\,\text{GeV}, \\ 17^\circ < \theta_e < 150^\circ \quad &\text{or} \quad 25^\circ < \theta_\mu < 145^\circ, \\ \mathrm{d}r_e < 0.5\,\text{cm} \quad &\text{and} \quad \mathrm{d}z_e < 2\,\text{cm}, \end{aligned} \tag{3.3.3}$$

   where $\mathcal{L}_\ell(\ell)$ is a particle identification likelihood ratio for the electron or muon hypothesis, $\vec{p}_\ell$ is the electron/muon momentum and the cuts on the polar angles θ_ℓ of the charged lepton are due to detector acceptances. For electrons, additional impact parameter requirements on $\mathrm{d}r$ and $\mathrm{d}z$ are applied as well.

   The cuts follow the general recommendations of the Belle CKM working group. Further optimization is not performed in order to have as few constraints on the $\bar{B}\longrightarrow D^*\ell^-\bar{\nu}_\ell$ phase space as possible.

   For electrons, care has to be taken with energy and momentum loss due to bremsstrahlung and final state radiation. To that end, photons in a cone with opening angle 5° around the electron trajectory are considered. The photon with the smallest opening angle inside this cone (if any) is taken into account for the calculation of the electron energy and momentum.

2. Next, candidates for D mesons are reconstructed using the channels shown in table 3.1a. Neutral pions are reconstructed from photon candidates (calorimeter clusters not matched to any track) with energy E_γ. The following cuts are applied:

$$\begin{aligned} \mathcal{L}_K(K) > 0.6 \quad &\text{and} \quad \mathcal{L}_K(\pi) < 0.6 \\ \mathrm{d}r_{K/\pi} < 0.5\,\text{cm} \quad &\text{and} \quad \mathrm{d}z_{K/\pi} < 2.0\,\text{cm} \\ E_\gamma > \begin{cases} 100\,\text{MeV} & \theta < 33^\circ \\ 50\,\text{MeV} & 33^\circ < \theta < 128^\circ \\ 150\,\text{MeV} & 128^\circ < \theta \end{cases} \\ 0.12\,\text{GeV} < m_{\pi^0} < 0.15\,\text{GeV} \\ |m_D^{\text{reco}} - m_D^{\text{true}}| < 14\,\text{MeV} \end{aligned} \tag{3.3.4}$$

   Furthermore, for neutral D mesons, no further charged track is allowed on the signal side, whereas for charged D mesons, at most one more charged track is allowed. The best candidate is selected by a vertex fit: Neutral pion candidates are fitted to the decay vertex (expected to be close to the interaction point) and the candidate with the lowest χ^2/n.d.f. is selected.

3. Finally, D^* meson candidates are reconstructed in the channels shown in figure 3.1b. The additional channel $\bar{D}^{*0}\longrightarrow\bar{D}^0\gamma$ which would make up for the remaining 35% branching ratio of the $\bar{D}^{*0}$ decay is not considered due to very low purity. Only one more cut is applied for the deviation from the expected mass difference, $\Delta m = m_{D^*} - m_D$:

$$|\Delta m^{\text{reco}} - \Delta m^{\text{true}}| < 0.01\,\text{GeV}. \tag{3.3.5}$$

   If there are multiple π^0 candidates that can be combined with the D meson reconstructed in step 2, the one with the value Δm^{reco} closest to the true value Δm^{true} is selected.

3.3.4 Backgrounds

After the event selection described in the previous selection, a fit in m^2_{miss} is performed to discriminate between signal and background, where

$$m^2_{\text{miss}} = (p_B - p_{D^*} - p_\ell)^2. \tag{3.3.6}$$

Channel	BR
$D^-\longrightarrow K^+\pi^-\pi^-$	9%
$\bar{D}^0\longrightarrow K^+\pi^-$	4%
$\bar{D}^0\longrightarrow K^+\pi^-\pi^0$	14%
$\bar{D}^0\longrightarrow K^+\pi^-\pi^-\pi^+$	8%

(a) D channels

Channel	BR
$\bar{D}^{*0}\longrightarrow \bar{D}^0\pi^0$	65%
$D^{*-}\longrightarrow \bar{D}^0\pi^-$	68%
$D^{*-}\longrightarrow D^-\pi^0$	31%

(b) D^* channels

Table 3.1: Considered decay channels for D and D^* mesons. The branching ratios displayed here are the (rounded) 2018 world averages [5].

Here, the signal side four-momentum p_B is calculated via the beam four-momentum as $p_B = p_{e^+e^-} - p_{\text{tag}}$. Figure 3.5 shows the m^2_{miss} distribution of the following notable event classes:

- Correctly reconstructed signal events $\bar{B}\longrightarrow D^*\ell^-\bar{\nu}_\ell$: Since $m^2_{\text{miss}} = p^2_\nu = 0$, the m^2_{miss}-distribution peaks at zero.
- True signal $\bar{B}\longrightarrow D^*\ell^-\bar{\nu}_\ell$ events in which the D or D^* meson has not been reconstructed correctly. These events also peak around vanishing m^2_{miss}, but form a broader peak than correctly reconstructed events.

  Due to the similar distribution shape in m^2_{miss}, separating these events from correctly reconstructed signal events via the fit to m^2_{miss} would lead to strongly anticorrelated event yields. Therefore, rather than doing so, we fit them together with correctly reconstructed signal events (henceforth referred to as signal+wrong or s+w for short). The drawback of this approach is higher migration between bins of the kinematic variables as discussed in section 3.7.
- $\bar{B}\longrightarrow D^{**}\ell^-\bar{\nu}_\ell$: Confused for signal if additional decay products from $D^{**}\longrightarrow D^*$ are missed. The missing particles shift the m^2_{miss} spectrum to positive values. The branching fraction for these events is poorly known and filled in by comparison of exclusive modes with inclusive $\bar{B}\longrightarrow X\ell^-\bar{\nu}$ measurements (cf. section 3.3.5.1).
- $\bar{B}\longrightarrow D\ell^-\bar{\nu}_\ell$: A combination of fake pions or pions from the tagging side is combined with the D meson from the signal side and results in a fake D^* meson candidate. Peaks at negative m^2_{miss}.
- Others: This contains all backgrounds not belonging to any other category. This background is dominated by B decays that are not semileptonic, e.g. cascade decays, where a lepton is produced in a secondary decay. Confusing them with $\bar{B}\longrightarrow D^*\ell^-\bar{\nu}_\ell$ means that particles are missed, hence a positive m^2_{miss} distribution is observed.
- Fake leptons: Either a different particle is mistaken for a charged lepton, or the flavor of a true charged lepton is incorrectly identified. This background shows a broad m^2_{miss} shape, but a very small yield.
- Continuum: As both B mesons are fully reconstructed, the continuum bacgkround (non-$B\bar{B}$ events) is fairly low. It is uniform in m^2_{miss}.

3.3.5 Monte Carlo data

Simulated collision events are employed heavily throughout the analysis, most notably to fit the signal and background m^2_{miss}-distribution shapes. This simulated data is referred to as *Monte*

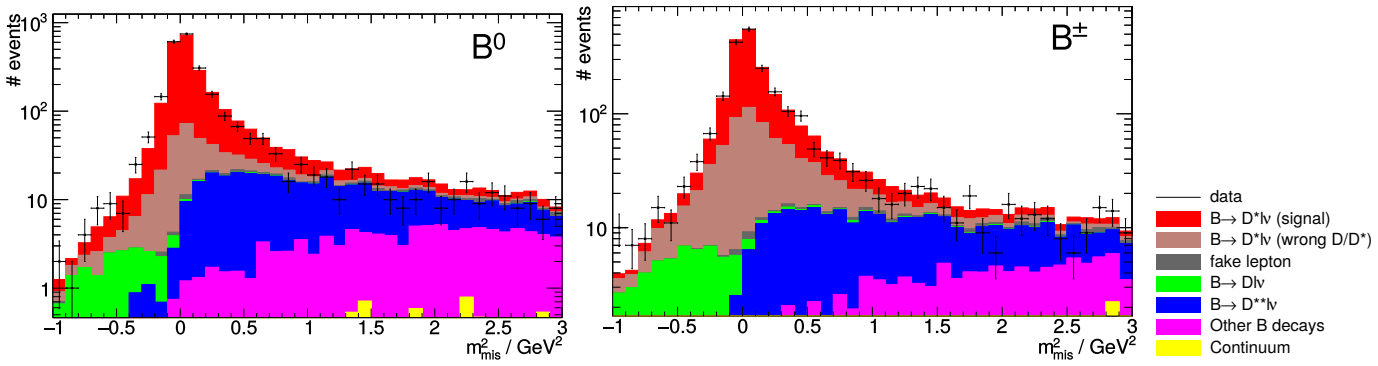


Figure 3.5: Signal and background distributions in m^2_{miss} [129].

Carlo (MC) data and has been generated by the `EvtGen` [138] package. To allow for a thorough comparison with data, a full detector simulation has been performed on these events using `GEANT3` [139] with final state radiation simulation provided by `PHOTOS` [140].

Information based on the reconstruction of MC events on which a full detector simulation had been performed is referred to as *reco* MC information, whereas the original (generator level) information is referred to as *truth/true* MC information. By comparing reco with truth MC, resolution effects and the precision of the applied reconstruction algorithm can be studied.

The MC events are divided into ten *streams* (i.e. subsamples), each split up in four subsamples with $B^\pm B^\mp$, $B^0\bar{B}^0$, $c\bar{c}$ continuum and $q\bar{q}$ ($q \neq c$) continuum events.[3] The total number of used MC data events is about ten times higher than the number of data events.

In order to update relevant branching fractions in the MC to their current averages, several corrections are necessary, which are applied to the MC datasets in the form of weights applied to each event. As the event frequencies scale with the branching fraction, these weights are of the form $\text{BR}_{\text{new}} / \text{BR}_{\text{MC}}$, where BR_{MC} is the branching ratio used in the generation of the MC and BR_{new} is the desired (updated) branching ratio.

3.3.5.1 Branching ratio corrections

The branching ratios for the considered D and D^* meson decay channels (table 3.1), $\bar{B} \longrightarrow D^{(*)}\ell^-\bar{\nu}_\ell$ and $\bar{B} \longrightarrow D^{**}\ell^-\bar{\nu}_\ell$ decays have been scaled up to the 2016 averages of the PDG group [141].[4] After these corrections, the branching ratio of $\bar{B} \longrightarrow X\ell^-\bar{\nu}_\ell$ is still smaller than the world average from [142]. As all relevant branching ratios but $\bar{B} \longrightarrow D^{**}\ell^-\bar{\nu}_\ell$ are well known, the $\bar{B} \longrightarrow D^{**}\ell^-\bar{\nu}_\ell$ branching ratio has been scaled up by a factor of 2.244 ($B^\pm$) and 2.601 (B^0) to meet the inclusive branching ratio.

Note that the numbers quoted in the Belle note [129] still refer to an earlier version of the code with correction factors calculated with the 2014 world averages [142].

3.3.5.2 Differential Branching Ratio Corrections

The shape of the differential branching ratios has been updated using the form factor parametrization from Caprini, Lellouch and Neubert (CLN) [60] and the 2014 world averages from the HFLAV group [143] as parameters. The full parametrization of the differential branching ratios is given in [129] and [61], but is in principle the same as the expressions discussed in appendix A (except that only SM contributions are considered).

[3] However, only the first 5 streams contain $B\bar{B}$ events.

[4] The value $\text{BR}(\bar{D}^{*0} \longrightarrow \bar{D}\pi)$ has not yet been updated since the 2014 PDG averages. This is also discussed in section 3.10.6.

In order to not interfere with the branching ratio correction described above, the weights are calculated as

$$\frac{\Gamma_{\mathrm{MC}}}{\Gamma_{\mathrm{new}}}\cdot\frac{\left.\frac{\mathrm{d}^4\Gamma}{\mathrm{d}w\,\mathrm{d}\cos\theta_\ell\,\mathrm{d}\cos\theta_{D^*}\,\mathrm{d}\chi}\right|_{\mathrm{new}}}{\left.\frac{\mathrm{d}^4\Gamma}{\mathrm{d}w\,\mathrm{d}\cos\theta_\ell\,\mathrm{d}\cos\theta_{D^*}\,\mathrm{d}\chi}\right|_{\mathrm{MC}}}, \tag{3.3.7}$$

where quantities with subscript "MC" are as assumed for the MC generation, whereas quantities with subscript "new" are the updated values as described.

3.3.5.3 Reconstruction and particle identification efficiency correction

Reconstruction and particle identification (PID) efficiencies are corrected using data driven methods:

- PID efficiencies: The standard Belle efficiency tables are used.
- Slow pion efficiency: The momentum-dependent efficiency correction is taken from a study of $\bar{B}\longrightarrow D\pi$ samples [144]
- Tagging efficiency: This calibration is performed using $\bar{B}\longrightarrow X\ell\bar{\nu}_\ell$ decays (since our signal mode $\bar{B}\longrightarrow D^*\ell^-\bar{\nu}_\ell$ is a part of this calibration sample, this introduces a small bias). The correction factors have been calculated as part of a Bachelor thesis [145] based on a procedure described in [146]: $\bar{B}\longrightarrow X\ell^-\bar{\nu}_\ell$ events are binned with respect to their decay topology and in bins of the multivariate classifier log(NB). The correction factor is then calculated separately per bin by comparing the number of $\bar{B}\longrightarrow X\ell\bar{\nu}_\ell$ events in data with the number of events in MC.

3.4 Fitting

The $\bar{B}\longrightarrow D^*\ell^-\bar{\nu}_\ell$ signal+wrong yield[5] is extracted using an unbinned likelihood fit to the m^2_{miss} distribution.[6] This is done separately for the disjoint samples that contain either only $B^\pm$ or B^0 decays and, if we distinguish between the lepton type, only either electrons or muons in the final state of the signal side.

All events of the subsample are projected in bins of the kinematic variable under consideration (in this context also referred to as *binning variable*), leading to $n = 10$ subsamples.

The signal+wrong and background yield $N^j_{\mathrm{s+w}}$ and N^j_{b} of each subsample j is extracted by maximizing the likelihood [61, p. 13]

$$\begin{aligned}\mathcal{L}(m^2_{\mathrm{miss}};N^j_{\mathrm{s+w}},N^j_{\mathrm{b}}) = \frac{1}{n!}\exp(-N^j_{\mathrm{s+w}}-N^j_{\mathrm{b}})\times\\ \times\prod_{i=1}^{n}\Big[N^j_{\mathrm{s+w}}\,\mathrm{pdf}^j_{\mathrm{s+w}}(m^2_{\mathrm{miss}}(i))+N^j_{\mathrm{b}}\,\mathrm{pdf}^j_{\mathrm{b}}(m^2_{\mathrm{miss}}(i))\Big],\end{aligned} \tag{3.4.1}$$

where $m^2_{\mathrm{miss}}(i)$ is the value of m^2_{miss} of the i-th event in the subsample and $\mathrm{pdf}^j_{\mathrm{s+w}}$ ($\mathrm{pdf}^j_{\mathrm{b}}$) are the probability density functions (PDFs) of the m^2_{miss} distributions for signal+wrong (background) in bin j. These PDFs are also called *fit templates* and are obtained from fitting the MC m^2_{miss} distributions.[7]

[5] Correctly reconstructed signal events are treated together with signal events, where the D^* or D meson have been incorrectly reconstructed as explained in section 3.3.4.

[6] In the following, it is important not to confuse the fitting variable (m^2_{miss}) and the binning/kinematic variables w, $\cos\theta_\ell$, $\cos\theta_{D^*}$ and χ.

[7] The different types of background have different distributions in m^2_{miss}. In order to fit the total background PDF, $\mathrm{pdf}^j_{\mathrm{b}}$, the background composition (i.e. the different normalization factors of the distributions) have been fixed to their values in MC.

Both the maximization and the fitting of the templates is carried out numerically using the `RooFit` package [147]. [61, p. 13] claims that the number of signal events is *not* constrained to be positive, but that is not true for the latest version of the code. However, after anomalies showed up for the closure test explained in section 3.5.3, the positivity constraint was relaxed and negative values were to some extent allowed for both signal+wrong and background.

To obtain a smooth PDF as fit template from the MC data, Gaussian kernel estimators [148] are used via the `RooKeysPdf` class (that is, the PDF is constructed as a sum of Gaussians). The corresponding *smoothing parameter* is chosen to be 0.75 as a compromise to smooth statistical fluctuations on the one hand, while not discarding physics information on the other hand. This is less of a concern for the signal shape, where high MC statistics lead to a well-defined shape and more for the background, where the MC data is limited and strong "wiggles" are observed.

Some examples for the fitting of the PDF templates are shown in figure 3.6; examples for a fit to data using these PDFs are presented in figure 3.7. Exemplary fit results for w are shown in figure 3.8, the full set of 24 plots is collected on pages 118 to 121.

3.5 Fit Closure Tests

To demonstrate that the fit procedure is working as expected, three *closure tests* are performed. Adopting the names from [129], the three tests are called "closure toys", "closure test 1" and "closure test 2". Schematic descriptions in the form of flowcharts are given in figure 3.9 to 3.11.

In addition to the tests with MC data as described in section 3.3.5, the same closure tests are performed with MC, where the reweighting was modified by varying the $\bar{B}\longrightarrow D^*\ell^-\bar{\nu}_\ell$ form factor parameters by 3σ (that is, the same MC variations that are also used to estimate the model error in section 3.8.2). This is done to confirm that the unbiasedness of the fit is a stable feature independent of small variations in the MC samples.

The thereby significantly increased number of performed closure tests require a more systematic analysis. To that end, additional plotting routines and analysis tools have been implemented using `pyROOT` [149].

3.5.1 Unbiasedness of the fit

The test "closure toys" compares MC truth yields with results of fits to distributions of toy events generated from the same MC dataset: Taking the MC dataset, PDFs for signal and background in each bin i ($\mathrm{pdf}^i_{\mathrm{s+w}}$ and $\mathrm{pdf}^i_{\mathrm{b}}$) are calculated as for the fitting procedure. In addition the "total" PDF $\mathrm{pdf}^i_{\mathrm{tot}}$ is calculated using the ratio of signal to background in MC:

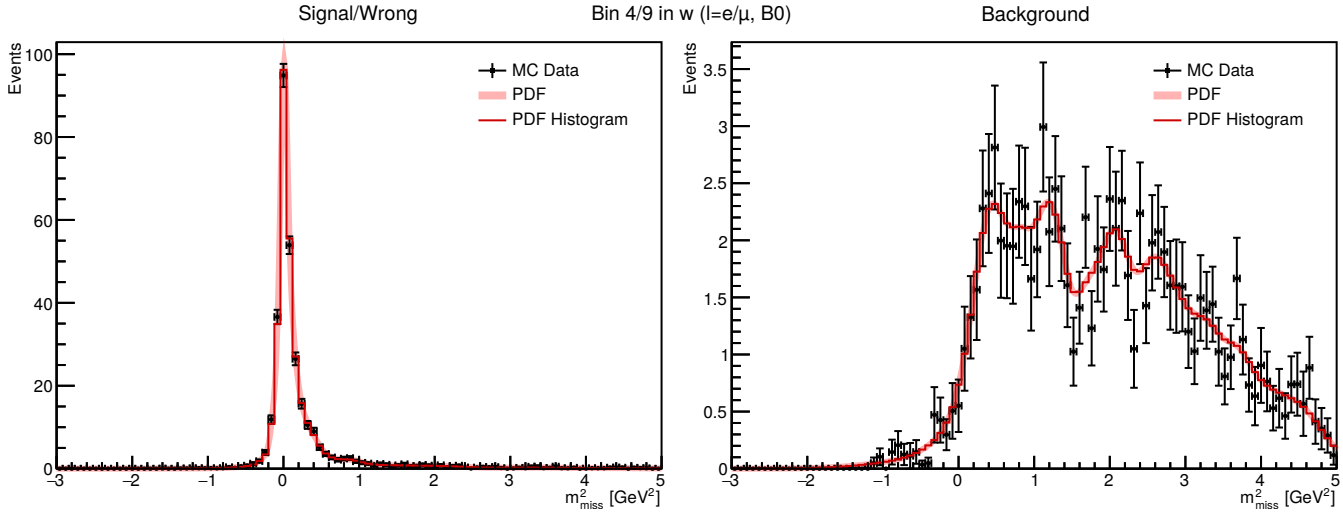
$$\mathrm{pdf}^i_{\mathrm{tot}} = \frac{N^{i,\mathrm{MC}}_{\mathrm{b}} \cdot \mathrm{pdf}^i_{\mathrm{b}} + N^{i,\mathrm{MC}}_{\mathrm{s+w}} \cdot \mathrm{pdf}^i_{\mathrm{s+w}}}{N^{i,\mathrm{MC}}_{\mathrm{b}} + N^{i,\mathrm{MC}}_{\mathrm{s+w}}}. \tag{3.5.1}$$

Using this PDF, toy events are generated based on a Poisson distribution with the total number of events given by the total number of events of this bin in MC.[8]

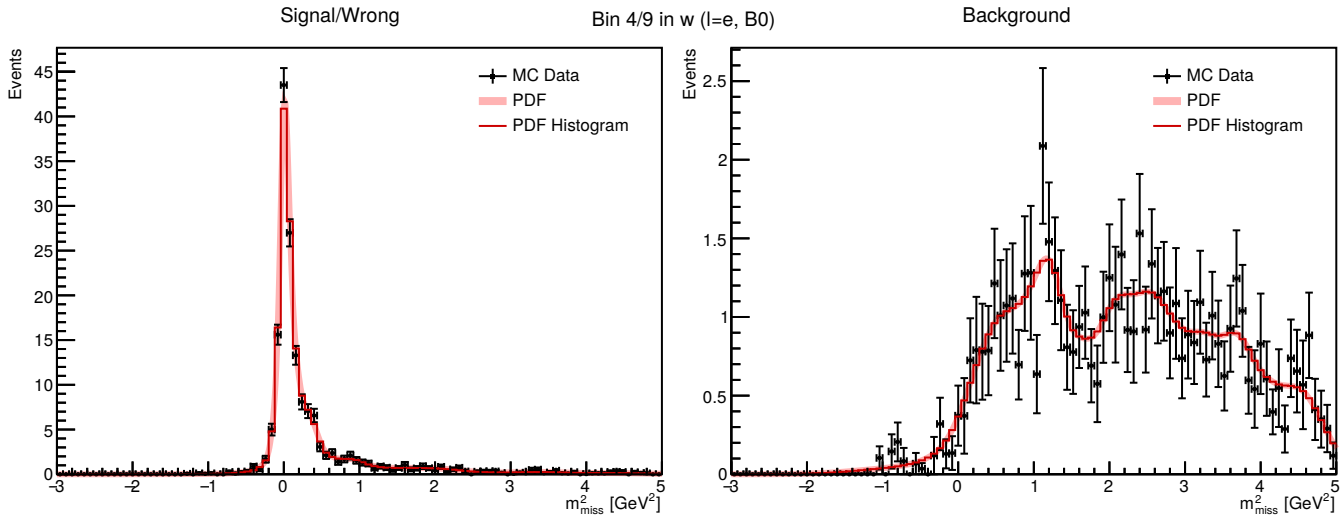
The m^2_{miss} spectrum of the toy events is then fitted using $N^i_{\mathrm{fit,\,s+w}} \cdot \mathrm{pdf}^i_{\mathrm{s+w}} + N^i_{\mathrm{fit,\,b}} \cdot \mathrm{pdf}^i_{\mathrm{b}}$, yielding $N^i_{\mathrm{fit,s+w/b}}$ and the fit error $\sigma^i_{\mathrm{fit,\,s+w/b}}$. These fit yields are compared with the number of (truth) MC events for signal and background in each bin, $N^i_{\mathrm{true,\,s+w/b}}$, by considering the "pull":

$$\mathrm{pull}^i_{\mathrm{s+w/b}} := \frac{N^i_{\mathrm{fit,\,s+w/b}} - N^i_{\mathrm{true,\,s+w/b}}}{\sigma^i_{\mathrm{fit,\,s+w/b}}} \tag{3.5.2}$$

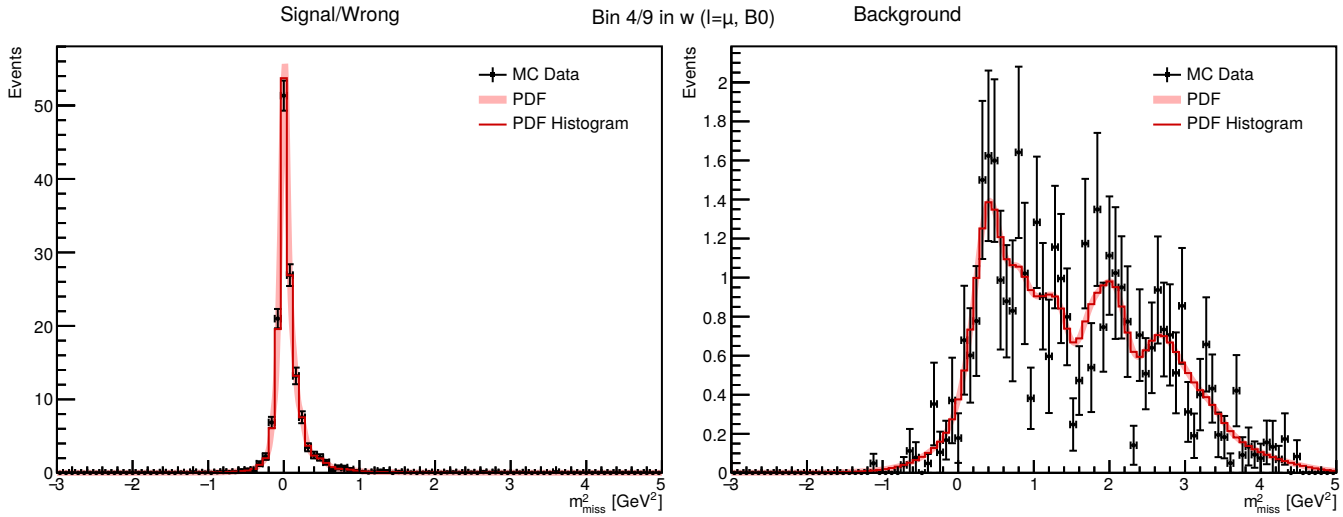
[8] The internal documentation [129] can be misunderstood to imply that the total number of events is varied by a Poisson statistics as well, but that is not the case.



(a) $\ell = e/\mu$



(b) $\ell = e$



(c) $\ell = \mu$

Figure 3.6: Fitting PDFs for signal+wrong and background in bin 4 in w for the combined fit $\ell = e/\mu$ and for the separate fits $\ell = e$ and $\ell = \mu$.

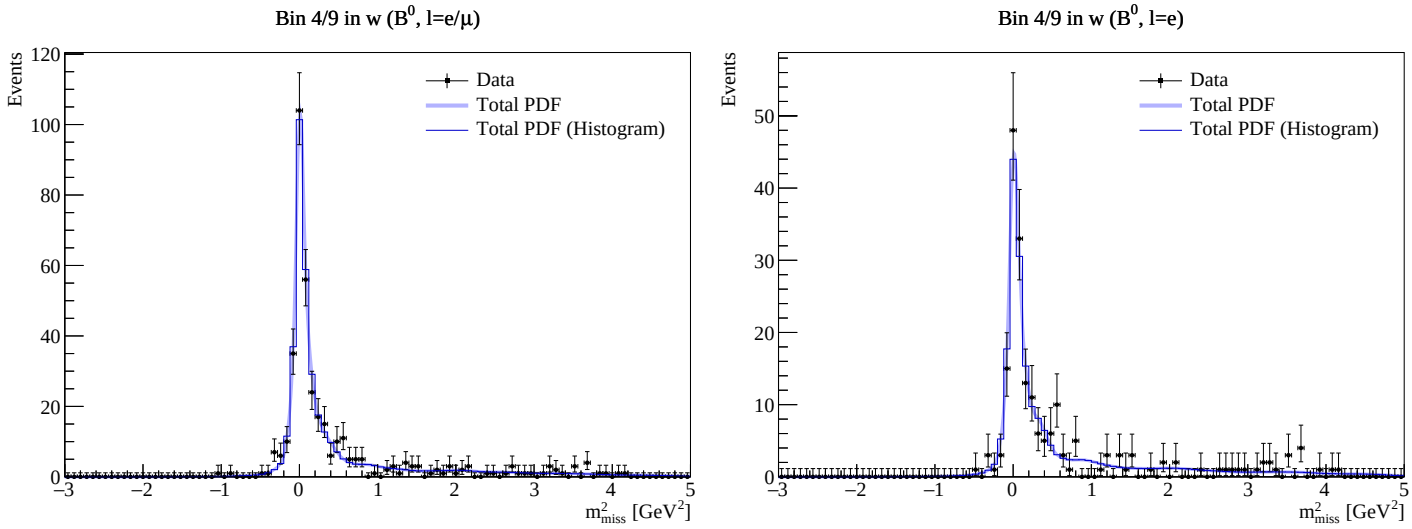


Figure 3.7: Examples for the fit of PDF templates to real data.

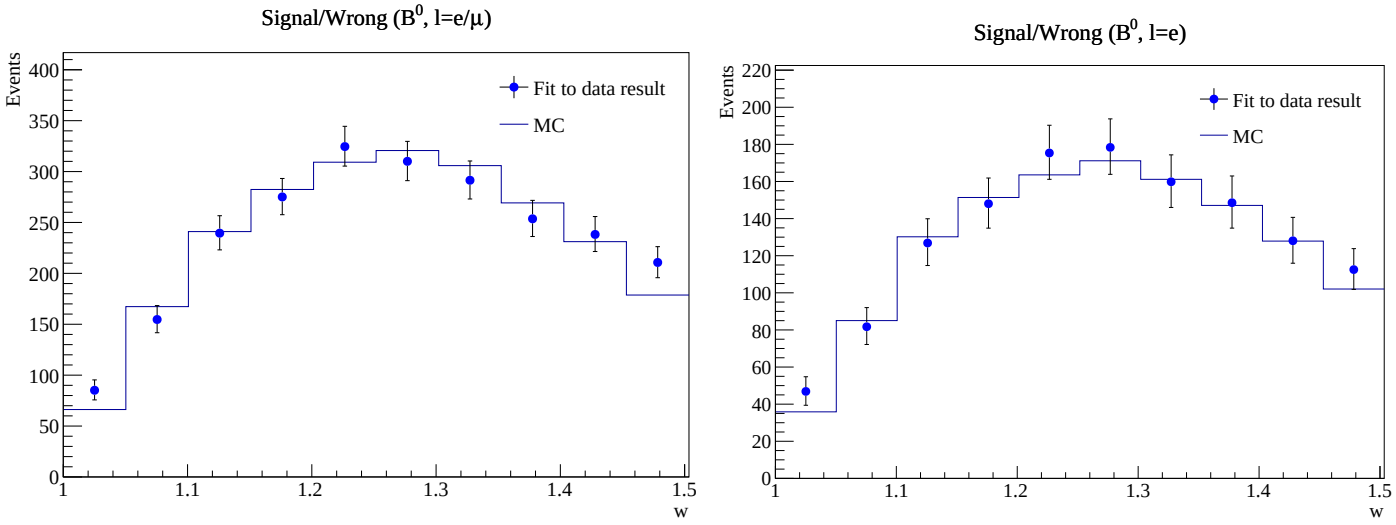


Figure 3.8: Exemplary fit results for w.

The pull distributions are considered separately for e, μ, $e+\mu$ and for B^0 and $B^\pm$. One of the 240 generated histograms is shown in figure 3.12. For an ideal unbiased fit, the distribution should approach a Gaussian with vanishing mean and unit standard deviation. While this is generally the case, most distributions show an enlarged negative tail, leading to a slightly negative mean of around -0.05 (1000 toys, $\ell = e/\mu$) and -0.03 (1000 toys, $\ell = e + \mu$).

Previous studies showed that this effect vanishes if the experiment is repeated with a 100 times higher number of generated toy events for each toy sample, making it seem likely that the observed bias is only due to non-Gaussian errors caused by low statistics.

To make it easier to spot possible biases, overview plots of the mean and the standard deviation of the distributions, as well as the χ^2/n.d.f. value of the fit with Gaussians, are shown in plots such as 3.13. The full set of plots can be found on pages 123 to 124.

3.5.2 Universality of the PDFs

For this closure test ("closure test 1"), subsets of the MC data are fitted using PDFs generated from the complementary MC subset. The resulting yields are compared and are expected to be

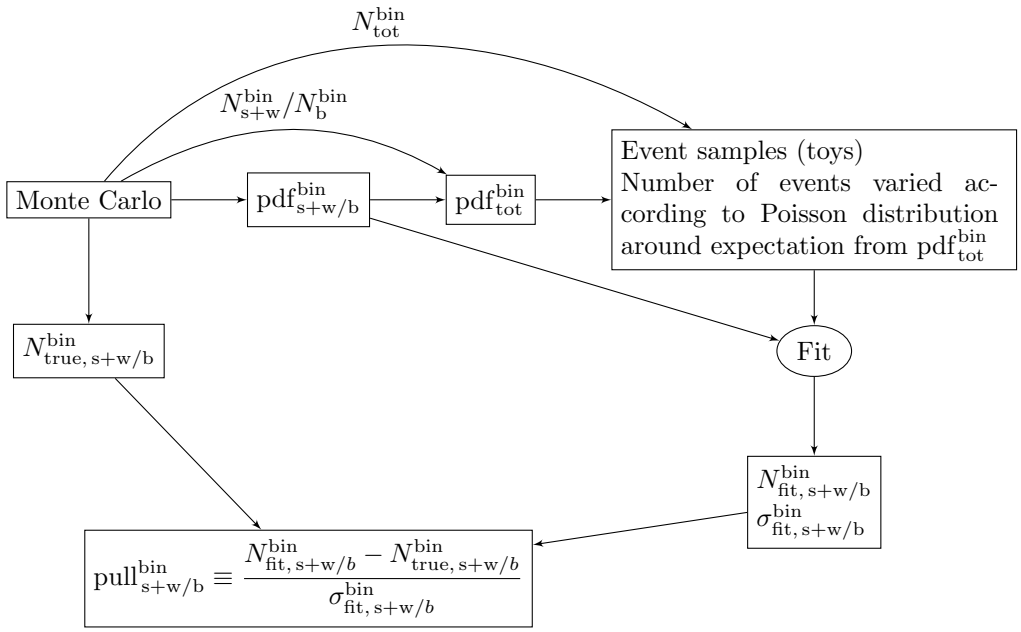


Figure 3.9: Flowchart for the closure test "closure toys".

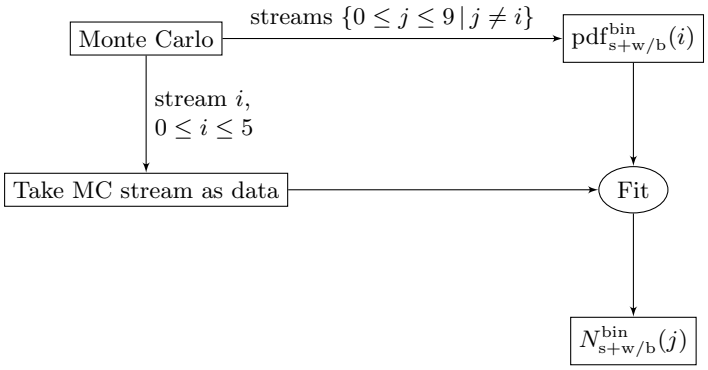


Figure 3.10: Flowchart for the closure test "closure test 1".

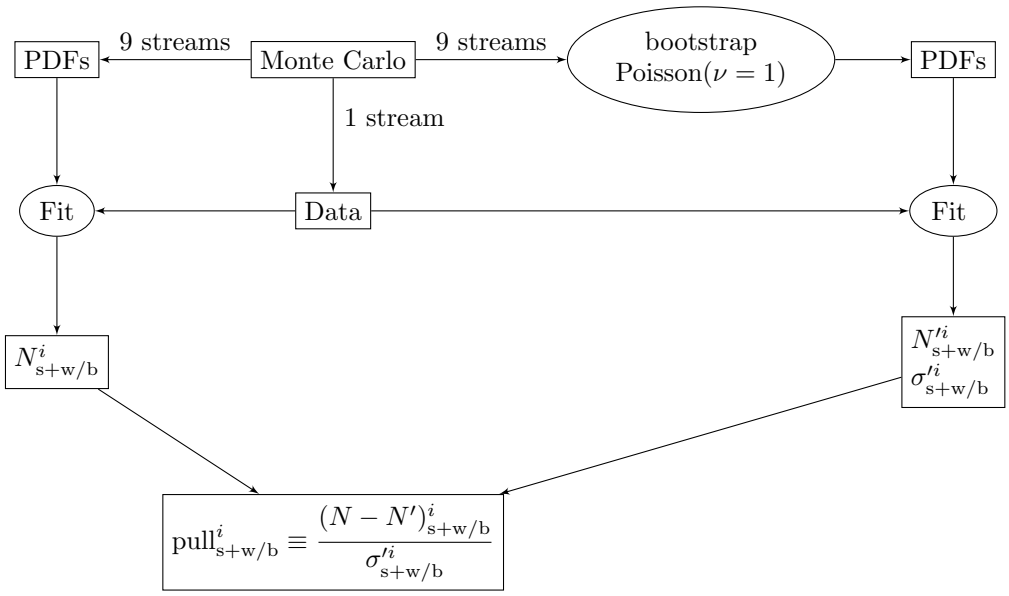


Figure 3.11: Flowchart for the closure test "closure test 2".

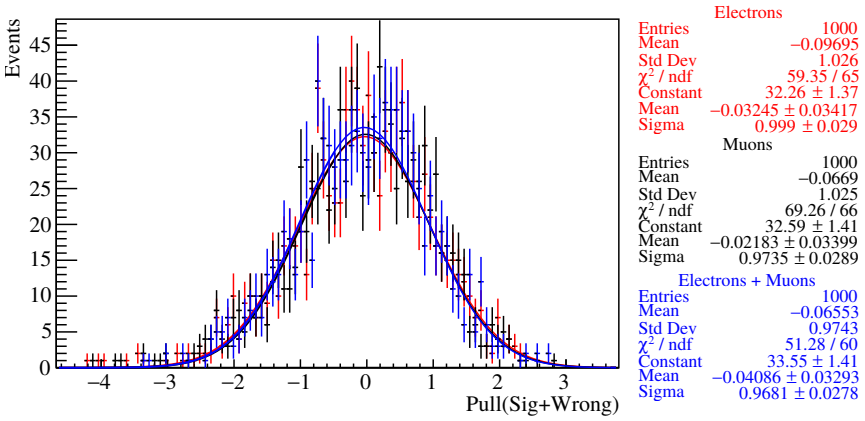
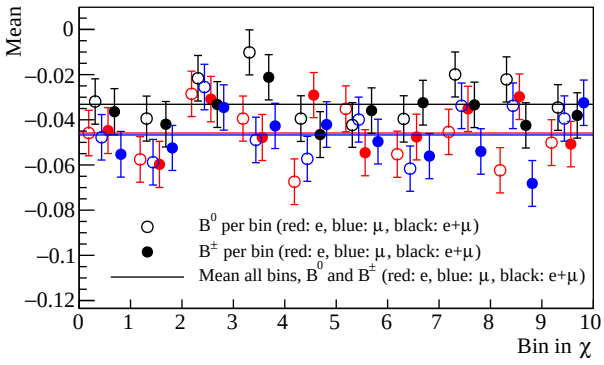
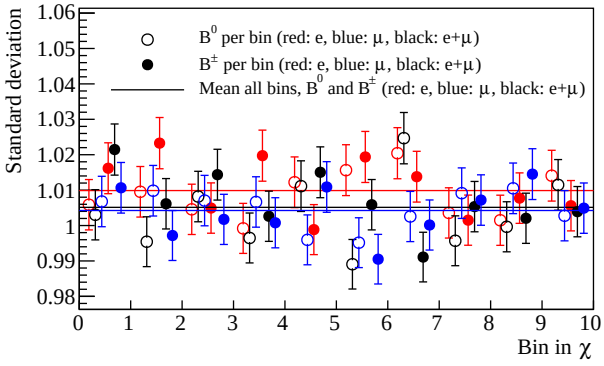


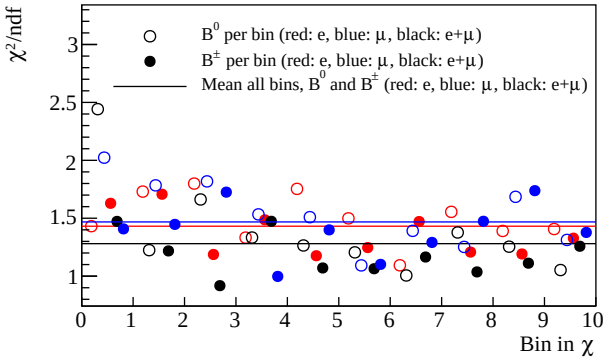
Figure 3.12: Pull of toy experiments for bin 0 in χ (B^0).



(a) Mean



(b) Standard deviation



(c) χ^2/n.d.f. for the fit of the Gaussian to the histogram

Figure 3.13: Overview of the shape of the pull distributions of the toys for the kinematic variable χ. Note: To facilitate readability, the markers for electrons, muons and the combined fit are slightly shifted. A comparison of all kinematic variables can be found on page 123 to 124.

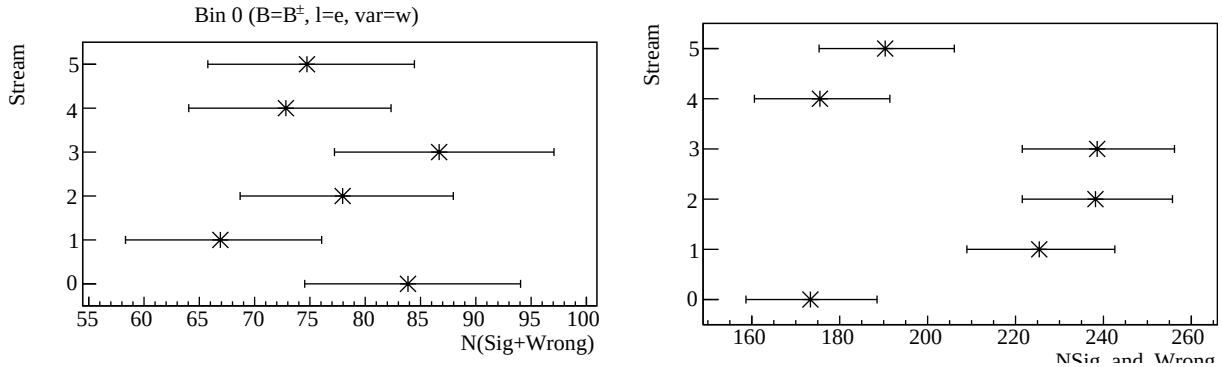
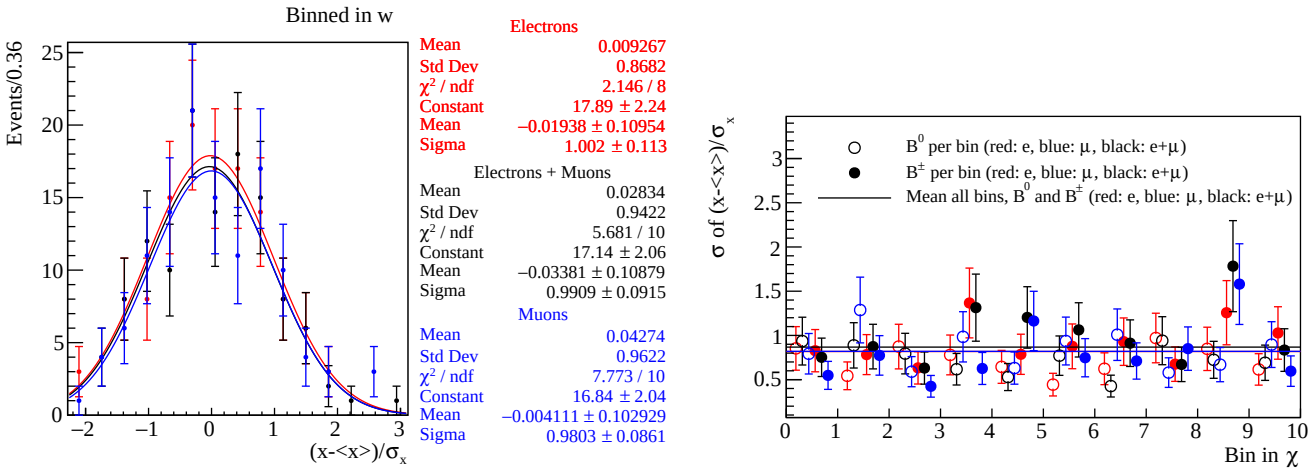


Figure 3.14: Closure test 1 for bin 0 in w ($B^\pm$).



(a) Combined histogram χ (B^0 and $B^\pm$ and all bins combined).

(b) Standard deviation of the 6 pull results in each bin for each B type and lepton type

Figure 3.15: Two examples of overview plots for "closure test 1". The full set of plots can be found on page 126.

compatible within the error margins.

The MC data is divided into 10 streams. Four of those (MC stream 5 to 9) do not contain $B\bar{B}$ events, leaving six "complete" streams (0 to 5). Taking stream $0 \leq j \leq 5$ as data, PDFs are calculated using the remaining nine streams. The data stream is then fitted with these PDFs, resulting in yields $N^i_{\mathrm{s+w/b}}(j)$ and corresponding errors $\sigma^i_{\mathrm{s+w/b}}(j)$. A visualization for two examples is shown in figure 3.14. To quantify the results, a pull quantity similar to equation (3.5.2) is introduced, taking the average of the results as the expectation value:

$$\mathrm{pull}^i_{\mathrm{s+w/b}}(j) := \frac{N^i_{\mathrm{s+w/b}}(j) - \frac{1}{6}\sum_{j'=1}^{6} N^i_{\mathrm{s+w/b}}(j')}{\sigma^i_{\mathrm{s+w/b}}(j)}. \tag{3.5.3}$$

Again, the pull distribution should resemble a Gaussian with unit standard deviation around zero. As only six values per bin are available, this is hard to test on a bin-to-bin basis. Nonetheless, plots such as 3.15b have been generated and checked. Additional fits to the combined pull distributions of all bins and both B^0 and $B^\pm$ show that the 60 pull values indeed form a Gaussian distribution (figure 3.15a). The full set of plots can be found on page 126.

3.5.3 Influence of the PDF shape uncertainty

The test "closure test 2" compares fit results obtained using two different PDFs corresponding to slightly varied MC datasets in order to quantify the effect of the uncertainty on the exact PDF shapes.

As in the last closure test, one complete stream (stream $0 \leq i \leq 5$) of the MC dataset is taken as data to be fitted and the remaining nine streams are used to build PDFs. This is done in two different ways. For the first result, the nine streams are fitted as they are ($\mathrm{pdf}^i_{\mathrm{s+w/b}}$). For the other results, each event in the nine streams is reweighted with Poisson weights of unit expectation value, before fitting the PDFs ($\mathrm{pdf}'^i_{\mathrm{s+w/b}}$). The fit results $N^i_{\mathrm{s+w/b}}$ and $N'^i_{\mathrm{s+w/b}}$ to the respective PDFs are then used to define the pull:

$$\mathrm{pull}^i_{\mathrm{s+w/b}} := \frac{N^i_{\mathrm{s+w/b}} - N'^i_{\mathrm{s+w/b}}}{\sigma'^i_{\mathrm{s+w/b}}}, \tag{3.5.4}$$

Running this repeatedly with different seeds for the generation of the Poisson weights results in a distribution $N'^i_{\mathrm{s+w/b}}$ and thus $\mathrm{pull}^i_{\mathrm{s+w/b}}$, which again should converge towards a Gaussian with vanishing mean. This time, however, the standard deviation is not expected to be close to one: Rather, it is taken as an estimator for the error introduced by the uncertainty on the PDF shape.

Initial runs of "closure test 2" for separate fits of electrons and muons lead to a number of distributions that looked like cut-off Gaussians. This could be traced back to the fit values being constrained to positive values. For separate fits to electrons and muons, the low number of background events can lead to an almost vanishing background yield in some bins, resulting in a distribution with a clear cut-off at zero. Since signal+wrong and background event yields together are fixed by the total number of events considered, this means that the signal yield and thus the pull distribution were similarly cut off towards the right side. Relaxing the lower bound on signal+wrong and background yields resolved this issue completely.

3.6 Correlation and Covariance between Kinematic Variables

Since the events in bin i of kinematic variable v and in bin j of kinematic variable $v \neq w$ overlap, the fitted event yields $N^{v,i}_{\mathrm{s+w}}$ and $N^{w,j}_{\mathrm{s+w}}$ are correlated, leading to correlated statistical errors $E^{v,i}_{\mathrm{stat}}$.

Correlations between random variables X and Y with mean $\mu_{X/Y}$ and standard deviation $\sigma_{X/Y}$ can be quantified using the Pearson correlation coefficient:

$$\mathrm{Corr}(X,Y) = \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y}, \tag{3.6.1}$$

respectively for a sample:

$$\mathrm{Corr}(X,Y) = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i-\bar{x})^2}\sqrt{\sum_{i=1}^N (y_i-\bar{y})^2}}, \tag{3.6.2}$$

where $\bar{x} = \frac{1}{N}\sum_{i=1}^N x_i$ (and $\bar{y}$ analogous). To generate a sample for the fit results $N^{v,i}_{\mathrm{s+w}}$ and $N^{w,j}_{\mathrm{s+w}}$, we generate toy samples by "bootstraping" the data events, i.e. applying a Poisson weight with unit expectation value to each event in the dataset. Repeating that procedure with N_{toys} different random seeds, we generate N_{toys} toy data sets, which are fitted and then used to calculate the correlation via equation (3.6.2).

The errors on the correlation coefficients $\rho := \mathrm{Corr}(E^{v,i}_{\mathrm{stat}}, E^{w,j}_{\mathrm{stat}})$ are estimated as[9]

$$\Delta\rho = \tanh\Big(\tanh^{-1}\rho + (N_{\mathrm{toys}}-3)^{-1/2}\Big) - \rho. \tag{3.6.3}$$

[9] In this context, $\tanh^{-1}$ is called a Fisher transformation. The Fisher transformation of the correlation of two normally distributed random variables X and Y, $\tanh^{-1}\rho(X,Y)$ is normally distributed with standard deviation $(N-3)^{-1/2}$, where N is the sample size. Applying tanh for the inverse transformation and subtracting the mean value ρ thus gives the 1σ error of ρ as stated.

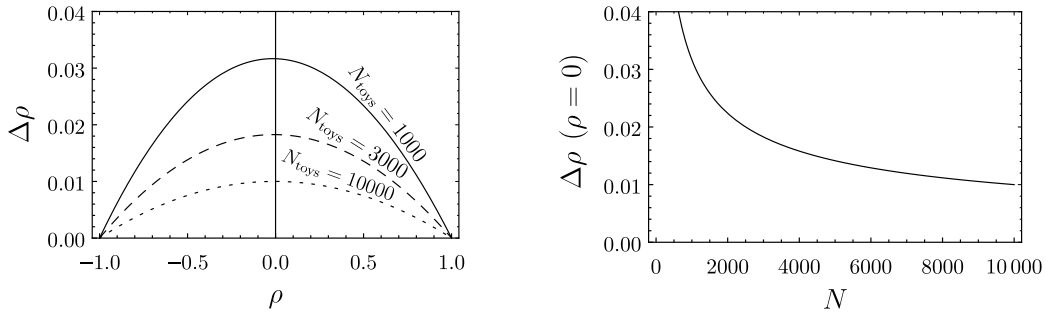


Figure 3.16: Uncertainty on the correlation coefficient.

A plot of the uncertainty $\Delta\rho$ in relation to ρ is shown in figure 3.16. The uncertainty is highest for vanishing correlation and vanishes for maximum correlation.

In the current setup, the correlation is calculated with $N_{\text{toys}} = 1000$, corresponding to an error $\Delta\rho = 0.03$ on most non-diagonal entries, because the correlation is less than 20% for all values. More specifically, we set

$$\text{Corr}(N^{v,i}, N^{v',i'}) = \begin{cases} 1 & v = v' \text{ and } i = i' \\ 0 & v = v' \text{ and } i \neq i' \\ 0 & v \neq v' \text{ and } \rho < 2\Delta\rho \text{ or } \rho < 0, \\ \rho & \text{else} \end{cases} \tag{3.6.4}$$

where ρ is the calculated correlation. This uses the a priori fact that, since the fit results for different bins of the same kinematic variable are obtained using disjoint sets of data and MC (namely projections based on disjoint (reco) value ranges of the kinematic variable), the corresponding correlation coefficients must vanish. $\text{Corr}(N^{v,i}, N^{v',i'}) = 1$ is obvious and negative correlations are unphysical,[10] thus are due to the uncertainty and can be set to zero.

An example of a correlation matrix obtained in this fashion is shown in figure C.5. The average of the non-diagonal elements of the matrices obtained for $B^\pm$ or B^0 and for $\ell = e$, $\ell = \mu$ or $\ell = e + \mu$ are nearly identical 6%. The correlation can reach up to 20% for single bins.

From the correlation of the statistical errors, we can calculate the corresponding covariance matrix as

$$\text{Cov}(E^{v,i}_{\text{stat}}, E^{v',i'}_{\text{stat}}) = E^{v,i}_{\text{stat}} E^{v',i'}_{\text{stat}}\, \text{Corr}(E^{v,i}_{\text{stat}}, E^{v',i'}_{\text{stat}}) \tag{3.6.5}$$

3.7 Migration Matrices

Limited resolution and accuracy in the measurement of a kinematic variable v (w, $\cos\theta_\ell$, $\cos\theta_{D^*}$, χ) can lead to events being mistakenly put in bin i_{reco} even though they belong in a different bin i_{true} (this is called "migration"). This effect is summarized in the *migration matrix* $\mathcal{M}^v_{\text{mig}}$:

$$\mathcal{M}^v_{\text{mig}} = \begin{pmatrix} N(1_{\text{reco}}, 1_{\text{true}}) & \cdots & N(1_{\text{reco}}, n_{\text{true}}) \\ \vdots & \ddots & \vdots \\ N(n_{\text{reco}}, 1_{\text{true}}) & \cdots & N(n_{\text{reco}}, n_{\text{true}}) \end{pmatrix}, \tag{3.7.1}$$

where n is the number of bins in v and $N(i_{\text{reco}}, j_{\text{true}})$ is the number signal+wrong events that belong in bin j_{true} but are put in bin i_{reco} by the analysis. Oftentimes the migration matrix is

[10] The correlation between the variables is due to the "overlap", i.e. events that are both in bin i of variable v and in bin i' of v'. Thus, an increased yield $N^{v,i}$ leads to an also increased yield $N^{v',i'}$.

	B^0			$B^\pm$		
Variable	e	μ	$e+\mu$	e	μ	$e+\mu$
w	89.8	89.9	89.9	84.7	85.2	84.9
χ	73.5	73.4	73.5	62.4	62.2	62.3
$\cos\theta_{D^*}$	71.2	71.6	71.4	60.6	60.5	60.5
$\cos\theta_\ell$	85.2	85.7	85.4	81.6	83.2	82.3

Table 3.2: Average value $\frac{1}{10}\sum_{i=1}^{10}(\mathcal{M}^v_{\text{mig}})_{ii}$ of the diagonal elements of each migration matrix.

presented with normalized columns. Because of $\sum_i(\mathcal{M}^v_{\text{mig}})_{ij} = N(j_{\text{true}})$, the number of events in bin j_{true}, this corresponds to $(\mathcal{M}^v_{\text{mig}})_{ij} := N(i_{\text{reco}}, j_{\text{true}})/N(j_{\text{true}})$.

Migration matrices are calculated from MC separately for each kinematic variable by performing a joint binning in bins of the reconstructed value of the kinematic variable and in bins of the true value of the kinematic variable.[11]

Distinguishing between B^0 and $B^\pm$ as well as between $\ell = e$, $\ell = \mu$ and $\ell = e+\mu$ leads to a total of 24 migration matrices. Two examples are shown in figure 3.17, the full set is shown on pages 132 to 135.

The average of the diagonal elements (table 3.2) corresponds to the fraction of correctly binned entries. The lowest migration is obtained for binnings in w (90% and 85% correctly binned events for B^0 and $B^\pm$, respectively) followed closely by $\cos\theta_\ell$, whereas binnings in χ and $\cos\theta_{D^*}$ are more problematic (70% and 60%).

In general, more migration is observed for the charged B mesons due to the higher fraction of wrongly reconstructed $\bar{B}\longrightarrow D^*\ell^-\bar{\nu}_\ell$ events. As the reconstruction of the D^* mostly affects the angles χ and θ_{D^*}, the differences between B^0 and $B^\pm$ are more pronounced in these variables (on average around 10% points difference per diagonal element), while ranging up to 5% points for w and θ_ℓ. The differences between the electron, muon and combined fit are very small, mostly ranging up to one percentage point for the diagonal entries.

3.8 Unfolding

In order to compare the results of the analysis with theory prediction and to fit $|V_{cb}|$ and form factor parameters, two strategies can be employed:

1. Folding of theory distributions: The predictions of the bin contents $N^{v,i,\text{theo}}_{\text{s+w}}$ are folded with the migration matrix, i.e. we consider $\tilde{N}^{v,i,\text{theo}}_{\text{s+w}} := \sum_j (M^v_{\text{mig}})_{ij}\, N^{v,j,\text{theo}}_{\text{s+w}}$ and directly compare $\tilde{N}^{v,i,\text{theo}}_{\text{s+w}}$ with the measured distribution $N^{v,i,\text{exp}}_{\text{s+w}}$.

2. Unfolding of measured distributions: Here, we try to undo the effect of the bin migration ("unfolding"), for example by considering $\tilde{N}^{v,i,\text{exp}}_{\text{s+w}} := \sum_j (\mathcal{M}^v_{\text{mig}})^{-1}_{ij}\, N^{v,j,\text{exp}}_{\text{s+w}}$. The unfolded measured distribution can then be directly compared to theoretical predictions.

Generally, the first method is conceptually simpler, less model dependent and does not need a careful treatment of numerical instabilities. Thus, for the fit to V_{cb} and the form factors performed in this analysis, the first strategy is applied (not discussed in detail in this thesis). However, to preserve the results in an experiment-independent form that allows for combinations

[11] For this, the signal+wrong events are selected by MC truth information, rather than a fitting procedure.

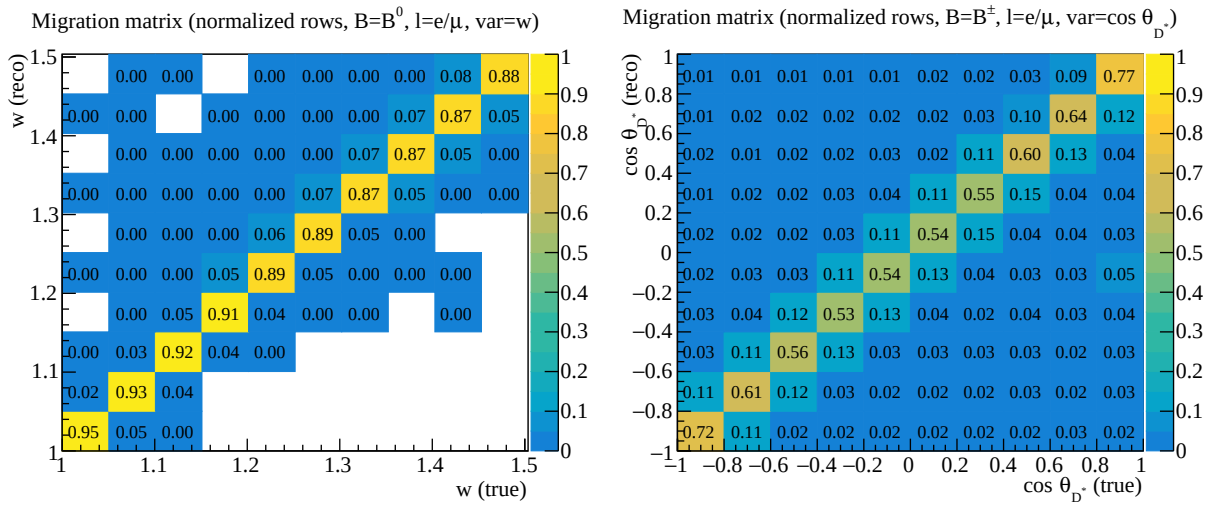


Figure 3.17: The migration matrix with lowest migration (B^0, $l = e + \mu$, w) and with highest migration ($B^{\pm}$, $l = e + \mu$, $\cos\theta_{D^*}$). The full set of migration matrices is shown on page 132 to 135.

and comparisons with other measurements, as well as for independent fits using different models, the unfolded measured distributions are needed.

3.8.1 SVD unfolding

3.8.1.1 Motivation

The problem with a simple unfolding via matrix inversion $\tilde{N}^{v,\mathrm{exp},i}_{\mathrm{s+w}} := \sum_j (\mathcal{M}^{v}_{\mathrm{mig}})^{-1}_{ij} N^{v,\mathrm{exp},j}_{\mathrm{s+w}}$, are numerical instabilities that lead to an amplification of uncertainties ("oscillations"). To address this issue, a method employing singular value decompositions (SVDs) is used. This method, *SVD Unfolding* is described in an application oriented way in [150].

3.8.1.2 Implementation

The method is implemented in `ROOT` as `TSVDUnfold` [151] and takes the following input:

- Data distribution (the distribution to unfold)
- Covariance matrix of the data distribution "pre-unfold covariance matrix", see section 3.10.3
- MC truth distribution
- MC reco distribution: In fact, it is not clear from the original paper [150], why this parameter is necessary and inspection of the source code has shown that this parameter is not used. This has also been tested with the corresponding tutorial examples.
- Migration matrix as calculated in section 3.7 (unnormalized, i.e. containing the event yields)

Note that, in principle, truth and reco MC distributions are the column and row sums of the migration matrix, thus the corresponding parameters can be considered to be superfluous for basic applications as well.

3.8.1.3 Theory of SVD unfolding

In the SVD unfolding procedure, the problem is first rotated and scaled to a linear equation $\tilde{A}w = \tilde{b}$ with $\mathrm{Cov}(b_i, b_j) = \delta_{ij}$, where (employing the notation from [150]) $\tilde{b}$ is the transformed measured spectrum $N_{\mathrm{s+w}}^{v,\mathrm{exp},j}$, w is the ratio of the unfolded spectrum (to be determined) to truth MC (per bin) and $\tilde{A}$ is the transformed unnormalized (!) migration matrix. This transformation makes sure that highly correlated elements of the input vector and elements with high uncertainty do not disproportionately influence the results.

Solving this matrix equation for w corresponds to a minimization of the squared differences $(\tilde{A}w - \tilde{b})^T(\tilde{A}w - \tilde{b})$. In order to suppress oscillations due to numerical instabilities, we add a *regularization term* to the minimization problem:

$$(\tilde{A}w - \tilde{b})^T(\tilde{A}w - \tilde{b}) + \tau \cdot (Cw)^T Cw = \text{min.}, \tag{3.8.1}$$

where the parameter $\tau \in \mathbb{R}$ controls the strength of the suppression and the symmetric matrix C is a discretization of the second derivative (albeit slightly modified in order to be non-singular).[12] The sum of squared second derivatives of the solutions, $(Cw)^T Cw$, is a measure of the "curvature" of the resulting distribution w, i.e. by adding this regularization term, we reward *smooth* solution histograms w with little bin-to-bin variations (which is a heuristically motivated a priori assumption on the shape of the true solution distribution).

In fact, the solution of (3.8.1) can be determined analytically and studying the influence of τ on the solution is instructive to motivate the optimal choice of τ. Finding the minimum of (3.8.1) is equivalent to the equation

$$\begin{pmatrix} \tilde{A}C^{-1} \\ \mathbb{1}\sqrt{\tau} \end{pmatrix} Cw = \begin{pmatrix} \tilde{b} \\ 0 \end{pmatrix}. \tag{3.8.2}$$

Of course, (3.8.2) could be directly solved by $w = \tilde{A}^{-1}\tilde{b}$, but the effect of $\tau \neq 0$ can be better understood by considering the singular value decomposition of $\tilde{A}C^{-1}$:

$$\tilde{A}C^{-1} = USV^{\mathrm{T}}, \tag{3.8.3}$$

where U and V are orthogonal matrices and $S = \mathrm{diag}(s_1, \cdots, s_n)$ is a diagonal matrix with $s_i \geq 0$. Note that for $\tau = 0$, the solution of (3.8.2) is given as $(V^{\mathrm{T}}Cw)_i = (U^{\mathrm{T}}\tilde{b})_i/s_i$ and numerical instabilities correspond to the small *singular values* s_i. In general, the solution of (3.8.2) takes the similar form

$$(V^{\mathrm{T}}Cw)_i^{(\tau)} = (U^{\mathrm{T}}\tilde{b})_i \cdot \frac{s_i}{s_i^2 + \tau}. \tag{3.8.4}$$

From this, it is easy to see that an appropriately sized parameter τ "regularizes" small singular values, i.e. dampens their influence strongly, while leaving the contributions with higher s_i relatively unchanged. This is similar to the conceptually simpler, but also more aggressive approach of *dimensional reduction*, where information from the dimensions associated with small singular values is simply discarded.

The central question is the choice of the value for τ. Following [150], we assume without loss of generality (after possible rearrangement of the dimensions), that $s_1 \geq s_2 \geq \cdots \geq s_n$ and set $\tau = s_k^2$ with $1 \leq k \leq n$.

The choice of the value k (and more generally τ) will always be a compromise between a low error on the result and a low model dependency:

[12] If $\tilde{C}$ is the matrix corresponding to the sum of squares of discretized second derivatives $(Cw)^{\mathrm{T}}Cw = \sum_i[(w_{i+1} - w_i) - (w_i - w_{i-1})]^2$, then we consider $C = \tilde{C} + \xi\mathbb{1}$, where $\mathbb{1}$ is the unit matrix and ξ is a positive value, usually between 10^{-3} and 10^{-4}.

- Low k (high τ) corresponds to strong dampening of small and not-so-small singular values. While this reduces numerical instabilities (oscillations) and thereby leads to small errors on the results, it also discards experimental information, i.e. introduces a bias towards MC and thereby a strong model dependency.

  The fact that the results are biased *towards* MC, i.e. that the unfolded distribution will start to resemble the truth MC distribution more and more for high τ can be understood from remembering that w is the ratio of the unfolded values to the truth MC values. Thus, the regularization term $\tau \cdot (Cw)^{\mathrm{T}} Cw$ from (3.8.1) does not punish the curvature of the solution *itself*, but rather of the ratio to the truth MC, i.e. punishes non-linear deviations of the results from MC truth. As the regularization term scales with τ, this bias also increases with higher values of τ.

- High k (low τ) corresponds to weak dampening of small singular values. This amplifies the errors of the input measurement, i.e. leads to larger errors on the results.

Note that the choice $k = 1$ is not sensible, as this implies that even the first singular value s_i is cut in halves. Even in the best case, where all information is concentrated in s_1, i.e $s_{i\neq 1} = 0$, this means that only half of the expected yields are recovered. Therefore, only the values $2 \leq k \leq 10$ are considered in the following.

3.8.2 Model uncertainties

The choice of the regularization parameter k depends on the estimate of the model uncertainties. Even when unfolding data using matrix inversion, the results are always model dependent, because the migration matrix is generated from MC truth vs reco distributions whose shape is subject to uncertainties on parameters employed in the MC reweighting (see section 3.3.5). Therefore, we expect the model uncertainty vs k distribution to become flat for a sufficiently high $k \geq k_0$, i.e. expect that for $k > k_0$, only negligible improvements in model independence over $k = k_0$ can be achieved.

For a conservative estimate of the model uncertainty, the MC dataset is reweighted with 3σ variations of the three orthogonal form factor parameters and the corresponding MC truth and reco distributions $N^{v,i}_{\text{truth}}(p_i \pm \sigma_{p_i})$ and $N^{v,i}_{\text{reco}}(p_i \pm \sigma_{p_i})$ are calculated (by counting the number of MC signal+wrong events projected into bins of the truth and reco level variables[13]).

The reco distributions are then unfolded (with the covariance matrix $\mathcal{B}^{v,i}$ of the data, the original migration matrix $\mathcal{M}^{v}_{\text{mig}}$, MC truth $N^{v,i}_{\text{truth}}$ and reco distribution $N^{v,i}_{\text{reco}}$ as parameters) and the result is compared with the varied MC truth distribution. Explicitly, the model error is calculated as follows (where $\mathcal{U}$ is the unfolding algorithm, i.e. `TSVDUnfold`):

$$\begin{aligned} E^{v,i}_{\text{model}\pm} &:= \sqrt{\sum_{i=1}^{3}\Big[\mathcal{U}(N^{v,i}_{\text{reco}}(p_i \pm \sigma_{p_i}), \mathcal{B}^{v,i}, N^{v,i}_{\text{truth}}, N^{v,i}_{\text{reco}}, \mathcal{M}^{v,i}_{\text{mig}}) - N^{v,i}_{\text{truth}}(p_i \pm \sigma_{p_i})\Big]^2}, \\ E^{v,i}_{\text{model}} &= \max\Big\{E^{v,i}_{\text{model}+}, E^{v,i}_{\text{model}-}\Big\} \end{aligned} \tag{3.8.5}$$

The corresponding relative error is calculated in the same fashion, only by normalizing each of the differences to $N^{v,i}_{\text{truth}}(p_i \pm \sigma_i)$.

A plot of the absolute model uncertainty is shown in figure 3.21, the full set of plots is shown on page 136 to 139.

[13] Here, the signal+wrong events are selected based on MC truth information, i.e. without performing any fit.

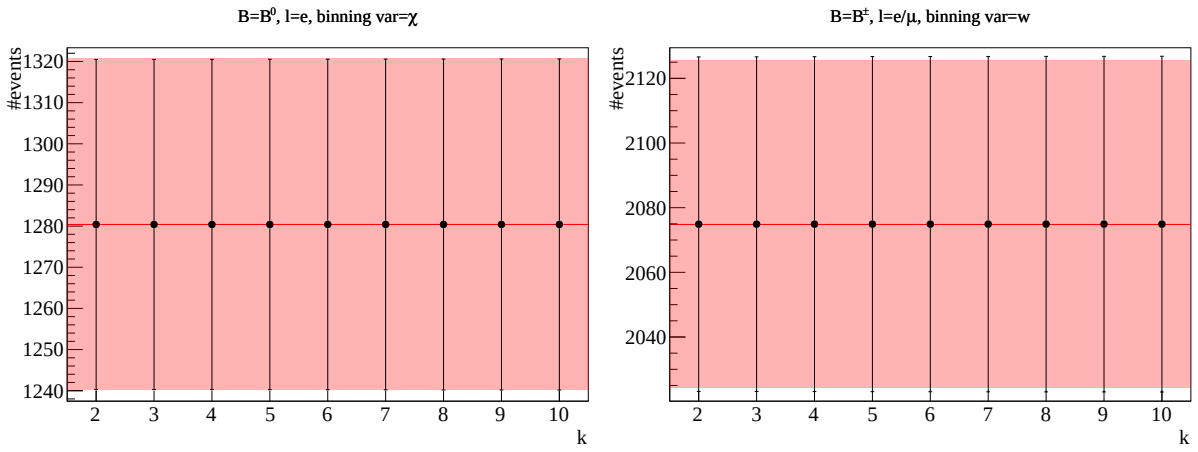


Figure 3.18: Two examples of errors on the total signal yield. The plot markers and vertical error bars correspond to the error on the integrated signal yield as calculated by `TSVDUnfold` with $k \geq 2$. The light red rectangle corresponds to the symmetric error calculated via the sum of the squared errors on the not-unfolded bin contents.

3.8.3 Propagation of integral error in the unfolding process

Because the unfolding procedure is in principle a linear transformation that leaves the total signal yield unchanged (and only changes the *shape* of the signal yield distribution), the error on the total signal yield is expected to stay (almost) invariant.[14] Two examples of a comparison of the integral errors are shown in figure 3.18. No significant deviation could be observed in any of the 22 other cases.

3.8.4 Unfolding of MC data

As a closure test, reco MC data is unfolded (only taking statistical errors into account). As the migration matrix is calculated by a joint binning of MC reco vs MC truth, i.e. $N^{v,\mathrm{reco},i}_{\mathrm{s+w}} = \sum_j (M^v_{\mathrm{mig}})_{ij} N^{v,\mathrm{truth},j}_{\mathrm{s+w}}$ (assuming normalized columns of $\mathcal{M}^v_{\mathrm{mig}}$), the unfolding procedure is expected to recover the truth distribution. This holds exactly for unfolding by matrix inversion and approximately for SVD unfolding, where the inverse of the migration matrix is only approximated.

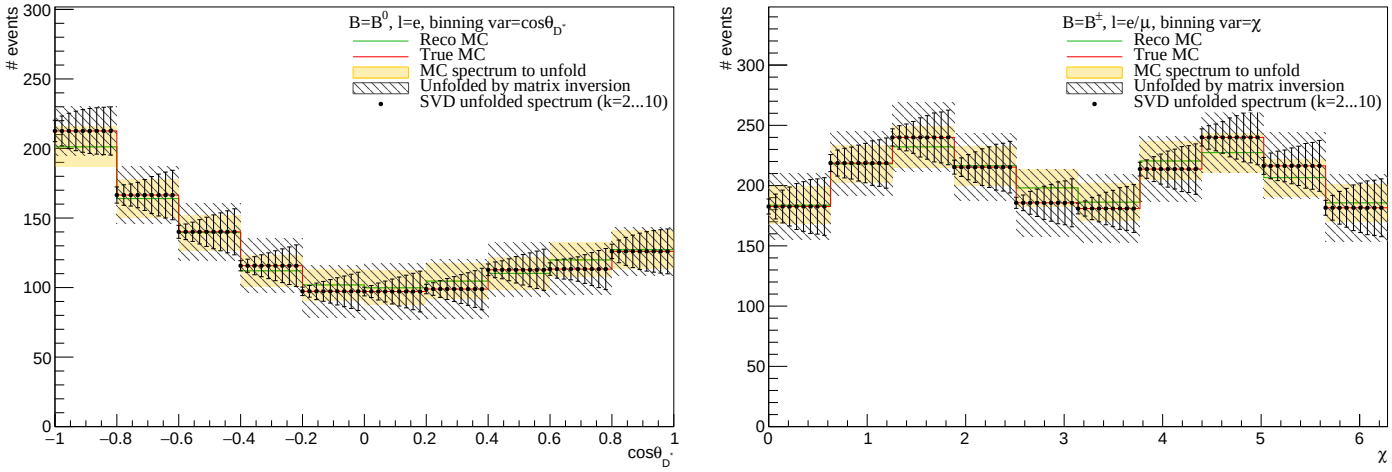
If most of the information is contained in the dimension corresponding to the highest singular value, the value of the regularization parameter $k \geq 2$ is expected to have little influence. This can, for example, be understood by considering the minimization of the term shown in equation (3.8.1): w is the ratio of the unfolded yields to truth MC yields, i.e. as soon as we come close to the MC truth distribution, $w \approx \mathrm{const} = 1$ and $(Cw)^{\mathrm{T}} Cw \approx 0$, i.e. the parameter τ becomes irrelevant.[15]

As is shown with two examples in figure 3.19a, both unfolding by matrix inversion and SVD unfolding for different values of $k \geq 2$ recover the truth MC distribution with high precision. Indeed, almost no dependency on k for $k \geq 2$ is visible and the statistical errors on the unfolded bin contents increase with k. This behavior holds true without exception for the remaining 22 plots as well.

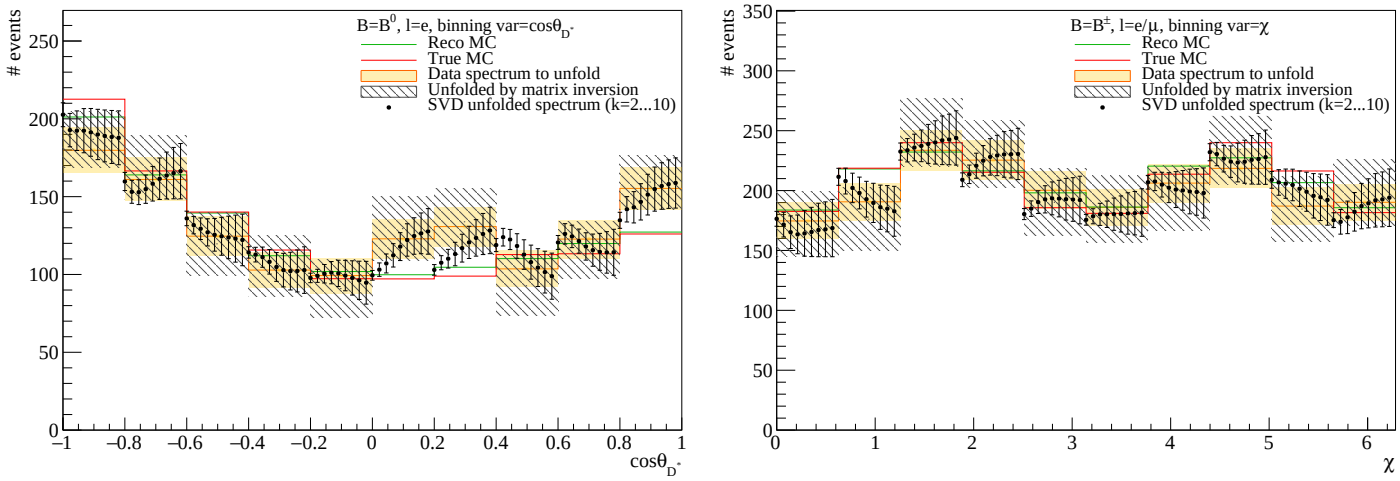
[14] For unfolding by matrix inversion this is exact and can be seen as follows: The migration matrix leaves the total yield invariant, i.e. $\sum_i (\mathcal{M}v)_i = \sum_j \left(\sum_i \mathcal{M}_{ij}\right) v_j \overset{!}{=} \sum_i v_i$, implying $\sum_i \mathcal{M}_{ij} = \mathbb{1}$. This property is inherited by the inverse $\mathcal{M}^{-1}$: We have $\sum_{i'} \mathcal{M}^{-1}_{ii'} \mathcal{M}_{i'j} = \delta_{ij}$, thus $1 = \sum_j \delta_{ij} = \sum_{i'} \sum_j \mathcal{M}^{-1}_{ii'} \mathcal{M}_{i'j} = \sum_{i'} \mathcal{M}^{-1}_{ii'}$.

Furthermore, $\mathrm{Cov}(\mathcal{M}v, \mathcal{M}v) = \mathcal{M}\,\mathrm{Cov}(v,v)\mathcal{M}^{\mathrm{T}}$ and $\mathrm{Var}\left(\sum_i v_i\right) = \sum_{ii'} \mathrm{Cov}(v_i, v_{i'})$, thus $\mathrm{Var}(\sum_i (\mathcal{M}v)_i) = \sum_{iji'j'} \mathcal{M}_{ii'}\,\mathrm{Cov}(v'_i, v_{j'})(\mathcal{M}^{\mathrm{T}})_{j'j} = \sum_{i'j'} \mathrm{Cov}(v'_i, v'_j) = \mathrm{Var}\left(\sum_i v_i\right)$. As the "inverse" of $\mathcal{M}$ as calculated in the SVD Unfolding algorithm approximates the real inverse $\mathcal{M}^{-1}$, the relation $\sum_{i'} \mathcal{M}^{-1}_{ii'}$ is still approximately true there.

[15] More accurately: For $w = w_0 = \mathrm{const}$, we have $(Cw)^{\mathrm{T}} Cw = n\xi w_0$.



(a) Unfolding of MC data



(b) Unfolding of real data

Figure 3.19: Unfolding of MC data (a) and real data (b). Both unfolded results using matrix inversion (hatched rectangles) and using SVD unfolding with different regularization parameters k (black data points with error bars) are shown. The data to be unfolded is shown with a golden histogram and a golden shaded rectangle for its errors. For comparison, reco MC data (green histogram) and truth MC data (red histogram) are shown as well. Note that for the unfolding of MC data, the reco MC histogram and the input data histogram are identical by definition and only the latter is visible..

As another quick consistency check regarding the comment made in section 3.8.1, one can also consider the distribution for $k = 1$. Here, the distribution shape is roughly recovered (which is good, because it means that the relevant information is indeed concentrated in the dimension associated with the highest singular value), while we miss the total yield by a factor of 2 as expected.

3.8.5 Unfolding of data

In contrast to the unfolding of MC data, the regularization parameter k also influences the central value of the unfolded measured data distribution.

This is demonstrated in figure 3.21. As expected a low k leads to a bias towards MC in most

cases, while for high k a trend towards the original input distribution is observed. The full set of unfolding plots for real data is shown on pages 140 to 143.

The influence of the parameter k is also investigated using toy experiments, where the input distribution is varied based on its covariance matrix (described in detail in section 3.10.4) and these "toy samples" are unfolded. An example is shown in figure 3.20. As a nice closure test, one can see that the average of the toys for tests with MC data also agrees perfectly with the truth distribution. Furthermore (for both data and MC), the standard deviation of the toy distributions agrees well with the error as calculated by `TSVDUnfold`.

3.8.6 Choice of regularization parameter k

An exemplary plot of the k dependency of the model uncertainty (described in section 3.8.2) is shown in figure 3.21. The full set of all 24 plots is shown on pages 136 to 139. For B^0, all plots show a downwards trend of the average error $\frac{1}{10}\sum_i$ model uncertainty(bin i), where the losses are large for small k and quickly decrease for $k \geq 5$. In particular for χ and θ_{D^*}, the distributions become flat for $k \geq 5$. Based on this observation, a value of $k = 5$ is chosen.

3.9 Efficiency Correction

So far, *signal yields* have been fitted and unfolded. To compare the kinematic distributions directly with theoretical results, these yields have to be converted to (binned) differential cross sections, taking into account the reconstruction efficiency.

We write $\Delta\Gamma_i$ for the integrated differential cross section corresponding to bin i in the kinematic variable v (χ, θ_ℓ, θ_{D^*}, w) with edges $v_i \leq v \leq v_{i+1}$, i.e.

$$\Delta\Gamma_i := \int_{v_i}^{v_{i+1}} \mathrm{d}v \, \frac{\mathrm{d}\Gamma}{\mathrm{d}v}. \tag{3.9.1}$$

Similarly, we also write $\Delta\mathrm{BR}_i$ for the binned differential branching ratios. Branching ratios can be converted into differential cross sections via $\hbar\tau_B^{-1}$:

$$\Delta\Gamma_i = \hbar\tau_B^{-1}\Delta\mathrm{BR}_i, \tag{3.9.2}$$

where τ_B is the mean life time of the B meson ($\tau_{B^\pm} = (1.638 \pm 0.004)\,\mathrm{ps}$ and $\tau_{B^0} = (1.520 \pm 0.004)\,\mathrm{ps}$ [5]) and $\hbar = 6.58 \times 10^{-25}\,\mathrm{GeV}\cdot\mathrm{s}$ is the reduced Planck constant. The event yields $N^{v,i}$ are related to the branching fraction as

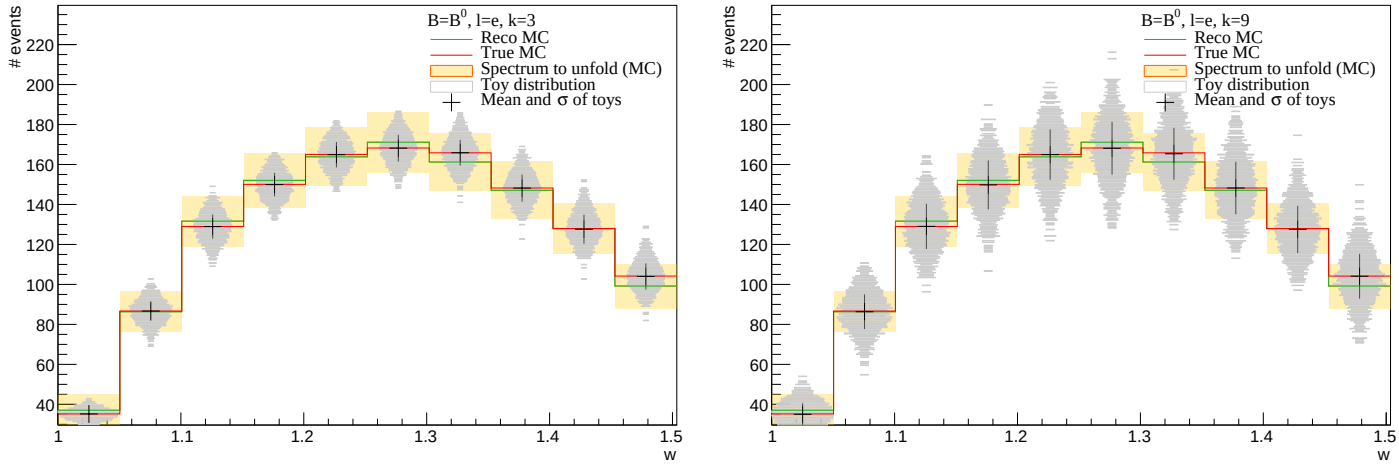
$$N_{\mathrm{s+w}}^{v,i} = 2n_f N_{B^+B^-/B^0\bar{B}^0}\Delta\mathrm{BR}_i, \tag{3.9.3}$$

where n_f is the number of lepton flavors we are considering (i.e. $n_f = 2$ for the lepton flavor agnostic fit and $n_f = 1$ for the separate fits) and the factor 2 is due to the conjugated decay modes of B^+/B^- and $B/\bar{B}$ (that is, because our results for $\bar{B}\longrightarrow D^*\ell^-\bar{\nu}_\ell$ are calculated as CP averages). The number of $B^\pm B^\mp$ and $B^0\bar{B}^0$ pairs, $N_{B^\pm B^\mp}$ and $N_{B^0\bar{B}^0}$, is related to the total number of $B\bar{B}$ pairs, $N_{B\bar{B}}$, using

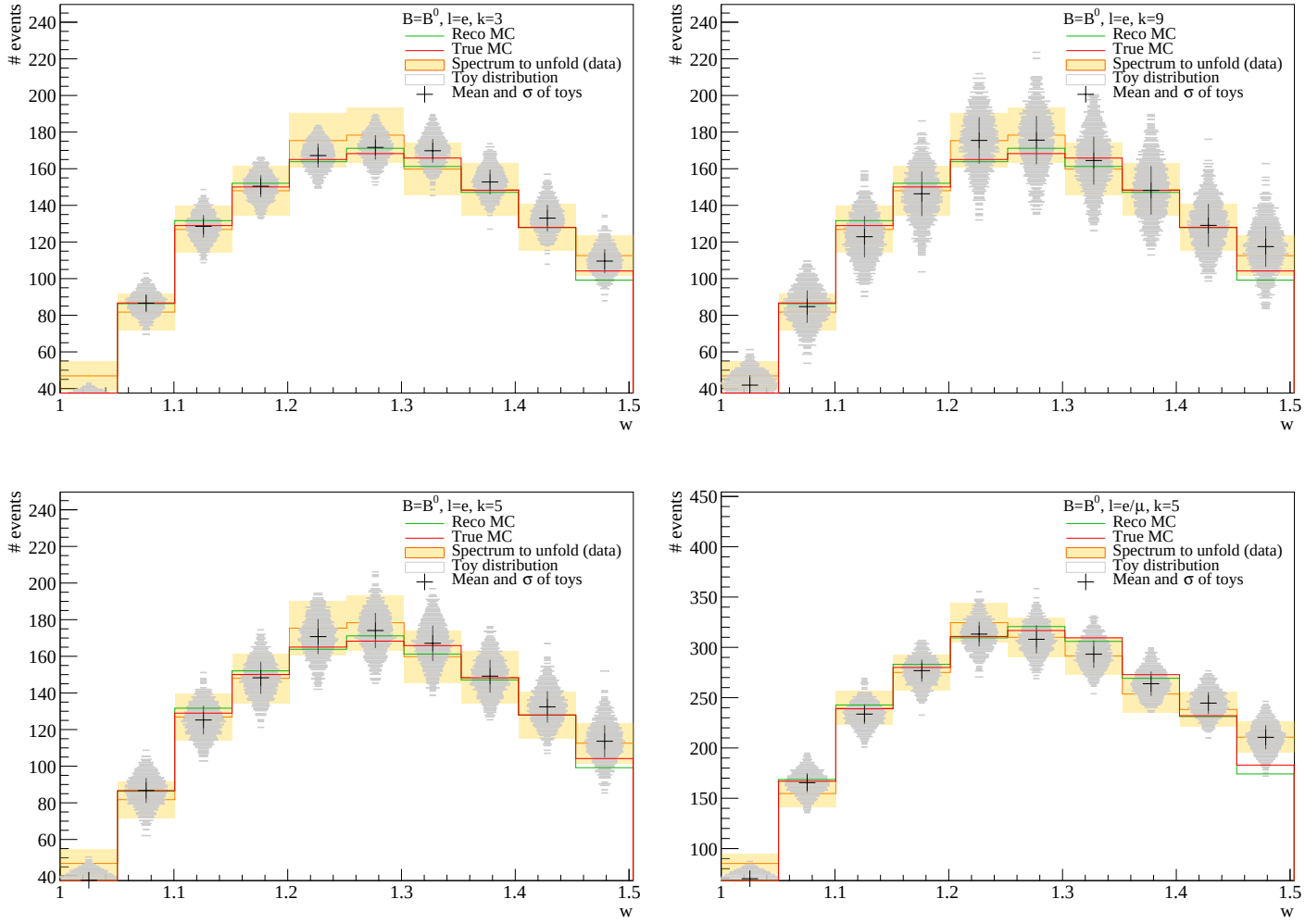
$$\begin{aligned} N_{B^0\bar{B}^0} = \frac{N_{B\bar{B}}}{1+f_{+0}} \quad &\Longleftrightarrow \quad N_{B^+B^-} = \frac{N_{B\bar{B}}}{1+f_{+0}^{-1}}, \\ \text{with} \quad f_{+0} := \frac{\mathrm{BR}(\Upsilon(4S)\longrightarrow B^+B^-)}{\mathrm{BR}(\Upsilon(4S)\longrightarrow B^0\bar{B}^0)} &= \frac{N_{B^+B^-}}{N_{B^0\bar{B}^0}}. \end{aligned} \tag{3.9.4}$$

The efficiency $\epsilon_{\mathrm{reco}}\epsilon_{\mathrm{tag}}$ can be determined by comparing the MC yield with the yield expectation as calculated with equation (3.9.3):

$$\epsilon_{\mathrm{reco}}\epsilon_{\mathrm{tag}}\big|_{\mathrm{MC}} = N_{\mathrm{s+w}}^{\mathrm{truth},v,i} \cdot \left[2n_f N_{B^+B^-/B^0\bar{B}^0}\hbar^{-1}\tau_B\Delta\Gamma_i\right]_{\mathrm{MC}}^{-1}, \tag{3.9.5}$$



(a) Unfolding of toys based on MC data



(b) Unfolding of toys based on real data

Figure 3.20: Unfolding of toy distributions and the influence of the regularization parameter k

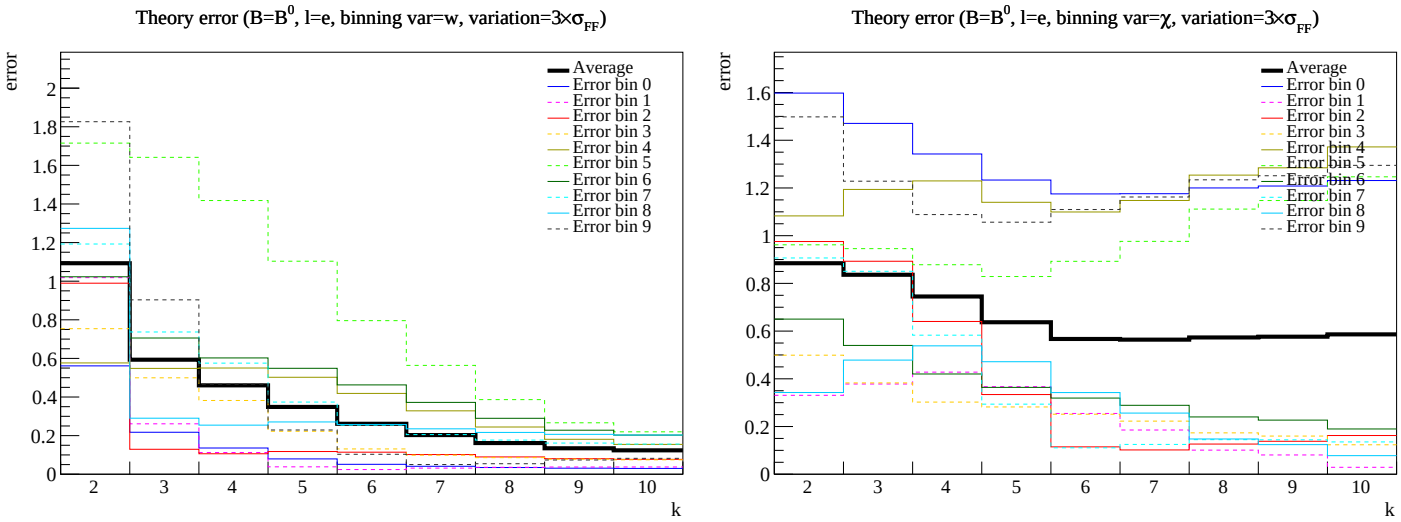


Figure 3.21: Model uncertainty for $2 \leq k \leq 10$. The full set of figures for all 24 cases is given on page 136 to 139.

B	ℓ	$N_{\text{s+w}}^{\text{data}}$	$N_{\text{s+w}}^{\text{MC}}$	$\epsilon_{\text{reco}}\epsilon_{\text{tag}}$
B^0	$e+\mu$	2367 ± 53	2377.1	3.19×10^{-5}
	e	1306 ± 40	1280.4	3.44×10^{-5}
	μ	1066 ± 34	1096.7	2.95×10^{-5}
$B^\pm$	$e+\mu$	2013 ± 50	2075.0	2.59×10^{-5}
	e	1068 ± 37	1110.4	2.77×10^{-5}
	μ	947 ± 33	964.5	2.41×10^{-5}

Table 3.3: Signal yields in data and (truth) MC and the efficiency $\epsilon_{\text{reco}}\epsilon_{\text{tag}}$.

where $N_{\text{truth}}^{v,i}$ is the truth MC yield in bin i of variable v (i.e. the number of events projected into bin i based on the truth value of v) and all parameters of the expression $[\cdots]_{\text{MC}}$ are calculated using the same parameters that were used to generate/reweigh the MC data, ensuring the model independence of $\epsilon_{\text{reco}}\epsilon_{\text{tag}}$ (in particular, we do not need to use f_{+0} here). Efficiency values obtained from the inclusive fit are shown in table 3.3. Figure 3.22 shows examples of the kinematic dependencies of the efficiencies. The full set of plots is shown on page 144.

Once the efficiency has been calculated, the efficiency corrected unfolded event yields are given by

$$N_{\text{s+w}}^{\text{e.corr.},v,i} = [\epsilon_{\text{reco}}\epsilon_{\text{tag}}]_{\text{MC}}^{-1} N_{\text{s+w}}^{\text{unf.},v,i}, \tag{3.9.6}$$

which we convert to the binned differential decay width via

$$\Delta\Gamma_i\Big|_{\text{final}}^{\text{meas.}} \overset{\substack{(3.9.3)\\(3.9.2)}}{=} N_{\text{s+w}}^{\text{e.corr.},v,i} \cdot \frac{\hbar\tau_B^{-1}\Big|_{\text{pdg}}}{2n_f\, N_{B^+B^-/B^0\bar{B}^0}\Big|_{\text{data}}}, \tag{3.9.7}$$

where the parameters of $\cdots|_{\text{pdg}}$ are the current world averages and $\cdots|_{\text{data}}$ refers to the measured values in real data. Combining equations (3.9.7), (3.9.6) and (3.9.5), the expression for our final results is (where we canceled constants modeled precisely in MC):

$$\Delta\Gamma_i\Big|_{\text{final}}^{\text{meas.}} = \frac{N_{\text{s+w}}^{\text{unf.},v,i}}{N_{\text{s+w}}^{\text{truth},v,i}} \cdot \Delta\Gamma_i|_{\text{MC}} \cdot \frac{N_{B^+B^-/B^0\bar{B}^0}\Big|_{\text{MC}}}{N_{B\bar{B}}|_{\text{data}}} \cdot \begin{cases} [1+f_{+0}]_{\text{pdg}} & B^0\bar{B}^0 \\ [1+f_{+0}^{-1}]_{\text{pdg}} & B^\pm B^\mp, \end{cases} \tag{3.9.8}$$

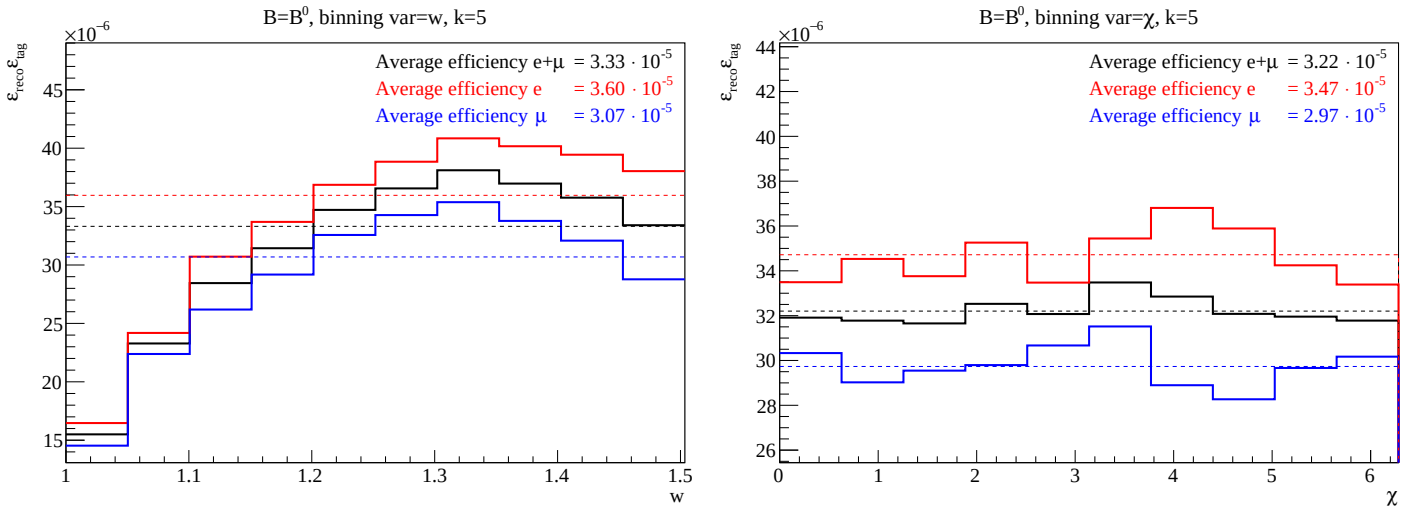


Figure 3.22: Efficiency $\epsilon_{\text{reco}}\epsilon_{\text{tag}}$ per bin for electrons (red), muons (blue) and the combined fit (black). The dashed line corresponds to the average efficiency (calculated with event yields as weights). The full set of plots is shown on page 144.

With $N_{B\bar{B}} = (772 \pm 11) \times 10^6$, $N_{B^+B^-/B^0\bar{B}^0}\big|_{\text{MC}} = 377.43 \times 10^6$ and the 2018 world average of $f_{+0} = 1.058 \pm 0.024$ [5], the correction term in equation (3.9.8) is calculated to be

$$\frac{N_{B^+B^-/B^0\bar{B}^0}\big|_{\text{MC}}}{N_{B\bar{B}}|_{\text{data}}} \cdot \begin{cases} [1+f_{+0}]_{\text{pdg}} & B^0\bar{B}^0 \\ [1+f_{+0}^{-1}]_{\text{pdg}} & B^\pm B^\mp \end{cases} = \begin{cases} 1.006 & B^0\bar{B}^0 \\ 0.951 & B^\pm B^\mp. \end{cases} \tag{3.9.9}$$

As a closure test, the efficiency corrected unfolded reco MC results converted to binned differential decay widths can be compared with the binned theory calculation of the differential decay width as used for the MC reweighting. The unfolded reco MC distributions perfectly matched the MC truth distributions (figure 3.19a), thus it is expected that both differential cross sections as calculated above agree as well. This can be confirmed perfectly as shown in two examples in figure 3.23.

3.10 Systematics

3.10.1 PDF shape uncertainty

Limited statistics in MC leads to an uncertainty on the fitted shape of the PDF for the m^2_{miss} distribution of signal+wrong and especially of background (due to the lower number of simulated events). This uncertainty also depends on the choice of the smoothing parameter of the Gaussian kernel estimation. Thus, the uncertainty induced in the fitted yields $N^i_{\text{s+w/b}}$ has to be evaluated. For this, the same setup as for closure test 2 is used, only that real data is taken instead of one of the MC streams (and all ten streams of MC are taken together to fit the PDFs).

The PDF shape uncertainty error is correlated between bins of different kinematic variables. Using a common set of seeds for the bootstrapping of the toys for different kinematic variables, the correlation of the error can be calculated via the sample correlation as described in 3.6.

However, in the original code, this method was abandoned due to technical difficulties.[16] Instead,

[16] Due to technical reasons, the order in which the random number generator was invoked when reweighting events for the bootstrapping of MC differed between kinematic variables.

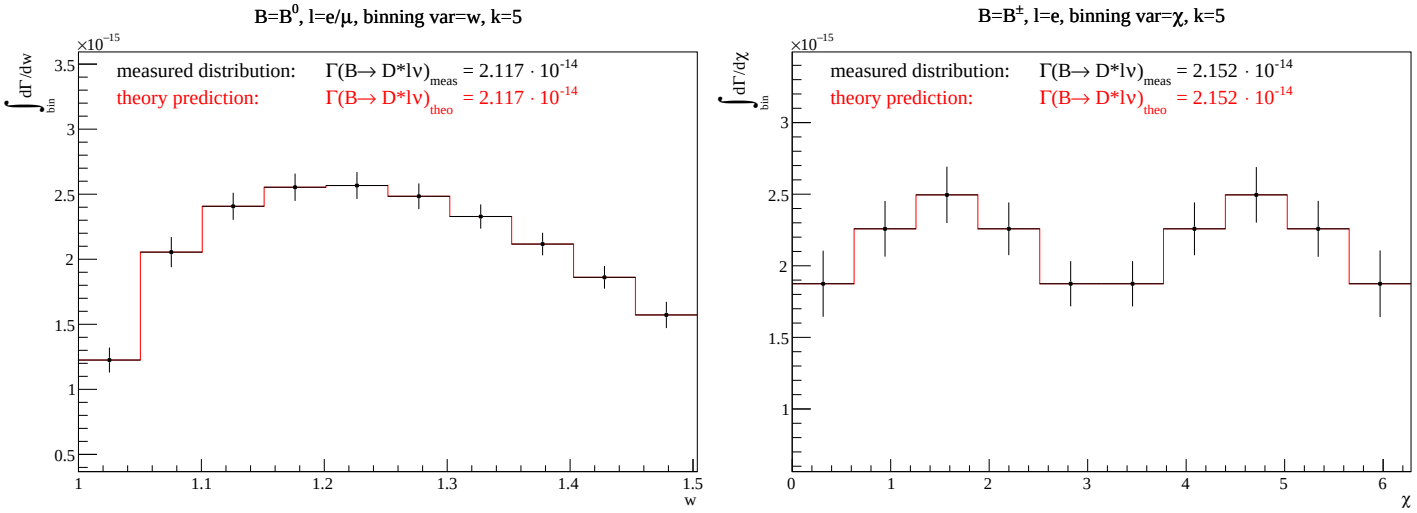


Figure 3.23: Closure test for the efficiency correction: Distribution of the differential decay width calculated from unfolded, efficiency corrected reco MC in black and the theoretically assumed distribution for the MC reweighting in red. Both distributions agree perfectly, thus only one common distribution is visible.

the correlation between the PDF shape errors $E^{v,i}_{\text{pdf}}$ was calculated as

$$\text{Corr}(E^{v,i}_{\text{pdf}}, E^{v',i'}_{\text{pdf}}) = \begin{cases} 0 & v = v', i \neq i' \\ 1 & v = v', i = i' \\ \rho(v,i;v',i') & \text{else,} \end{cases} \tag{3.10.1}$$

where the *overlap matrix* ρ was defined as

$$\rho(v,i;v',i') = \frac{\#(\text{events in bin } i \text{ of } v \cap \text{events in bin } i' \text{ of } v')}{\sqrt{\#(\text{events in bin } i \text{ of } v)}\sqrt{\#(\text{events in bin } i' \text{ of } v')}} \tag{3.10.2}$$

and was calculated with MC data. While a switch to the standard method using toys is intended (and a clean solution for the technical difficulties is now available), the results presented in this thesis are still calculated with the overlap matrix.

The covariance matrix $\text{Cov}(E^{v,i}_{\text{pdf}}, E^{v',i'}_{\text{pdf}})$ is calculated from $E^{v,i}_{\text{pdf}}$ and the correlation matrix.

3.10.2 Uncertainties due to MC reweighting

As described in section 3.3.5, corrections to the MC dataset are applied in form of event weights. There are several sources of corrections to be taken into account, such as branching fractions, form factors and efficiencies, each of which depends on one or more parameters (which are often further split up in bins of one or more variables, e.g. the neural network classifiers or detector regions for the efficiencies).

The basic idea is to vary these parameters according to their uncertainties and to observe their effect on the considered quantity, thereby determining the corresponding systematic uncertainty. Care has to be taken if the systematic errors on the parameters are correlated as they can no longer be independently varied. To that end, two different strategies are applied:

1. The straightforward strategy ("variation") is to vary each parameter p_i of by $\pm\sigma_{p_i}$ and observe the effect. If the errors on the parameters are correlated (across bins), the parameters are first transformed into eigenvectors $\tilde{p}_i$ of the correlation matrix, which can then be varied independently by $\pm\sigma_{\tilde{p}_i}$.

The error σ_s on a quantity $f(\tilde{p}_1,\ldots,\tilde{p}_n)$ corresponding to a systematics source S is then calculated as

$$\sigma_S = \sqrt{\sum_{i=1}^{n}\left(\frac{f(\tilde{p}_1,\ldots,\tilde{p}_i+\sigma_{\tilde{p}_i},\ldots,\tilde{p}_n)-f(\tilde{p}_1,\ldots,\tilde{p}_i-\sigma_{\tilde{p}_i},\ldots,\tilde{p}_n)}{2}\right)^2}. \tag{3.10.3}$$

2. If the transformation into the eigenspace proves impractical, because the errors on a parameter p are split across many bins of several variables (which makes the inversion of the correlation matrix numerically unstable), a different strategy ("toys") is employed: The errors on the weights w_i for events in bin i are split up[17] in an uncorrelated part $\Delta w_i^{\text{uncorr.}}$ and a fully correlated part $\Delta w_i^{\text{corr.}}$. For each toy, one random number κ_1 is picked from a Gaussian distribution Gauss$(0,1)$, which is shared across all events. In addition, another random number $\kappa_1(\text{toy},\text{event})$ is picked from a Gaussian distribution Gauss$(0,1)$ for each event. The weight is now calculated as

$$w_i(\text{toy},\text{event}) = w_i + \kappa_1(\text{toy})\,\Delta w_i^{\text{uncorr.}} + \kappa_2(\text{toy},\text{event})\,\Delta w_i^{\text{corr.}}. \tag{3.10.4}$$

The error σ_S on a quantity f corresponding to the systematics source S is then calculated as

$$\sigma_S = \frac{1}{N_{\text{toy}}}\sqrt{\sum_{i_{\text{toy}}=1}^{N_{\text{toy}}}\left(\overline{f}-f(i_{\text{toy}})\right)^2} \quad\text{with}\quad \overline{f} = \frac{1}{N_{\text{toy}}}\sum_{i_{\text{toy}}}^{N_{\text{toy}}} f(i_{\text{toy}}), \tag{3.10.5}$$

where $f(i_{\text{toy}})$ is the result of f for a specific toy.

A list of systematic sources and their methods is shown in table 3.4.

Uncertainties due to MC reweighting have to be considered at two points of the unfolding:

1. before the unfolding (to build the pre-unfolding covariance matrix, see section 3.10.3): Here their effect on the PDF shape has to be taken into consideration (but since the PDFs are normalized, the effect of the parameters on the overall efficiency is irrelevant).

2. after the unfolding (to build the post-unfolding covariance matrix, see section 3.10.4): Here their effect on the efficiencies is considered. This is described in more detail in section 3.10.5.

In both cases, the corresponding errors are considered to be fully correlated across all bins of all variables.[18]

3.10.3 Pre-unfolding covariance matrix

The 40×40 covariance matrix of the fit results from section 3.4 is calculated separately for $B^\pm$ and B^0 as well as for $\ell = e$, $\ell = \mu$ and $\ell = e+\mu$.

Three types of uncertainties have to be incorporated:

[17] To split up the errors on w_i into the uncorrelated and fully correlated part, consider the dependency of w_i on input measurements. Whenever the same measurement is used for the calculation of w_i and w_j, the corresponding error is fully correlated. On the other hand, a measurement that only is used for the calculation of one Δw_i, will contribute to the uncorrelated part of the error.

[18] This is a consequence of the systematics being parametrized by single parameters across all bins of all kinematic variables: Consider the variation of a parameter p by σ_p around the nominal value p_0 and dependent functions $f_i(p)$ (where i is for example an enumeration of the different bins and binning variables). For variations Δp in the range of sufficiently small uncertainties σ_p of the parameter p, the change in f_i, $f_i(p_0+\Delta p)-f_i(p_0)$ can be approximated as $\Delta p f_i'(p_0)$. Thus $\text{Cov}(f_i,f_j)=\sigma_p^2 f_i'(p_0)f_j'(p_0)$ and $\text{Corr}(f_i,f_j)=\text{Cov}(f_i,f_j)/\sqrt{\text{Cov}(f_i,f_i)\,\text{Cov}(f_j,f_j)}=1$.

Name	What	Method
`Ddecay`	BR($D \longrightarrow K\pi(\pi)(\pi)$)	variation
`Dsdecay`	BR($D^* \longrightarrow D\pi$)	variation
`BtoDsslnu`	BR($\bar{B} \longrightarrow D^{**}\ell^-\bar{\nu}_\ell$)	variation
`BtoDlnu`	BR($\bar{B} \longrightarrow D\ell^-\bar{\nu}_\ell$)	variation
`DsFF`	$\bar{B} \longrightarrow D^*\ell^-\bar{\nu}_\ell$ form factors	variation
`DssFF`	$\bar{B} \longrightarrow D^{**}\ell^-\bar{\nu}_\ell$ form factors	variation
`slowPi`	π_{slow} Efficiency	variation
`lepEff_el`	e PID	toys
`lepEff_mu`	μ PID	toys
`lepFake`	Lepton fake rates	toys
`KPiEff`	K PID	toys
`tagEff`	Tagging calibration	toys

Table 3.4: Systematics of the MC reweighting.

- Statistical errors on the fit results calculated together with the nominal fit results. The corresponding covariance matrix was calculated in section 3.6.
- The PDF shape uncertainty was calculated together with its covariance matrix in section 3.10.1.
- Other systematics as calculated by observing the effect of MC reweighting variations (as described in section 3.10.2) on the fitted signal yields, assuming full correlation, i.e.

$$\text{Cov}(E^{v,i}_{\text{reweigh}}, E^{v',i'}_{\text{reweigh}}) = \sum_S E^{v,i}_S \cdot E^{v',i'}_S, \tag{3.10.6}$$

  with $E^{v,i}_{\text{S}}$ calculated as in equation (3.10.3) or (3.10.5) and S running through the different systematic sources of table 3.4

The total pre-unfolding covariance matrix is then given as the sum of the three individual covariance matrices:

$$\begin{aligned}\text{Cov}(E^{v,i}_{\text{pre-unfold}}, E^{v',i'}_{\text{pre-unfold}}) = \\ \text{Cov}(E^{v,i}_{\text{stat}}, E^{v',i'}_{\text{stat}}) + \text{Cov}(E^{v,i}_{\text{pdf}}, E^{v',i'}_{\text{pdf}}) + \text{Cov}(E^{v,i}_{\text{reweigh}}, E^{v',i'}_{\text{reweigh}}).\end{aligned} \tag{3.10.7}$$

3.10.4 Post-unfolding covariance matrix

The errors and correlations between the yields per bin of the different variables need to be propagated through the unfolding process. This is done using toy experiments.

As before, $B^\pm$ and B^0 samples as well as $\ell = e$, $\ell = \mu$ and $\ell = e + \mu$ are treated separately. The nominal yields of all variables are considered as one vector (of dimension 40) together with the 40×40 covariance matrix. This yield vector is rotated into the eigenspace of the covariance matrix and a random variation is applied, i.e. a random number from a Gaussian distribution around zero with variance given by the corresponding eigenvalue of the covariance matrix is added. Rotating back into the previous system then gives us one toy distribution, the different dimensions corresponding to the yields $N^{\text{unf.},v,i}_{\text{s+w}}(j)$ of variable v in bin i, where j denotes the number of the toy distribution.

The covariance matrix (and thereby errors and correlations) is then calculated from $N = 1000$ toy samples $N_{\text{s+w}}^{\text{unf.},v,i}(j)$ with the maximum likelihood estimator

$$\begin{aligned}\text{Cov}(N_{\text{s+w}}^{\text{unf.},v,i}, N_{\text{s+w}}^{\text{unf.},v',i'}) &= \frac{1}{N}\sum_{j=1}^{N}\Big(N_{\text{s+w}}^{\text{unf.},v,i}(j) - \overline{N}_{\text{s+w}}^{\text{unf.},v,i}\Big)\Big(N_{\text{s+w}}^{\text{unf.},v',i'}(j) - \overline{N}_{\text{s+w}}^{\text{unf.},v',i'}\Big),\\ \text{with}\quad \overline{N}_{\text{s+w}}^{\text{unf.},v^{(\prime)},i^{(\prime)}} &= \frac{1}{N}\sum_{j=1}^{N} N_{\text{s+w}}^{\text{unf.},v^{(\prime)},i^{(\prime)}}(j),\end{aligned} \tag{3.10.8}$$

Because the migration is strongest for neighboring bins and the unfolding procedure reverts this process, the correlation matrix between different bins of the same variable as calculated from this covariance matrix will show non-negligible correlations particularly in the close-to-diagonal entries. This is in contrast to the pre-unfolding correlation matrix (section 3.10.3), where the off-diagonal correlations of statistical errors were assumed to vanish and (with the exception of the PDF shape uncertainty), all other systematics were assumed to be maximally correlated, thus resulting in approximately constant off-diagonal elements.

3.10.5 Uncertainties on the efficiency

As described in section 3.9, the efficiency correction is applied as $N_{\text{s+w}}^{\text{e.corr.},v,i} = [\epsilon_{\text{reco}}\epsilon_{\text{tag}}]^{-1}_{\text{MC}}\, N_{\text{s+w}}^{\text{unf.},v,i}$, where the MC based value $\epsilon_{\text{reco}}\epsilon_{\text{tag}}\big|_{\text{MC}}$ serves as an estimate for the real experimental reconstruction efficiency $\epsilon_{\text{reco}}\epsilon_{\text{tag}}\big|_{\text{real}}$, defined by $N_{\text{s+w}}^{\text{real},v,i} = [\epsilon_{\text{reco}}\epsilon_{\text{tag}}]^{-1}_{\text{real}}\, N_{\text{s+w}}^{\text{unf},v,i}$. Thus

$$N_{\text{s+w}}^{\text{real},v,i} = \frac{\epsilon_{\text{reco}}\epsilon_{\text{tag}}\big|_{\text{real}}}{\epsilon_{\text{reco}}\epsilon_{\text{tag}}\big|_{\text{MC}}}\cdot N_{\text{s+w}}^{\text{e.corr.},v,i}, \tag{3.10.9}$$

where we assume $\epsilon_{\text{reco}}\epsilon_{\text{tag}}\big|_{\text{real}} = \epsilon_{\text{reco}}\epsilon_{\text{tag}}\big|_{\text{MC}} \pm \Delta\,\epsilon_{\text{reco}}\epsilon_{\text{tag}}\big|_{\text{MC}}$ with $\Delta\,\epsilon_{\text{reco}}\epsilon_{\text{tag}}\big|_{\text{MC}}$ the error on the reconstruction efficiency, i.e.

$$N_{\text{s+w}}^{\text{real},v,i} = N_{\text{s+w}}^{\text{e.corr.},v,i}\left(1 \pm \frac{\Delta\,\epsilon_{\text{reco}}\epsilon_{\text{tag}}\big|_{\text{MC}}}{\epsilon_{\text{reco}}\epsilon_{\text{tag}}\big|_{\text{MC}}}\right). \tag{3.10.10}$$

Thus the relative error induced on our estimate $N_{\text{real}}^{v,i}$ is exactly the relative error on the efficiency $\epsilon_{\text{reco}}\epsilon_{\text{tag}}\big|_{\text{MC}}$. This error is calculated in the next three subsections.

3.10.5.1 MC reweighting

As $\epsilon_{\text{reco}}\epsilon_{\text{tag}}\big|_{\text{MC}}$ is proportional to the yield $N_{\text{s+w}}^{\text{truth},v,i}$ (number of total signal+wrong events projected in bin i of v based on MC truth information), the relative error on $\epsilon_{\text{reco}}\epsilon_{\text{tag}}\big|_{\text{MC}}$ corresponding to the uncertainty on the parameters used for the MC reweighting (as introduced in 3.10.2) is the same as that on $N_{\text{s+w}}^{\text{truth},v,i}$.

In principle, the estimation of the relative errors on $N_{\text{truth}}^{v,i}$ is the same as described in section 3.10.2, however a few differences have to be noted:

- The errors corresponding to the sources $\text{BR}(D\longrightarrow K\pi(\pi)(\pi))$, $\text{BR}(D^*\longrightarrow D\pi)$, K PID and tagging calibration are assumed to be flat in bins of the kinematic variable. Thus, in order to suppress the effect of statistical fluctuations, the relative error is calculated based on the total yields $\sum_i N_{\text{s+w}}^{\text{truth},v,i}$.
- The errors for sources of systematic errors that are treated using the "variation" strategy (see section 3.10.2) are calculated as

$$\begin{aligned}\sigma_{\pm S} &= \sqrt{\sum_i^n \big(f(\tilde{p}_1,\ldots,\tilde{p}_n) - f(\tilde{p}_1,\ldots,\tilde{p}_i \pm \sigma_{\tilde{p}_i},\ldots,\tilde{p}_n\big)^2},\\ \sigma_S &= \max\{\sigma_{+S},\sigma_{-S}\},\end{aligned} \tag{3.10.11}$$

where $\tilde{p}_i$ are the varied parameters. The values for $\sigma_{\pm S}$ are expected to be very close. This can be confirmed.[19]

The covariance matrix is calculated under the assumption of full correlation between all bins and variables.

In comparison with the results from [61, 129], several mistakes have been corrected:

- The errors for BR($D \longrightarrow K\pi(\pi)(\pi)$), BR($D^* \longrightarrow D\pi$), π_{slow} efficiency, e PID , μ PID, lepton fake rates and K PID had been mistakenly divided by 3 (however all of these errors are small, thus after taking the squared sum, this makes virtually no difference).

- The error for the π_{slow} efficiency had been mistakenly calculated using the root mean square, even though the errors are estimated using the "variation" strategy (section 3.10.2). Correction of this resulted in a tenfold increased error (although the effect on the final error budget is small as well).

The original scripts to calculate the errors were rewritten from scratch and integrated with the main analysis framework.

3.10.5.2 Charged track efficiency uncertainty

The uncertainty on the charged track reconstruction efficiency is 0.35% per charged track. This number needs to be scaled up by the average number of charged tracks per event. Since the track reconstruction efficiency depends on the true value of the kinematic variable (rather than the reconstructed one), this number is calculated by considering the truth MC event subsample belonging in the bin under consideration and (for B^0 mesons and $\ell = e + \mu$) ranges between 4.2 and 4.7. This number has not yet been recalculated by the author, nor have different numbers been calculated for $\ell = e, \mu, e/\mu$ or B^0 and $B^{\pm}$. Differences between the lepton flavors are expected to be small, but a more careful study is necessary for the $B^{\pm}$ case.

Full correlation accross bins and variables is assumed.

3.10.5.3 π^0 reconstruction efficiency uncertainty

π^0 reconstruction efficiency uncertainty: The π^0 reconstruction efficiency is 2%. This number is scaled up by the average number of π^0 in each event, similar to the calculation of the charged track efficiency uncertainty. This number is found to range from 0.3 to 0.5 (for B^0 mesons and $\ell = e + \mu$). It has not yet been recalculated by the author, nor have different numbers been calculated for $\ell = e, \mu, e/\mu$ or B^0 and $B^{\pm}$ (but differences between the lepton flavors are expected to be small).

Full correlation accross bins and variables is assumed.

3.10.5.4 Uncertainties on the number of B mesons

The number of charged or neutral B meson pairs recorded by Belle, $N_{B^0\bar{B}^0/B^{\pm}B^{\mp}}$ is calculated from the number of recorded B meson pairs (of any charge), $N_{B\bar{B}}$ and the world average for f_{+0} as described in section 3.10.5. This leads to a flat error with full correlation across all bins and all variables.

[19] Previous issues with vanishing errors for BR($D^* \longrightarrow D\pi$) with negative variations have been resolved.

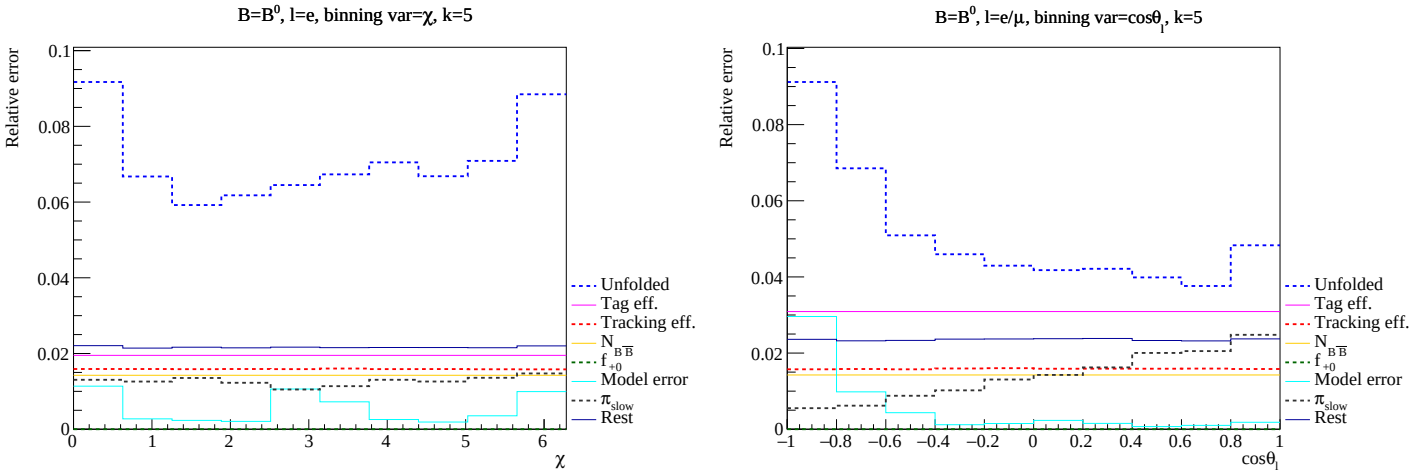


Figure 3.24: Comparison of relative errors per bin. Note that the errors will be added squared, thus the relative impact of large errors is (even) bigger than this figure suggests. The full set of plots is shown on pages 145 to 148.

3.10.6 Final covariance matrix

The final covariance matrix is the sum of the individual covariance matrices as follows:

- Post-unfolding covariance matrix as calculated in section 3.10.4: Statistical errors of the unbinned likelihood fit to m^2_{miss} and systematics affecting the PDF shape that were propagated through the unfolding algorithm.
- Model uncertainty covariance matrix: Built from the *model error* (error introduced by unfolding data using a erroneous migration matrix as discussed in section 3.8.2), assuming full correlation across bins and variables. Previously, the model uncertainty was calculated with both CLN and BGL parametrization of the form factors and both errors were added quadratically. The results presented here currently only estimate the errors with CLN parametrization. However, due to the small size of the error, this makes virtually no difference for the total errors.
- Efficiency correction uncertainty covariance matrix: The covariance matrix is built from the errors as discussed in section 3.10.5, assuming full correlation across bins and variables.
- Covariance matrix corresponding to the uncertainty on the number of charged or neutral B meson pairs recorded by Belle as described in section 3.10.5.4.

All of the individual covariance matrices are calculated based on relative errors and scaled up to the final results using the measured central values. In the course of this, the results in [129] and [61] erroneously included differing powers of a correction factor of 0.980. This has been fixed. The corresponding scripts have been rewritten from scratch and integrated with the main code base.

Two examples of the composition of the relative errors per bin are shown in figure 3.24. The influence of each error source, averaged over all bins and kinematic variables is also summarized in table 3.5. The total errors per bin are found to be of order 10%. For neutral (charged) B mesons, the combined lepton result has an average uncertainty of 7% (9%) per bin and around 8% (11%) for the lepton flavor specific results. This difference comes mainly from the larger statistical errors (part of the "unfolded errors"), while the effect of the tagging efficiency uncertainty is reduced (all other errors are approximately the same). The overall error on $B^\pm$ is higher due to the higher unfolded errors.

	B^0			$B^\pm$		
Variable	e	μ	$e+\mu$	e	μ	$e+\mu$
Unfolded errors	6.87%	7.25%	4.99%	9.15%	9.01%	6.43%
Tag eff.	1.95%	1.87%	3.09%	2.40%	2.40%	2.94%
Charged tracks efficiency	1.58%	1.58%	1.58%	1.58%	1.58%	1.58%
$N_{B\bar{B}}$	1.42%	1.42%	1.42%	1.42%	1.42%	1.42%
π_{slow}	1.39%	1.43%	1.40%	2.10%	2.08%	2.09%
BR($D \longrightarrow K\pi(\pi)(\pi)$	1.18%	1.19%	1.18%	1.41%	1.40%	1.41%
f_{+0}	1.17%	1.17%	1.17%	1.10%	1.10%	1.10%
e PID	1.38%	$< 0.01\%$	1.09%	1.35%	$< 0.01\%$	1.06%
K PID	0.78%	0.77%	1.02%	0.64%	0.68%	0.80%
μ PID	$< 0.01\%$	1.52%	1.01%	$< 0.01\%$	1.56%	1.05%
BR($D^* \longrightarrow D\pi$)	0.64%	0.64%	0.64%	4.31%	4.31%	4.31%
π^0 efficiency	0.55%	0.55%	0.55%	0.55%	0.55%	0.55%
$B \longrightarrow D^*\ell\bar{\nu}_\ell$ form factors	0.43%	0.43%	0.43%	0.47%	0.41%	0.42%
Model error	0.42%	0.37%	0.34%	0.52%	0.36%	0.38%
Fake leptons	$< 0.01\%$	$< 0.01\%$	$< 0.01\%$	$< 0.01\%$	$< 0.01\%$	$< 0.01\%$
BR($B \longrightarrow D^{**}\bar{\nu}_\ell$)	$< 0.01\%$	$< 0.01\%$	$< 0.01\%$	$< 0.01\%$	$< 0.01\%$	$< 0.01\%$
BR($B \longrightarrow D\ell\bar{\nu}_\ell$)	$< 0.01\%$	$< 0.01\%$	$< 0.01\%$	$< 0.01\%$	$< 0.01\%$	$< 0.01\%$
$B \longrightarrow D^{**}\ell\bar{\nu}_\ell$ form factors	$< 0.01\%$	$< 0.01\%$	$< 0.01\%$	$< 0.01\%$	$< 0.01\%$	$< 0.01\%$
Total error	**8.05%**	**8.40%**	**7.00%**	**11.20%**	**11.10%**	**9.26%**

Table 3.5: Average error per error source for the efficiency corrected unfolded decay rates. Sorted by the errors for B^0 and $\ell = e + \mu$.

Note that the BR($D^{*0} \longrightarrow D\pi$) error relevant for $B^\pm$ is not up to date: it still corresponds to a reweighting of the Belle I MC using the 2014 PDG averages [142], where BR($D^{*0} \longrightarrow D^0\pi^0$) was reported as $(61.9 \pm 2.9)\%$ (i.e. a 4.7% error, consistent with the 4.31% error shown above). This measurement has since been updated to $(64.7 \pm 0.9)\%$ (2018 average [5]), which is expected to result in a similar error compared to B^0 mesons.

3.11 Results

The final results together with their errors are summarized in table 3.6 and visualized together with the MC distributions in figure 3.25. A comparison with the previous results from [61, 129] is shown in figure 3.26. While several changes were introduced in the calculation of the errors, the total error remains largely unchanged.

The correlation matrices can be found on page 149 to 154. The average non-diagonal correlation matrix element is around 40% (44%) for lepton flavor specific fits of neutral (charged) B mesons and 57% (58%) for the lepton flavor agnostic fit.

$$\frac{\Delta\Gamma_i}{\Delta v_i}\ [10^{-15}\ \mathrm{GeV}]$$

		B^0			$B^\pm$		
v	Bin	e	μ	$e+\mu$	e	μ	$e+\mu$
w	0	$x.xx \pm 0.16$	$x.xx \pm 0.15$	$x.xx \pm 0.12$	$x.xx \pm 0.15$	$x.xx \pm 0.14$	$x.xx \pm 0.12$
	1	$x.xx \pm 0.19$	$x.xx \pm 0.19$	$x.xx \pm 0.16$	$x.xx \pm 0.20$	$x.xx \pm 0.19$	$x.xx \pm 0.17$
	2	$x.xx \pm 0.18$	$x.xx \pm 0.18$	$x.xx \pm 0.16$	$x.xx \pm 0.21$	$x.xx \pm 0.22$	$x.xx \pm 0.20$
	3	$x.xx \pm 0.18$	$x.xx \pm 0.18$	$x.xx \pm 0.16$	$x.xx \pm 0.22$	$x.xx \pm 0.23$	$x.xx \pm 0.20$
	4	$x.xx \pm 0.18$	$x.xx \pm 0.18$	$x.xx \pm 0.16$	$x.xx \pm 0.22$	$x.xx \pm 0.21$	$x.xx \pm 0.19$
	5	$x.xx \pm 0.17$	$x.xx \pm 0.16$	$x.xx \pm 0.15$	$x.xx \pm 0.21$	$x.xx \pm 0.20$	$x.xx \pm 0.18$
	6	$x.xx \pm 0.16$	$x.xx \pm 0.15$	$x.xx \pm 0.14$	$x.xx \pm 0.20$	$x.xx \pm 0.19$	$x.xx \pm 0.17$
	7	$x.xx \pm 0.15$	$x.xx \pm 0.14$	$x.xx \pm 0.13$	$x.xx \pm 0.18$	$x.xx \pm 0.18$	$x.xx \pm 0.16$
	8	$x.xx \pm 0.14$	$x.xx \pm 0.15$	$x.xx \pm 0.13$	$x.xx \pm 0.17$	$x.xx \pm 0.19$	$x.xx \pm 0.15$
	9	$x.xx \pm 0.15$	$x.xx \pm 0.17$	$x.xx \pm 0.13$	$x.xx \pm 0.18$	$x.xx \pm 0.20$	$x.xx \pm 0.15$
χ	0	$x.xx \pm 0.20$	$x.xx \pm 0.19$	$x.xx \pm 0.15$	$x.xx \pm 0.25$	$x.xx \pm 0.26$	$x.xx \pm 0.20$
	1	$x.xx \pm 0.18$	$x.xx \pm 0.17$	$x.xx \pm 0.15$	$x.xx \pm 0.23$	$x.xx \pm 0.22$	$x.xx \pm 0.19$
	2	$x.xx \pm 0.19$	$x.xx \pm 0.18$	$x.xx \pm 0.16$	$x.xx \pm 0.25$	$x.xx \pm 0.23$	$x.xx \pm 0.21$
	3	$x.xx \pm 0.18$	$x.xx \pm 0.17$	$x.xx \pm 0.15$	$x.xx \pm 0.23$	$x.xx \pm 0.24$	$x.xx \pm 0.20$
	4	$x.xx \pm 0.15$	$x.xx \pm 0.14$	$x.xx \pm 0.13$	$x.xx \pm 0.19$	$x.xx \pm 0.21$	$x.xx \pm 0.17$
	5	$x.xx \pm 0.15$	$x.xx \pm 0.14$	$x.xx \pm 0.13$	$x.xx \pm 0.19$	$x.xx \pm 0.20$	$x.xx \pm 0.16$
	6	$x.xx \pm 0.16$	$x.xx \pm 0.17$	$x.xx \pm 0.14$	$x.xx \pm 0.22$	$x.xx \pm 0.23$	$x.xx \pm 0.19$
	7	$x.xx \pm 0.17$	$x.xx \pm 0.18$	$x.xx \pm 0.15$	$x.xx \pm 0.24$	$x.xx \pm 0.23$	$x.xx \pm 0.20$
	8	$x.xx \pm 0.17$	$x.xx \pm 0.17$	$x.xx \pm 0.15$	$x.xx \pm 0.24$	$x.xx \pm 0.22$	$x.xx \pm 0.19$
	9	$x.xx \pm 0.20$	$x.xx \pm 0.18$	$x.xx \pm 0.16$	$x.xx \pm 0.26$	$x.xx \pm 0.26$	$x.xx \pm 0.20$
$\cos\theta_{D^*}$	0	$x.xx \pm 0.24$	$x.xx \pm 0.25$	$x.xx \pm 0.21$	$x.xx \pm 0.32$	$x.xx \pm 0.36$	$x.xx \pm 0.28$
	1	$x.xx \pm 0.16$	$x.xx \pm 0.17$	$x.xx \pm 0.15$	$x.xx \pm 0.21$	$x.xx \pm 0.24$	$x.xx \pm 0.19$
	2	$x.xx \pm 0.14$	$x.xx \pm 0.15$	$x.xx \pm 0.13$	$x.xx \pm 0.20$	$x.xx \pm 0.21$	$x.xx \pm 0.17$
	3	$x.xx \pm 0.13$	$x.xx \pm 0.14$	$x.xx \pm 0.11$	$x.xx \pm 0.18$	$x.xx \pm 0.19$	$x.xx \pm 0.15$
	4	$x.xx \pm 0.13$	$x.xx \pm 0.13$	$x.xx \pm 0.11$	$x.xx \pm 0.17$	$x.xx \pm 0.17$	$x.xx \pm 0.14$
	5	$x.xx \pm 0.14$	$x.xx \pm 0.13$	$x.xx \pm 0.11$	$x.xx \pm 0.16$	$x.xx \pm 0.17$	$x.xx \pm 0.13$
	6	$x.xx \pm 0.15$	$x.xx \pm 0.14$	$x.xx \pm 0.12$	$x.xx \pm 0.18$	$x.xx \pm 0.18$	$x.xx \pm 0.14$
	7	$x.xx \pm 0.16$	$x.xx \pm 0.15$	$x.xx \pm 0.13$	$x.xx \pm 0.21$	$x.xx \pm 0.20$	$x.xx \pm 0.17$
	8	$x.xx \pm 0.20$	$x.xx \pm 0.18$	$x.xx \pm 0.17$	$x.xx \pm 0.25$	$x.xx \pm 0.24$	$x.xx \pm 0.21$
	9	$x.xx \pm 0.31$	$x.xx \pm 0.28$	$x.xx \pm 0.24$	$x.xx \pm 0.37$	$x.xx \pm 0.33$	$x.xx \pm 0.29$
$\cos\theta_\ell$	0	$x.xx \pm 0.09$	$x.xx \pm 0.11$	$x.xx \pm 0.08$	$x.xx \pm 0.09$	$x.xx \pm 0.12$	$x.xx \pm 0.08$
	1	$x.xx \pm 0.11$	$x.xx \pm 0.13$	$x.xx \pm 0.10$	$x.xx \pm 0.12$	$x.xx \pm 0.15$	$x.xx \pm 0.11$
	2	$x.xx \pm 0.13$	$x.xx \pm 0.14$	$x.xx \pm 0.11$	$x.xx \pm 0.15$	$x.xx \pm 0.17$	$x.xx \pm 0.14$
	3	$x.xx \pm 0.15$	$x.xx \pm 0.16$	$x.xx \pm 0.13$	$x.xx \pm 0.19$	$x.xx \pm 0.19$	$x.xx \pm 0.16$
	4	$x.xx \pm 0.16$	$x.xx \pm 0.17$	$x.xx \pm 0.15$	$x.xx \pm 0.21$	$x.xx \pm 0.21$	$x.xx \pm 0.18$
	5	$x.xx \pm 0.18$	$x.xx \pm 0.18$	$x.xx \pm 0.16$	$x.xx \pm 0.22$	$x.xx \pm 0.22$	$x.xx \pm 0.20$
	6	$x.xx \pm 0.18$	$x.xx \pm 0.18$	$x.xx \pm 0.16$	$x.xx \pm 0.23$	$x.xx \pm 0.24$	$x.xx \pm 0.21$
	7	$x.xx \pm 0.19$	$x.xx \pm 0.18$	$x.xx \pm 0.17$	$x.xx \pm 0.23$	$x.xx \pm 0.24$	$x.xx \pm 0.21$
	8	$x.xx \pm 0.20$	$x.xx \pm 0.18$	$x.xx \pm 0.18$	$x.xx \pm 0.24$	$x.xx \pm 0.23$	$x.xx \pm 0.21$
	9	$x.xx \pm 0.24$	$x.xx \pm 0.22$	$x.xx \pm 0.20$	$x.xx \pm 0.27$	$x.xx \pm 0.26$	$x.xx \pm 0.23$

Table 3.6:]
Unfolded kinematic distributions [CENSORED].

Censored

Censored

Censored

Censored

Figure 3.25: Final results for the differential cross sections. The dashed line represents the MC data.

Censored

Censored

Figure 3.26: Comparison of the results as calculated by the author ("new", black data points) with the results from [61, 129] ("old", red dashed line and shaded rectangles). The MC distribution is shown as green dashed line.

3.12 Summary

An update on the hadronically tagged analysis of $\bar{B}\longrightarrow D^*\ell^-\bar{\nu}_\ell$ using data from Belle I was given. Unfolded differential decay rates in four kinematic variables were presented separately for $\ell = e$ and $\ell = \mu$ in the final state. The consistency, usability, and documentation of the code base were improved, some bugs were corrected and several parts have been rewritten from scratch. While some small details still need investigation, the analysis will be complete with the final V_{cb} and form factor fit.

Appendix A

Full Expression For the Differential Cross Section

The full angular dependency of the differential cross section of $\bar{B} \longrightarrow D^*(\rightarrow D\pi) l^- \bar{\nu}_l$ is given in [116, p. 5] as follows:

$$\begin{aligned}\frac{\mathrm{d}^4\Gamma}{\mathrm{d}q^2\,\mathrm{dcos}\,\theta_l\,\mathrm{dcos}\,\theta_{D^*}\,\mathrm{d}\chi} = \frac{9}{32\pi} N_F \Bigg(& \cos^2\theta_{D^*}\left(V_1^0 + V_2^0\cos 2\theta_\ell + V_3^0\cos\theta_\ell\right) \\ & + \sin^2\theta_{D^*}\left(V_1^T + V_2^T\cos 2\theta_\ell + V_3^T\cos\theta_\ell\right) \\ & + V_4^T\sin^2\theta_{D^*}\sin^2\theta_\ell\cos 2\chi + V_1^{0T}\sin 2\theta_{D^*}\sin 2\theta_\ell\cos\chi \\ & + V_2^{0T}\sin 2\theta_{D^*}\sin\theta_\ell\cos\chi + V_5^T\sin^2\theta_{D^*}\sin^2\theta_\ell\sin 2\chi \\ & + V_3^{0T}\sin 2\theta_{D^*}\sin\theta_\ell\sin\chi + V_4^{0T}\sin 2\theta_{D^*}\sin 2\theta_\ell\sin\chi\Bigg).\end{aligned} \tag{A.0.1}$$

Thus, the following six functions in θ_ℓ appear:

$$1, \cos\theta_\ell, \cos 2\theta_\ell, \sin\theta_\ell, \sin 2\theta_\ell, \sin^2\theta_\ell \tag{A.0.2}$$

Note however, that these functions are not linearly independent, because $\cos 2\theta_\ell = \cos^2\theta_\ell - \sin^2\theta_\ell = 1 - 2\sin\theta_\ell$. Thus, the above dissection is inconvenient for our needs[1]. To remedy this, we expand $\cos 2\theta_\ell$ in the above expression and get

$$\begin{aligned}&\frac{\mathrm{d}^4\Gamma}{\mathrm{d}q^2\,\mathrm{dcos}\,\theta_l\,\mathrm{dcos}\,\theta_{D^*}\,\mathrm{d}\chi} \propto \\ &\quad\propto \cos^2\theta_{D^*}(V_1^0 + V_2^0\cos 2\theta_\ell + \cdots) + \sin^2\theta_{D^*}(V_1^T + V_2^T\cos 2\theta_\ell + \cdots) + \cdots = \\ &\quad= \cos^2\theta_{D^*}((V_1^0 + V_2^0) - 2V_2^0\sin^2\theta_\ell + \cdots) + \sin^2\theta_{D^*}((V_1^T + V_2^T) - 2V_2^T\sin^2\theta_\ell) + \cdots\end{aligned} \tag{A.0.3}$$

Thus, it makes sense to consider the combination $V_1^0 + V_2^0$ $(V_1^T + V_2^T)$ and $-2V_2^0$ $(-2V_2^T)$ as new variables. To distinguish the two different conventions, we write:

$$\begin{aligned} W_1^0 &:= V_1^0 + V_2^0, & W_1^T &:= V_1^T + V_2^T, \\ W_2^0 &:= -2V_2^0, & W_2^T &:= -2V_2^T, \\ W_a &:= V_a \quad \forall\, a \neq \{10, 20, 1T, 2T\}. \end{aligned} \tag{A.0.4}$$

Furthermore, note that chapter 2 considered the cross section

$$\frac{\mathrm{d}^4\Gamma}{\mathrm{d}q^2\,\mathrm{d}\theta_l\,\mathrm{d}\theta_{D^*}\,\mathrm{d}\chi}, \tag{A.0.5}$$

[1] Each of the V_a terms is still linearly independent if all three angles are considered together, so the dissection does make sense as far as the analysis in [116] is concerned. However, after integrations over χ and θ_{D^*} linear dependences can appear.

which accounts for a uniform factor of $\sin(\theta_\ell)\sin(\theta_{D^*})$ in the angle dependencies compared to equation (A.0.2). With this, we arrive at the result of $\sum_a W_a B_a$ of chapter 2.

A.1 Dependency on NP Coupling Constants

In [116], the expressions for the V_a are split up in terms of hadronic helicity amplitudes $\mathcal{A}_0$, $\mathcal{A}_t$, $\mathcal{A}_P$, $\mathcal{A}_{0T}$, $\mathcal{A}_{\pm T}$, and $\mathcal{A}_\pm$, which depend on the coupling constants g_A, g_V, g_S, g_P and T_L. Trying to make the dependency on the coupling constants of the W_a more apparent, we first split up the hadronic helicity amplitudes in terms of the coupling constants:

$$\begin{gathered}\mathcal{A}_0 = (1-g_A)\mathcal{A}_0^{(A)}, \qquad \mathcal{A}_\pm = (1-g_A)\mathcal{A}_{\pm'}^{(A)} \mp (g_V+1)\mathcal{A}_{\pm'}^{(V)}, \\ \mathcal{A}_t = (1-g_A)\mathcal{A}_t^{(A)}, \\ \mathcal{A}_P = g_P\,\mathcal{A}_P^{(P)}, \\ \mathcal{A}_{0T} = T_L\,\mathcal{A}_{0T}^{(T)}, \qquad \mathcal{A}_{\pm T} = T_L\,\mathcal{A}_{\pm T}^{(T)}.\end{gathered} \tag{A.1.1}$$

Substituting this into the expressions given for the V_a or W_a, we split them up as $W_a = \sum_{i=1}^{13} W_a^{(i)} c_i$, where the 13 c_i are real/imaginary parts/absolute values of bilinears in the coupling constants (as showed in table 2.11). The expressions for the $W_a^{(i)}$ are given in table A.1.

A.2 Dependency on q^2

The factor N_F is given by:

$$N_F = \frac{G_F^2 |p_{D^*}||V_{cb}|^2 q^2}{3\times 2^6\pi^3 m_B^2}\Big(1-\frac{m_l^2}{q^2}\Big)^2 \,\mathrm{Br}(D^*\to D\pi), \tag{A.2.1}$$

where we relate $|p_{D^*}|$ to q^2 via [59, p. 7]:

$$|p_{D^*}| = \frac{1}{2m_b}\sqrt{m_B^4 + m_{D^*}^4 + q^4 - 2(m_B^2 m_{D^*}^2 + m_{D^*}^2 q^2 + m_B^2 q^2)} \tag{A.2.2}$$

The hadronic helicity amplitudes are given in terms of $B\longrightarrow D^*$ form factors as follows:

$$\begin{aligned}\mathcal{A}_0^{(A)} &= \frac{(m_{D^*}+m_B)}{2\sqrt{q^2}m_{D^*}}\Big(\big(-m_{D^*}^2+m_B^2-q^2\big)A_1 - \frac{\lambda_{D^*}}{(m_{D^*}+m_B)^2}A_2\Big), \\ \mathcal{A}_\pm^{(A)} &= (m_{D^*}+m_B)A_1, \\ \mathcal{A}_\pm^{(V)} &= \frac{\sqrt{\lambda_{D^*}}}{m_{D^*}+m_B}V, \\ \mathcal{A}_t^{(A)} &= \frac{\sqrt{\lambda_{D^*}}}{\sqrt{q^2}}A_0, \\ \mathcal{A}_P^{(P)} &= \frac{\sqrt{\lambda_{D^*}}}{m_b+m_c}A_0, \\ \mathcal{A}_{0T}^{(T)} &= \frac{1}{2m_{D^*}}\Big(\big(3m_{D^*}^2+m_B^2-q^2\big)T_2 - \frac{\lambda_{D^*}}{m_B^2-m_{D^*}^2}T_3\Big), \\ \mathcal{A}_{\pm T}^{(T)} &= \pm\sqrt{\frac{\lambda_{D^*}}{q^2}}T_1 + \frac{m_B^2-m_{D^*}^2}{\sqrt{q^2}}T_2,\end{aligned} \tag{A.2.3}$$

where [118, p. 4]

$$\lambda_{D^*} = \big((m_B-m_{D^*})^2-q^2\big)\big((m_B+m_{D^*})^2-q^2\big) \overset{(A.2.2)}{=} \big(2m_B|p_{D^*}|\big)^2. \tag{A.2.4}$$

These form factor expressions are related to the form factor convention commonly used in Heavy Quark Effective Field Theory (HQET) [118, p. 19]:

$$V = \frac{m_{D^*} + m_B}{2\sqrt{m_B m_{D^*}}} h_V, \tag{A.2.5a}$$

$$A_0 = \frac{1}{2\sqrt{m_B m_{D^*}}} \left(\frac{(m_{D^*} + m_B)^2 - q^2}{2m_{D^*}} h_{A_1} - \frac{-m_{D^*}^2 + m_B^2 + q^2}{2m_B} h_{A_2} + \right.$$
$$\left. - \frac{-m_{D^*}^2 + m_B^2 - q^2}{2m_{D^*}} h_{A_3} \right), \tag{A.2.5b}$$

$$A_1 = \frac{(m_{D^*} + m_B)^2 - q^2}{2\sqrt{m_B m_{D^*}}(m_{D^*} + m_B)} h_{A_1},$$

$$A_2 = \frac{m_{D^*} + m_B}{2\sqrt{m_B m_{D^*}}} \left(\frac{m_{D^*}}{m_B} h_{A_2} + h_{A_3} \right),$$

$$T_1 = \frac{h_{T_1}(m_{D^*} + m_B) - h_{T_2}(m_B - m_{D^*})}{2\sqrt{m_B m_{D^*}}},$$

$$T_2 = \frac{1}{2\sqrt{m_B m_{D^*}}} \left(\frac{(m_{D^*} + m_B)^2 - q^2}{m_{D^*} + m_B} h_{T_1} - \frac{(m_B - m_{D^*})^2 - q^2}{m_B - m_{D^*}} h_{T_2} \right), \tag{A.2.5c}$$

$$T_3 = \frac{1}{2\sqrt{m_B m_{D^*}}} \left((m_B - m_{D^*}) h_{T_1} - (m_{D^*} + m_B) h_{T_2} - \frac{2(m_B^2 - m_{D^*}^2)}{m_B} h_{T_3} \right).$$

Using HQET relations, all h terms are then expressed in terms of a single function, h_{A_1} [118, pp. 20–21]:

$$h_{T_1} = \frac{1}{2(-2wr_{D^*} + r_{D^*}^2 + 1)} \left(\frac{(m_b - m_c)^2}{m_B - m_{D^*}} (1 - r_{D^*})(w+1) h_{A_1} + \right.$$
$$\left. - \frac{m_b + m_c}{m_{D^*} + m_B} (r_{D^*} + 1)^2 (w - 1) h_V \right),$$

$$h_{T_2} = \frac{1}{2(-2wr_{D^*} + r_{D^*}^2 + 1)} (w+1) \left(1 - r_{D^*}^2\right) \left(\frac{h_{A_1}(m_b - m_c)}{m_B - m_{D^*}} - \frac{h_V(m_b + m_c)}{m_{D^*} + m_B} \right), \tag{A.2.6a}$$

$$h_{T_3} = -\frac{1}{2(r_{D^*} + 1)(-2wr_{D^*} + r_{D^*}^2 + 1)} \left(\frac{m_b - m_c}{m_B - m_{D^*}} 2(w+1) r_{D^*} h_{A_1} + \right.$$
$$- \frac{m_b - m_c}{m_B - m_{D^*}} \left(-2wr_{D^*} + r_{D^*}^2 + 1\right) (h_{A_3} - r_{D^*} h_{A_2}) +$$
$$\left. - \frac{m_b + m_c}{m_{D^*} + m_B} (r_{D^*} + 1)^2 h_V \right),$$

$$h_V = R_1 h_{A_1}, \tag{A.2.6b}$$

$$h_{A_2} = \frac{R_2 - R_3}{2r_{D^*}} h_{A_1},$$

$$h_{A_3} = \frac{1}{2}(R_2 + R_3) h_{A_1}, \tag{A.2.6c}$$

where $r_{D^*} = m_{D^*}/m_B$ ([118, p. 20], do not confuse with $R_{D^*} = 2\sqrt{m_B m_{D^*}}/(m_B + m_{D^*})$ e.g. used by [59, p. 10]) and the q^2 dependency is expressed in terms of the product of the four-momenta of the B and D^*, i.e.

$$w = v_B v_{D^*} = \frac{m_B^2 + m_{D^*}^2 - q^2}{2m_B m_{D^*}}. \tag{A.2.7}$$

$$W_1^{0(1)} = \frac{4m_l^2\left(\left(\mathcal{A}_0^{(A)}\right)^2+\left(\mathcal{A}_t^{(A)}\right)^2\right)}{q^2}$$
$$W_1^{0(3)} = 4\left(\mathcal{A}_P^{(P)}\right)^2$$
$$W_1^{0(4)} = 64\left(\mathcal{A}_{0T}^{(T)}\right)^2$$
$$W_1^{0(6)} = \frac{8m_l\mathcal{A}_P^{(P)}\mathcal{A}_t^{(A)}}{\sqrt{q^2}}$$
$$W_1^{0(7)} = -\frac{32m_l\mathcal{A}_0^{(A)}\mathcal{A}_{0T}^{(T)}}{\sqrt{q^2}}$$
$$W_2^{0(1)} = 4\left(\mathcal{A}_0^{(A)}\right)^2\left(1-\frac{m_l^2}{q^2}\right)$$
$$W_2^{0(4)} = -64\left(\mathcal{A}_{0T}^{(T)}\right)^2\left(1-\frac{m_l^2}{q^2}\right)$$
$$W_3^{0(1)} = -\frac{8m_l^2\mathcal{A}_0^{(A)}\mathcal{A}_t^{(A)}}{q^2}$$
$$W_3^{0(6)} = -\frac{8m_l\mathcal{A}_0^{(A)}\mathcal{A}_P^{(P)}}{\sqrt{q^2}}$$
$$W_3^{0(7)} = \frac{32m_l\mathcal{A}_{0T}^{(T)}\mathcal{A}_t^{(A)}}{\sqrt{q^2}}$$
$$W_3^{0(9)} = 32\mathcal{A}_P^{(P)}\mathcal{A}_{0T}^{(T)}$$

$$W_1^{T(1)} = 4\left(\mathcal{A}_\pm^{(A)}\right)^2$$
$$W_1^{T(2)} = 4\left(\mathcal{A}_\pm^{(V)}\right)^2$$
$$W_1^{T(4)} = \frac{32m_l^2\left(\left(\mathcal{A}_{-T}^{(T)}\right)^2+\left(\mathcal{A}_{+T}^{(T)}\right)^2\right)}{q^2}$$
$$W_1^{T(7)} = -\frac{16m_l\left(\mathcal{A}_{-T}^{(T)}+\mathcal{A}_{+T}^{(T)}\right)\mathcal{A}_\pm^{(A)}}{\sqrt{q^2}}$$
$$W_1^{T(8)} = \frac{16m_l\left(\mathcal{A}_{+T}^{(T)}-\mathcal{A}_{-T}^{(T)}\right)\mathcal{A}_\pm^{(V)}}{\sqrt{q^2}}$$
$$W_2^{T(1)} = -2\left(\mathcal{A}_\pm^{(A)}\right)^2\left(1-\frac{m_l^2}{q^2}\right)$$
$$W_2^{T(2)} = -2\left(1-\frac{m_l^2}{q^2}\right)\left(\mathcal{A}_\pm^{(V)}\right)^2$$
$$W_2^{T(4)} = 8\left(\left(\mathcal{A}_{+T}^{(T)}-\mathcal{A}_{-T}^{(T)}\right)^2+\left(\mathcal{A}_{-T}^{(T)}+\mathcal{A}_{+T}^{(T)}\right)^2\right)\left(1-\frac{m_l^2}{q^2}\right)$$
$$W_3^{T(4)} = -\frac{32m_l^2\left(\mathcal{A}_{+T}^{(T)}-\mathcal{A}_{-T}^{(T)}\right)\left(\mathcal{A}_{-T}^{(T)}+\mathcal{A}_{+T}^{(T)}\right)}{q^2}$$
$$W_3^{T(5)} = 8\mathcal{A}_\pm^{(A)}\mathcal{A}_\pm^{(V)}$$
$$W_3^{T(7)} = \frac{16m_l\left(\mathcal{A}_{+T}^{(T)}-\mathcal{A}_{-T}^{(T)}\right)\mathcal{A}_\pm^{(A)}}{\sqrt{q^2}}$$
$$W_3^{T(8)} = -\frac{16m_l\left(\mathcal{A}_{-T}^{(T)}+\mathcal{A}_{+T}^{(T)}\right)\mathcal{A}_\pm^{(V)}}{\sqrt{q^2}}$$
$$W_4^{T(1)} = -2\left(\mathcal{A}_\pm^{(A)}\right)^2\left(1-\frac{m_l^2}{q^2}\right)$$
$$W_4^{T(2)} = 2\left(1-\frac{m_l^2}{q^2}\right)\left(\mathcal{A}_\pm^{(V)}\right)^2$$
$$W_4^{T(4)} = 8\left(\left(\mathcal{A}_{-T}^{(T)}+\mathcal{A}_{+T}^{(T)}\right)^2-\left(\mathcal{A}_{+T}^{(T)}-\mathcal{A}_{-T}^{(T)}\right)^2\right)\left(1-\frac{m_l^2}{q^2}\right)$$
$$W_5^{T(10)} = -4\mathcal{A}_\pm^{(A)}\left(1-\frac{m_l^2}{q^2}\right)\mathcal{A}_\pm^{(V)}$$

$$W_1^{0T(1)} = \frac{2\mathcal{A}_0^{(A)}\mathcal{A}_\pm^{(A)}\left(q^2-m_l^2\right)}{q^2}$$
$$W_1^{0T(4)} = -\frac{16\left(\mathcal{A}_{-T}^{(T)}+\mathcal{A}_{+T}^{(T)}\right)\mathcal{A}_{0T}^{(T)}\left(q^2-m_l^2\right)}{q^2}$$
$$W_2^{0T(1)} = \frac{4m_l^2\mathcal{A}_\pm^{(A)}\mathcal{A}_t^{(A)}}{q^2}$$
$$W_2^{0T(4)} = -\frac{32m_l^2\left(\mathcal{A}_{+T}^{(T)}-\mathcal{A}_{-T}^{(T)}\right)\mathcal{A}_{0T}^{(T)}}{q^2}$$
$$W_2^{0T(5)} = 4\mathcal{A}_0^{(A)}\mathcal{A}_\pm^{(V)}$$
$$W_2^{0T(6)} = \frac{4m_l\mathcal{A}_P^{(P)}\mathcal{A}_\pm^{(A)}}{\sqrt{q^2}}$$
$$W_2^{0T(7)} = \frac{8m_l}{\sqrt{q^2}}\Bigg(\mathcal{A}_0^{(A)}\left(\mathcal{A}_{+T}^{(T)}-\mathcal{A}_{-T}^{(T)}\right) - \left(\mathcal{A}_{-T}^{(T)}+\mathcal{A}_{+T}^{(T)}\right)\mathcal{A}_t^{(A)}\Bigg)$$
$$W_2^{0T(8)} = -\frac{16m_l\mathcal{A}_{0T}^{(T)}\mathcal{A}_\pm^{(V)}}{\sqrt{q^2}}$$
$$W_2^{0T(9)} = -8\mathcal{A}_P^{(P)}\left(\mathcal{A}_{-T}^{(T)}+\mathcal{A}_{+T}^{(T)}\right)$$
$$W_3^{0T(10)} = \frac{4m_l^2\mathcal{A}_t^{(A)}\mathcal{A}_\pm^{(V)}}{q^2}$$
$$W_3^{0T(11)} = -\frac{8m_l}{\sqrt{q^2}}\Bigg(\left(\mathcal{A}_{+T}^{(T)}-\mathcal{A}_{-T}^{(T)}\right)\mathcal{A}_t^{(A)} + - \mathcal{A}_0^{(A)}\left(\mathcal{A}_{-T}^{(T)}+\mathcal{A}_{+T}^{(T)}\right)+2\mathcal{A}_{0T}^{(T)}\mathcal{A}_\pm^{(A)}\Bigg)$$
$$W_3^{0T(12)} = -\frac{4m_l\mathcal{A}_P^{(P)}\mathcal{A}_\pm^{(V)}}{\sqrt{q^2}}$$
$$W_3^{0T(13)} = -8\mathcal{A}_P^{(P)}\left(\mathcal{A}_{+T}^{(T)}-\mathcal{A}_{-T}^{(T)}\right)$$
$$W_4^{0T(10)} = 2\mathcal{A}_0^{(A)}\left(1-\frac{m_l^2}{q^2}\right)\mathcal{A}_\pm^{(V)}$$

Table A.1: Expressions for the coefficients $W_a^{(i)}$ ($a \in A$, $1 \le i \le 13$) in terms of hadronic helicity amplitudes.

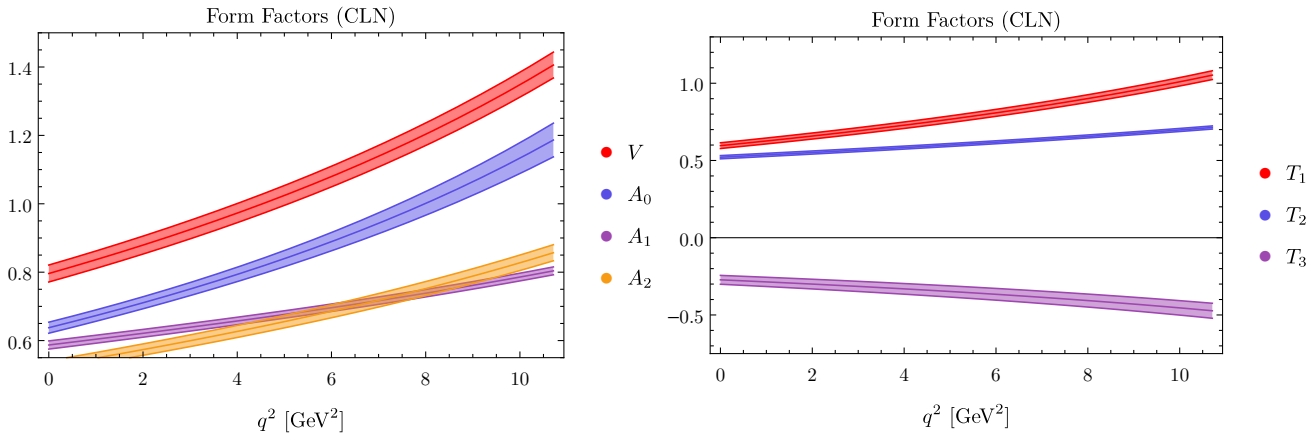


Figure A.1: Plots of the form factors using the CLN parametrization. The bands correspond to the pointwise error from the uncertainty on all parameters as calculated in section A.4.

Finally, using the parametrization introduced by Caprini, Lellouch and Neubert (CLN parametrization, [60]):

$$\begin{aligned} h_{A_1} &= h_{A_1}(1)\Big(-z^3\Big(231\rho_{D^*}^2 - 91\Big) + z^2\Big(53\rho_{D^*}^2 - 15\Big) - 8z\rho_{D^*}^2 + 1\Big), \\ R_1 &= R_1(1) + 0.05(w-1)^2 - 0.12(w-1), \\ R_2 &= R_2(1) - 0.06(w-1)^2 + 0.11(w-1), \\ R_3 &= R_3(1) + 0.026(w-1)^2 - 0.052(w-1), \end{aligned} \tag{A.2.8}$$

where the values of w have been conformally mapped on the open unit disc via [60, p. 7]

$$z(w) = \frac{\sqrt{w+1} - \sqrt{2}}{\sqrt{w+1} + \sqrt{2}}. \tag{A.2.9}$$

All parameters are discussed in section A.4. Finally, plots of the now fully specified form factors V, $A_{0,1,2}$ and $T_{1,2,3}$ are shown in figure A.1.

A.3 Comparison with other results

In the following, we compare this parametrization with the results for $\bar{B} \longrightarrow D^* l^- \bar{\nu}_l$ from [59]. Without considering the D^* decay, the χ angle dependency is trivial and the angle θ_{D^*} is not defined. Thus, we consider the cross section

$$\begin{aligned} \frac{\mathrm{d}^2\Gamma(\bar{B} \longrightarrow D^* \ell^- \bar{\nu}_\ell)}{\mathrm{d}q^2\, \mathrm{d}\cos\theta_\ell} &= \frac{1}{\mathrm{Br}(D^* \to D\pi)} \int_0^{2\pi} \mathrm{d}\chi \int_0^{\pi} \mathrm{d}\theta_{D^*} \frac{-1}{\sin\theta_\ell} \frac{\mathrm{d}^4\Gamma(\bar{B} \longrightarrow D^*(\to D\pi)\ell^- \bar{\nu}_\ell)}{\mathrm{d}q^2\, \mathrm{d}\theta_l\, \mathrm{d}\theta_{D^*}\, \mathrm{d}\chi} = \\ &= \frac{9}{32} \frac{N_F}{\mathrm{Br}(D^* \to D\pi)} \sum_a \sum_{i=1}^{13} c_i W_a^{(i)} \frac{-1}{\sin\theta_\ell} \int_0^{2\pi} \mathrm{d}\chi \int_0^{\pi} \mathrm{d}\theta_{D^*}\, B_a \end{aligned} \tag{A.3.1}$$

First note that only B_1^0, B_2^0, B_3^0 and B_1^T, B_2^T, B_3^T survive the χ integration. Furthermore, following [59], we assume $T_L = 0$ and $g_A = -g_V := g$, i.e. only $c_1 = c_2 = c_5 = |1-g|^2$ and $c_6 = \mathrm{Re}((1-g)g_P^*)$ are non-zero. Without further loss of generality we can therefore assume that g and g_P are both real. Thus, only 10 of the 41 $W_a^{(i)}$ coefficients in table A.1 are relevant: $W_1^{0(1)}$, $W_1^{0(6)}$, $W_2^{0(1)}$, $W_3^{0(1)}$, $W_3^{0(6)}$, $W_1^{T(1)}$, $W_1^{T(2)}$, $W_2^{T(1)}$, $W_2^{T(2)}$ and $W_3^{T(5)}$. With this, we get:

$$\frac{\mathrm{d}^2\Gamma(\bar{B} \longrightarrow D^* \ell^- \bar{\nu}_\ell)}{\mathrm{d}q^2\, \mathrm{d}\cos\theta_\ell} = \frac{9}{32} \frac{N_F}{\mathrm{Br}(D^* \to D\pi)} \cdot \frac{8\pi}{3} \cdot \left[\left(c_1 \frac{W_1^{0(1)}}{2} + c_6 \frac{W_1^{0(6)}}{2} + c_1 W_1^{T(1)} + c_2 W_1^{T(2)} \right) + \right.$$

$$+\sin^2\theta_\ell\Big(c_1\frac{W_2^{0(1)}}{2}+c_1W_2^{T(1)}+c_2W_2^{T(2)}\Big)+ \tag{A.3.2}$$
$$+\cos\theta_\ell\Big(c_1\frac{W_3^{0(1)}}{2}+c_6\frac{W_3^{0(6)}}{2}+c_5W_3^{T(5)}\Big)\Bigg].$$

Using the definition of N_F in (A.2.1), the factor in front becomes

$$A := \frac{G_F^2|p_{D^*}|^2|V_{cb}|^2q^2}{256\pi^3m_B^2}\Big(1-\frac{m_l^2}{q^2}\Big)^2, \tag{A.3.3}$$

matching the normalization factor of [59]. Comparing equations (A.2.3) with the form factor definitions of [59], we see that

$$\mathcal{A}_0^{(A)}=H_{00},\quad \mathcal{A}_t^{(A)}=H_{0t},\quad \mathcal{A}_P^{(P)}=\frac{\sqrt{q^2}}{m_b+m_c}H_{0t},\quad \mathcal{A}_{\pm'}^{(A)}\mp\mathcal{A}_{\pm'}^{(V)}=H_{\pm\pm}. \tag{A.3.4}$$

Plugging this in (A.3.2), we get

$$\begin{aligned}\frac{\mathrm{d}^2\Gamma}{\mathrm{d}q^2\,\mathrm{d}\cos\theta_\ell}=A\Bigg[&\Bigg(2c_1\frac{m_l^2}{q^2}(H_{0t}^2+H_{00}^2)+c_6\frac{4m_l}{m_b+m_c}H_{0t}^2+2c_1(H_{++}^2+H_{--}^2)\Bigg)+\\&+\sin^2\theta_\ell\Bigg(2c_1\Big(1-\frac{m_l^2}{q^2}\Big)H_{00}^2-c_1\Big(1-\frac{m_l^2}{q^2}\Big)(H_{++}^2+H_{--}^2)\Bigg)+\\&+\cos\theta_\ell\Big(-4c_1\frac{m_l^2}{q^2}H_{00}H_{0t}-c_6\frac{4m_l}{m_b+m_c}H_{00}H_{0t}+2c_5(H_{++}^2-H_{--}^2)\Big)\Bigg].\end{aligned} \tag{A.3.5}$$

Grouping by terms with and without the additional factor of m_ℓ^2/q^2 (which turns out to be equivalent to sorting by the lepton polarization):

$$\begin{aligned}\frac{\mathrm{d}^2\Gamma}{\mathrm{d}q^2\,\mathrm{d}\cos\theta_\ell}=&A(1-g)^2\Bigg[\Big((1-\cos\theta_\ell)^2H_{++}^2+(1+\cos\theta_\ell)^2H_{--}^2+2\sin^2\theta_\ell H_{00}^2\Big)+\\&+\frac{m_l^2}{q^2}\Bigg(\sin^2\theta_\ell(H_{++}^2+H_{--}^2)+2\cos^2\theta_\ell H_{00}^2+2H_{0t}^2\Big(1+\frac{g_P}{1-g}\frac{2q^2}{m_l(m_b+m_c)}\Big)+\\&-4H_{00}H_{0t}\cos\theta_\ell\Big(1+\frac{g_P}{1-g}\frac{q^2}{m_l(m_b+m_c)}\Big)\Bigg)\Bigg].\end{aligned} \tag{A.3.6}$$

Integrating over $\cos\theta_\ell$, we furthermore obtain:

$$\begin{aligned}\frac{\mathrm{d}\Gamma}{\mathrm{d}q^2}=\frac{8}{3}A(1-g)^2\Bigg[&(H_{++}^2+H_{--}^2+H_{00}^2)\Big(1+\frac{m_l^2}{2q^2}\Big)+\\&+\frac{3}{2}\frac{m_l^2}{q^2}\Big(1+\frac{g_P}{1-g}\frac{2q^2}{m_l(m_b+m_c)}\Big)H_{0t}^2\Bigg].\end{aligned} \tag{A.3.7}$$

This agrees with [59] up to a difference in the choice of definition for the coupling constants (as can be seen from their differing definition of the Lagrangian).

Another difference between [59] and our implementation following [116] is the implementation of the A_0 form factor, which is in direct proportionality to the amplitude H_{0t}. The paper [59] uses

$$A_0=\frac{m_B+m_{D^*}}{2\sqrt{m_Bm_{D^*}}}R_0(w)h_{A_1}(w), \tag{A.3.8}$$

with R_0 derived "using the results of [60]" as

$$R_0(w)=R_0(1)-0.11(w-1)+0.01(w-1)^2, \tag{A.3.9}$$

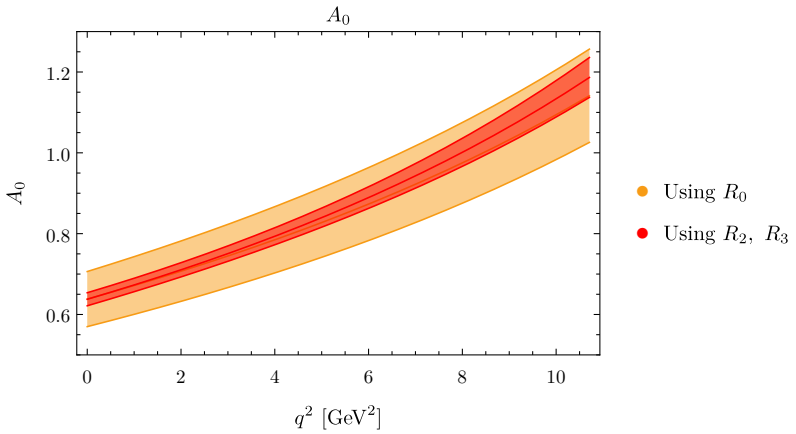


Figure A.2: Two different parametrizations of the form factor A_0 in [59] (black) vs [116, 118] (red)

where $R_0(1) = 1.14$. A 10% error is assumed on $R_0(1)$. In contrast, using the form factor parametrization of [118] given in equation (A.2.5b), we arrive at:

$$A_0 = \frac{m_B + m_{D^*}}{2\sqrt{m_B m_{D^*}}} \frac{1}{1 + r_{D^*}} \left((1 + w) - \frac{R_2 - R_3}{2} (r_{D^*}^{-1} - w) - \frac{R_2 + R_3}{2} (-r_{D^*} + w) \right), \quad \text{(A.3.10)}$$

a cubic expression in w. Both parametrizations are shown together in figure A.2. The differences are small for small q^2, then grow larger for higher q^2, probably due to the effect of the cubic term in the implementation following [118]. The effect of this discrepancy on the value of $R(D^*)$ is however small and the large error assumed in the implementation of [59] is enough to account for it.

A.4 Parameters and Errors

The constants for the CLN parametrization are taken from the 2016 results of the Heavy Flavor Averaging Group (HFLAV, [152, p. 138]):

$$\begin{aligned} \rho^2 &= 1.205 \pm 0.026, \\ R_1(1) &= 1.404 \pm 0.032, \\ R_2(1) &= 0.854 \pm 0.020. \end{aligned} \quad \text{(A.4.1)}$$

The form factor ratio R_3 cannot be extracted from $\bar{B} \longrightarrow D^* \ell^- \bar{\nu}_\ell$ studies, as it corresponds to the helicity suppressed part of the amplitude. We use $R_3(1) \simeq 1.19 - 0.26\eta(1) - 1.20\hat{\chi}_2(1)$ from [53, p. 10], where η and $\hat{\chi}_2$ are subleading Isgur-Wise functions. The same paper gives fit results for $\eta(1)$ and $\hat{\chi}_2(1)$ under various assumptions. Here, we have used the fit result using lattice calculations of $\mathcal{F}(1)$, $f_{+,0}(1)$ and $f_{+,0}(w > 1)$ as well as QCD sum rules, unitarity constraints and data from the Belle experiment, leading to $\eta(1) = 0.30 \pm 0.03$ and $\hat{\chi}_2(1) = -0.06 \pm 0.02$. It should be noted that fits without the QCD sum rules lead to very different results for $\hat{\chi}_2(1)$. This results in

$$R_3(1) = 1.184 \pm 0.121. \quad \text{(A.4.2)}$$

As noted in [53], the form factor parameters cannot be seen as completely independent. Thus our method to treat their uncertainties as independent sources of error (as previously assumed e.g. in [59]) is only a simplified (and flawed) approach.

For the form factor normalization $h_{A_1}(1)$, we take the result from the Fermilab Lattice and MILC collaborations [153, p. 29] (also quoted as average by [154, p. 163] in 2016):

$$h_{A_1}(1) = 0.906 \pm 0.004_{\text{stat}} \pm 0.012_{\text{syst}} = 0.906 \pm 0.013. \tag{A.4.3}$$

For the calculation of the differential cross section, the following additional constants from PDG 2018 [5] have been used:

$$\begin{aligned} m_{D^*} &= (2.00685 \pm 0.00005)\,\text{GeV}, \\ m_b &= 4.18^{+0.004}_{-0.003}\,\text{GeV}, \\ m_c &= 1.275^{+0.0025}_{-0.035}\,\text{GeV}, \\ m_\tau &= (1.77686 \pm 0.00012)\,\text{GeV}, \end{aligned} \tag{A.4.4}$$

$$\text{Br}(D^* \to D\pi) = (64.7 \pm 0.9)\%, \tag{A.4.5}$$

$$\frac{G_F}{(\hbar c)^3} = (1.1663787 \pm 0.0000005) \times 10^{-5}\,\text{GeV}^{-2}. \tag{A.4.6}$$

The mass of the B meson was taken as average of the masses of the B^0 and the $B^\pm$:

$$m_B = \frac{(5.227963 \pm 0.00015) + (5.27932 \pm 0.00014)}{2}\,\text{GeV} = (5.27948 \pm 0.00015)\,\text{GeV} \tag{A.4.7}$$

to which the difference $(m_{B^0} - m_{B^\pm})/2$ was added as additional error, resulting in a total error of 0.00031 GeV. The value of $|V_{cb}|$ has been taken to be $(39.05 \pm 0.47_{\text{exp}} \pm 0.58_{\text{th}}) \times 10^{-3}$ from [152, p. 139].

Errors on a function f of parameters p_i $(1 \le i \le N)$ with nominal value p_i^0 and lower (upper) absolute error p_i^- (p_i^+) have been calculated using the root of the sum of squares as follows:

$$f = f(p_1^0, \ldots, p_N^0) \pm \sqrt{\sum_{i=1}^{N} \max\left(\begin{matrix} (f(p_1^0,\ldots,p_N^0) - f(p_1^0,\ldots,p_i^0 - p_i^-,\ldots,p_N^0))^2, \\ (f(p_1^0,\ldots,p_N^0) - f(p_1^0,\ldots,p_i^0 + p_i^+,\ldots,p_N^0))^2 \end{matrix} \right)}. \tag{A.4.8}$$

The errors on the differential cross sections $\mathrm{d}\Gamma/\,\mathrm{d}v$ are shown in figure 2.1. The results for the integrated cross sections and $R(D^*)$ are:

$$\begin{aligned} \Gamma(\bar{B} \longrightarrow D^*(\to D\pi)\tau^-\bar{\nu}_\tau) &= (3.42 \quad \pm 0.18 \quad) \times 10^{-15}\,\text{GeV}, \\ \Gamma(\bar{B} \longrightarrow D^*(\to D\pi)\mu^-\bar{\nu}_\mu) &= (1.003 \pm 0.052) \times 10^{-14}\,\text{GeV}, \\ \Gamma(\bar{B} \longrightarrow D^*(\to D\pi)e^-\bar{\nu}_e) &= (1.005 \pm 0.052) \times 10^{-14}\,\text{GeV}, \\ R(D^*)|_\mu &= 0.2536 \pm 0.0020, \\ R(D^*)|_e &= 0.2528 \pm 0.0020, \\ R(D^*)|_\ell &= 0.2532 \pm 0.0020, \end{aligned} \tag{A.4.9}$$

where the value $R(D^*)|_\ell$ was calculated as

$$R(D^*)|_\ell = \frac{\int_{q^2_{\text{min}}}^{q^2_{\text{max}}} \frac{\mathrm{d}\Gamma(\bar{B} \longrightarrow D^*(\to D\pi)\tau^-\bar{\nu}_\tau)}{\mathrm{d}q^2}}{\frac{1}{2}\int_0^{q^2_{\text{max}}} \mathrm{d}q^2 \left(\frac{\mathrm{d}\Gamma(\bar{B} \longrightarrow D^*(\to D\pi)e^-\bar{\nu}_e)}{\mathrm{d}q^2} + \frac{\mathrm{d}\Gamma(\bar{B} \longrightarrow D^*(\to D\pi)\mu^-\bar{\nu}_\mu)}{\mathrm{d}q^2} \right)}. \tag{A.4.10}$$

To get a feeling for the importance of each error, table A.2 lists the most important sources of error for the total cross section $\Gamma(\bar{B} \longrightarrow D^*(\to D\pi)\tau^-\bar{\nu}_\tau)$ and $\Gamma(\bar{B} \longrightarrow D^*(\to D\pi)\ell^-\bar{\nu}_\ell)$ as well as $R(D^*)$.

Error	$\Gamma(\bar{B}\longrightarrow D^*(\rightarrow D\pi)\tau^-\bar{\nu}_\tau)$	$\Gamma(\bar{B}\longrightarrow D^*(\rightarrow D\pi)\ell^-\bar{\nu}_\ell)$	$R(D^*)$
V_{cb}	3.88%	3.88%	
$h_{A_1}(1)$	2.89%	2.89%	
$\mathrm{BR}(D^*\longrightarrow D\pi)$	1.39%	1.39%	
$\rho^2_{D^*}$	0.79%	1.04%	0.60%
$R_2(1)$	0.6%	0.57%	0.32%
$R_3(1)$	0.42%		0.42%
$R_1(1)$	0.26%	0.34%	0.06%
m_B	0.07%	0.04%	0.03%
m_τ	0.02%		0.02%
m_{D^*}	0.01%		0.01%
Total	5.14%	5.18%	0.80%

Table A.2: Sources of error for the total cross section $\Gamma(\bar{B}\longrightarrow D^*(\rightarrow D\pi)\tau^-\bar{\nu}_\tau)$ and $\Gamma(\bar{B}\longrightarrow D^*(\rightarrow D\pi)\ell^-\bar{\nu}_\ell)$ as well as $R(D^*)$.

Appendix B

Mathematical Background on Finding Weights

Lemma B.0.1. *Let $\{f_a\}_{1\leq a\leq n}$ be a set of integrable functions $f_a\colon \mathbb{R}^k \supseteq U \longrightarrow \mathbb{R}$ with U bounded. Furthermore, let $\{g_i\}_{1\leq i\leq N}$ be a set of integrable functions $g_i\colon U\longrightarrow\mathbb{R}$. Then the following two statements are equivalent:*

i) For any $1 \leq b \leq n$, there exist weights $\omega_b(i) \in \mathbb{R}$, $1 \leq i \leq N$ such that

$$\delta_{ab} \overset{!}{=} \int_U \mathrm{d}x^k \, f_a(\vec{x})\tilde{f}_b(\vec{x}) \quad \forall\, 1 \leq a \leq n \quad \textit{with} \quad \tilde{f}_b(\vec{x}) := \sum_{i=1}^{N} \omega_b(i) g_i(\vec{x}) \tag{B.0.1}$$

ii) The $n \times N$ matrix $(A_{ai})_{ai}$ with

$$A_{ai} := \int_U \mathrm{d}x^k \, f_a(\vec{x}) g_i(\vec{x}) \tag{B.0.2}$$

has rank n.

In particular both statements imply $N \geq n$. If $N = n$, then the weights $\omega_b(i)$ are unique. If $N > n$, then for each b, the possible choices $(w_b(i))_i \in \mathbb{R}^N$ form an affine vector space of dimension $N - n$.

Proof. Using the definitions of the matrix A and for $\tilde{f}_b$, we have:

$$\begin{aligned} \int_U \mathrm{d}x^k \, f_a(\vec{x})\tilde{f}_b(\vec{x}) = \int_U \mathrm{d}x^k \, f_a(\vec{x}) \sum_{i=1}^{N} \omega_b(i) g_i(\vec{x}) = \sum_{i=1}^{N} \omega_b(i) \int_U \mathrm{d}x^k \, f_a(\vec{x}) g_i(\vec{x}) = \\ = \sum_{i=1}^{N} \omega_b(i) A_{ai} = (A\Omega)_{ij} \end{aligned} \tag{B.0.3}$$

with the $N \times n$ matrix $\Omega_{ib} = \omega_b(i)$ (i.e. the matrix that has the weights $\omega_b(i)$ as columns). Thus statement i) is equivalent to $A\Omega = \mathbb{1}$ with $\mathbb{1}$ the $n \times n$ identity matrix, i.e. equivalent to Ω being a right inverse of A. But a right inverse of a matrix exists if and only if the matrix has full row rank, i.e. has more or equally many columns than rows and full rank. That shows the equivalence i) $\iff$ ii) and that both imply $N \geq n$.

Provided that ii) is met, $\Omega = A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}$ is an explicit example of a right inverse. Furthermore, if Ω is a right inverse, then Ω' is another right inverse if and only if

$$0 \overset{!}{=} A\Omega - A\Omega' = A(\Omega - \Omega') =: A\,\Delta\Omega, \tag{B.0.4}$$

i.e. if $\Delta\Omega$ is a projector on $\ker A$. $\Delta\Omega$ is a projector on $\ker A$ if and only if its columns (which are the images of the base vectors) all lie in $\ker A$, a vector space of $\dim \ker A = N - n$. But the weight $(w_b(i))_i$ is exactly the b-th column of Ω and thus lies in an affine vector space of dimension $N - n$. More precisely:

$$(\omega_b(i))_i \in A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}\hat{e}_b + \ker A \tag{B.0.5}$$

with $\hat{e}_b = (\delta_{bb'})_{1\leq b'\leq n} \in \mathbb{R}^n$. □

Remark B.0.2. Note that this result could be even more generalized in terms of linear independent elements of a general inner product space[1] V, spanned by vectors f_i, i.e.

$$X = \mathbb{K}f_1 + \cdots + \mathbb{K}f_n \tag{B.0.6}$$

for a field $\mathbb{K}$. Now consider the dual space V^* of linear maps $V \longrightarrow \mathbb{K}$. There is a dual base ϕ_i of V^* corresponding to the base f_i of V, defined via the values at the base elements

$$\phi_i \colon V \longrightarrow \mathbb{K} \quad \text{with} \quad \phi_i(f_j) = \delta_{ij}. \tag{B.0.7}$$

By theorem of Riesz, there are $\tilde{f}_i \in V$, such that $\phi_i = \langle \tilde{f}_i, \cdot \rangle$.

Thus, what Lemma B.0.1 really tells us, is the equivalence of the following two statements:

i) The elements $\langle g_i, \cdot \rangle$ span V^*

ii) The matrix $A_{ij} = \langle f_i, g_j \rangle$ has full rank n.

Remark B.0.3. The degrees of freedom $\alpha_a(1), \ldots, \alpha_a(N-n)$ in the case $N > n = \operatorname{rank} A$ of the previous theorem can also be handled in a slightly more "explicit" manner. After possible renaming, divide the set of the g_i into $\{g_i\}_{1\leq i\leq n}$ and $\{g_i\}_{n<i\leq N}$ in such a way, that the $n \times n$ matrix $A' := (A_{ai})_{1\leq a,i\leq n}$ (with A_{ai} defined as in equation (B.0.2)) has full rank n. Note that we have $\sum_{i=1}^n A_{ai}(A'^{-1})_{ib} = \delta_{ab}$. Now define

$$\omega_b(i) = \begin{cases} \alpha_b(i-n) & \text{for } i = n+1, \ldots, N \\ (A'^{-1})_{ib} - \sum_{c=1}^n \sum_{j=n+1}^N \alpha_b(j-n) A_{cj} (A'^{-1})_{ic} & \text{for } i = 1, \ldots, n \end{cases}, \tag{B.0.8}$$

i.e. use the degrees of freedom to directly fix some of the weights. The weights (B.0.8) are indeed a solution to (B.0.1):

$$\begin{aligned} \int_U \mathrm{d}x^k \, f_a(\vec{x}) \tilde{f}_b(\vec{x}) &= \int_U \mathrm{d}x^k x \, f_a(\vec{x}) \sum_{i=1}^N \omega_b(i) g_i(\vec{x}) = \sum_{i=1}^N \omega_b(i) A_{ai} = \\ &= \sum_{i=1}^n \left((A'^{-1})_{ib} A_{ai} - \sum_{c=1}^n \sum_{j=n+1}^N \alpha_b(j-n) A_{cj} (A'^{-1})_{ic} A_{ai} \right) + \sum_{i=n+1}^N \alpha_b(i-n) A_{ai} = \\ &= \delta_{ab} - \sum_{i=1}^n \sum_{c=1}^n \sum_{j=n+1}^N \alpha_b(j-n) A_{cj} \underbrace{A_{ai}(A'^{-1})_{ic}}_{\longrightarrow \delta_{ac}} + \sum_{i=n+1}^N \alpha_b(i-n) A_{ai} = \\ &= \delta_{ab} - \sum_{j=n+1}^N \alpha_b(j-n) A_{aj} + \sum_{i=n+1}^N \alpha_b(i) A_{ai} = \delta_{ab}. \end{aligned} \tag{B.0.9}$$

[1] In the above case the inner product was given by the an integral over U: $\langle f_i, f_j \rangle = \int_u \mathrm{d}x^k \, f_i(\vec{x}) f_j(\vec{x})$ for $f_i, f_j \in V$.

More specifically we now restrict ourselves to *step functions* for $\tilde{f}_b$ by considering lemma B.0.1 for the specific case of

$$g_i(\vec{x}) = \mathbb{1}_{U_i}(\vec{x}) = \begin{cases} 1 & \vec{x} \in U_i \\ 0 & \text{else} \end{cases} \tag{B.0.10}$$

for subsets $U_i \subseteq U$, $1 \leq i \leq N$. In this setting, lemma B.0.1 takes the form of

Corollary B.0.4. *Consider a set $\{f_a\}_{1\leq a\leq n}$ of integrable functions $f_a \colon \mathbb{R}^k \supseteq U \longrightarrow \mathbb{R}$ with U bounded and subsets "bins" $U_i \subseteq U$ with $1 \leq i \leq N$. Define the "bin content":*

$$F_a(i) = \int_{U_i} \mathrm{d}x^k \, f_a(\vec{x}) \tag{B.0.11}$$

Then the following two statements are equivalent:

i) For any $1 \leq b \leq n$, there exist weights $\omega_b(i) \in \mathbb{R}$, $1 \leq i \leq N$ such that

$$\delta_{ab} \overset{!}{=} \sum_{i=1}^{N} \omega_b(i) F_a(i) \quad \forall\, 1 \leq a \leq n. \tag{B.0.12}$$

ii) The $n \times N$ matrix $\left(F_a(i)\right)_{ai}$ has full rank n.

In particular, i) implies $N \geq n$. If $N = n$, then the weights $\omega_b(i)$ are unique. If $N > n$, then for each i, the possible choices $(\omega_b(i))_i \in \mathbb{R}^N$ form an affine vector space of dimension $N - n$.

Proof. Note that we could have also written (B.0.12) as

$$\delta_{ab} = \int_U \mathrm{d}x^k \, f_a(\vec{x}) \tilde{f}_b(\vec{x}) \quad \text{with} \quad \tilde{f}_b(\vec{x}) = \sum_a \omega_b(i) \mathbb{1}_{U_i}. \tag{B.0.13}$$

Thus this is Lemma B.0.1 with the functions g_i as in (B.0.10). □

One last thing that remains of interest is, whether we can actually always find $N = n$ bins U_i such that the two statements of Corollary B.0.4 are true. As a result of the following two lemmas, this is in fact always possible:

Lemma B.0.5. *Consider a set $\{f_a\}_{1\leq a\leq n}$ of integrable functions*

$$f_a \colon \mathbb{R}^k \supseteq U := [x_1^-, x_1^+] \times \cdots [x_n^-, x_n^+] \longrightarrow \mathbb{R}. \tag{B.0.14}$$

Then the following two statements are equivalent:

(i) The $\{f_a\}$ are linearly independent

(ii) The integrated functions

$$F_a \colon \begin{array}{l} U \longrightarrow \mathbb{R} \\ \vec{x} \longmapsto \displaystyle\int_{p_1}^{x_1} \mathrm{d}x_1' \cdots \int_{p_n}^{x_n} \mathrm{d}x_n' \; f_a(\vec{x}') \end{array} \tag{B.0.15}$$

are linearly independent.

Proof. Clear by the linearity of the integral and derivative. □

Lemma B.0.6. *Consider a set* $\{f_a\}_{1\leq a\leq n}$ *of functions* $\mathbb{R}^n \supseteq U \longrightarrow \mathbb{R}$. *Then the following statements are equivalent:*

(i) The $\{f_a\}$ *are linearly independent*

(ii) We can find $\vec{x}_b \in U$, $1 \leq b \leq n$ *such that the* $n \times n$ *matrix* $\left(f_a(\vec{x}_b)\right)_{1\leq a,b\leq n}$ *has full rank* n.

Proof. $\neg$(i)$\Longrightarrow\neg$(ii) is obvious. For (i) $\Longrightarrow$ (ii), suppose that the $\{f_a\}$ are linearly independent, but that for any n points $\vec{x}_a \in U$, the matrix $(f_a(x_b))_{1\leq a,b\leq n}$ has rank less than n. Consider the set

$$M = \{(f_a(\vec{x}))_{1\leq a\leq n} \mid \vec{x} \in U\} \subseteq \mathbb{R}^n. \tag{B.0.16}$$

Then our assertion states that any n vectors of M are not linearly independent. However, if any n points of a subset of a n dimensional vector space M are linearly dependent, then they lie in a hyperplane $\langle M\rangle$ of the vector space with $m := \dim\langle M\rangle < n$.[2] But that means that I can find fixed $0 \neq (\alpha_a)_{1\leq a\leq n}$, such that

$$0 = \sum \alpha_a\, f_a(\vec{x}) \quad \forall\, \vec{x} \in U, \tag{B.0.17}$$

i.e. the $\{f_a\}$ are not linearly independent. □

The following lemma will also be of some interest when exploiting symmetry properties of functions to find weights:

Definition B.0.7.

(i) A function $f\colon \mathbb{R} \supseteq [x_-, x_+] \longrightarrow \mathbb{R}$, $x_- < x_+$ is called (anti)symmetric if and only if $f(x) = \pm f(x_- + x_+ - x)$ for all $x_- \leq x \leq x_+$.

(ii) A binning $U_i = [x_i, x_{i+1}] \subseteq \mathbb{R}$, $x_i < x_{i+1}$, $0 \leq i \leq N-1$ is called symmetric if and only if $x_i + x_{N-i} = x_0 + x_N$ for all $0 \leq i \leq N$.

Lemma B.0.8. *Consider* N *symmetric bins* $U_i = [x_i, x_{i+1}] \subseteq \mathbb{R}$, $x_i < x_{i+1}$, $0 \leq i \leq N-1$ *and functions* $\{f_a\colon [x_0, x_{N+1}] \longrightarrow \mathbb{R}, 1 \leq a \leq n\}$. *Suppose now that the solutions* $\{\tilde{f}_a\}$ *from Corollary B.0.4 exist. Then*

(i) $\#\{f_a \mid \textit{symmetric}\} \leq \lceil N/2\rceil$ *and* $\#\{f_a \mid \textit{antisymmetric}\} \leq \lfloor N/2\rfloor$

(ii) If all functions are either symmetric or antisymmetric, a solution for $\{\tilde{f}_i\}$ *can be found by independently finding weights to distinguish between*

1. $\{f_a \mid \textit{symmetric}\}$ *on* $[x_0, x_1], \ldots, [x_{\lceil N/2\rceil - 1}, x_{\lceil N/2\rceil}]$
2. $\{f_a \mid \textit{antisymmetric}\}$ *on* $[x_0, x_1], \ldots, [x_{\lfloor N/2\rfloor - 1}, x_{\lfloor N/2\rfloor}]$,

leading to $\lceil N/2\rceil$ *or* $\lfloor N/2\rfloor$ *weights, respectively, and then defining the remaining weights by employing the symmetry requirements from (iii).*

(iii) If all functions are either symmetric or antisymmetric and $N = n$, *then* f_a *(anti)symmetric* $\Longleftrightarrow$ $\tilde{f}_a$ *(anti)symmetric* $\Longleftrightarrow$ $\omega_a(i) = \pm\omega_a(N+1-i)\ \forall\, 1 \leq i \leq N$

[2] Let $M \subseteq V$ be a subset of a $\mathbb{K}$ vector space and $n := \dim V$. Suppose that any n points $x_1, \ldots, x_n$ of M are linearly dependent and consider the corresponding hyperplanes $\langle x_1, \ldots, x_n\rangle := \mathbb{K}x_1 + \cdots + \mathbb{K}x_n$. Let A be the hyperplane with maximal dimension $m < n$. Suppose that $M \nsubseteq A$. Pick a random element $m \in M\backslash A$. Since the $x_1, \cdots x_n$ that generate A are linearly dependent, we can (after possible renaming) assume that $\langle x_1, \cdots x_{n-1}\rangle = A$. But then $\dim\langle x_1, \cdots x_{n-1}, a\rangle > \dim A$ in contradiction to $\dim A$ being maximal. Thus $M \subseteq A$, i.e. M is contained in a lower dimensional hyperplane of V.

Proof.

(i) For the "binned" functions $F_a(i)$, we have $F_a(N+1-i) = \pm F_a(i)$ for all $1 \leq i \leq N$. But then

$$\begin{aligned} \delta_{ab} \overset{!}{=} \sum_{i=1}^{N} \omega_b(i) F_a(i) = \sum_{i=1}^{\lfloor N/2 \rfloor} \omega_b(i) F_a(i) + \sum_{i=1}^{\lceil N/2 \rceil} \omega_b(N+1-i) F_a(N+1-i) = \\ = \sum_{i=1}^{\lfloor N/2 \rfloor} (\omega_b(i) \pm \omega_b(N+1-i)) F_a(i) + \begin{cases} \omega_b(\lceil N/2 \rceil) F_a(\lceil N/2 \rceil) & N \text{ odd} \\ 0 & \text{else.} \end{cases} \end{aligned} \tag{B.0.18}$$

Thus we found a way to distinguish between all antisymmetric functions in $\lfloor N/2 \rfloor$ (note that for antisymmetric f_i, we have $F_i(\lfloor N/2 \rfloor) = 0$ for odd N) and between all symmetric functions in $\lceil N/2 \rceil$ bins. But to distinguish between m functions we always need at least m bins (Corollary B.0.4).

(ii) Lemmata B.0.6 and B.0.5 together with (i) show that it is possible to find the weights for the two cases as described. It is obvious that this solution can then be extended to one for all $\{f_i\}$ by employing the symmetry properties as described.

(iii) If $N = n$, then the solutions $\{f_i\}$ are unique. Thus the only possible solution is the one found in (ii), which fulfills the claim by definition. □

Note that we can also do this recursively, thus reducing the whole problem to different symmetry properties. This applies for example to the functions in χ in uniform bins of $\pi/4$ width (see figure 2.13 and table 2.2).

Appendix C

Supplementary Figures

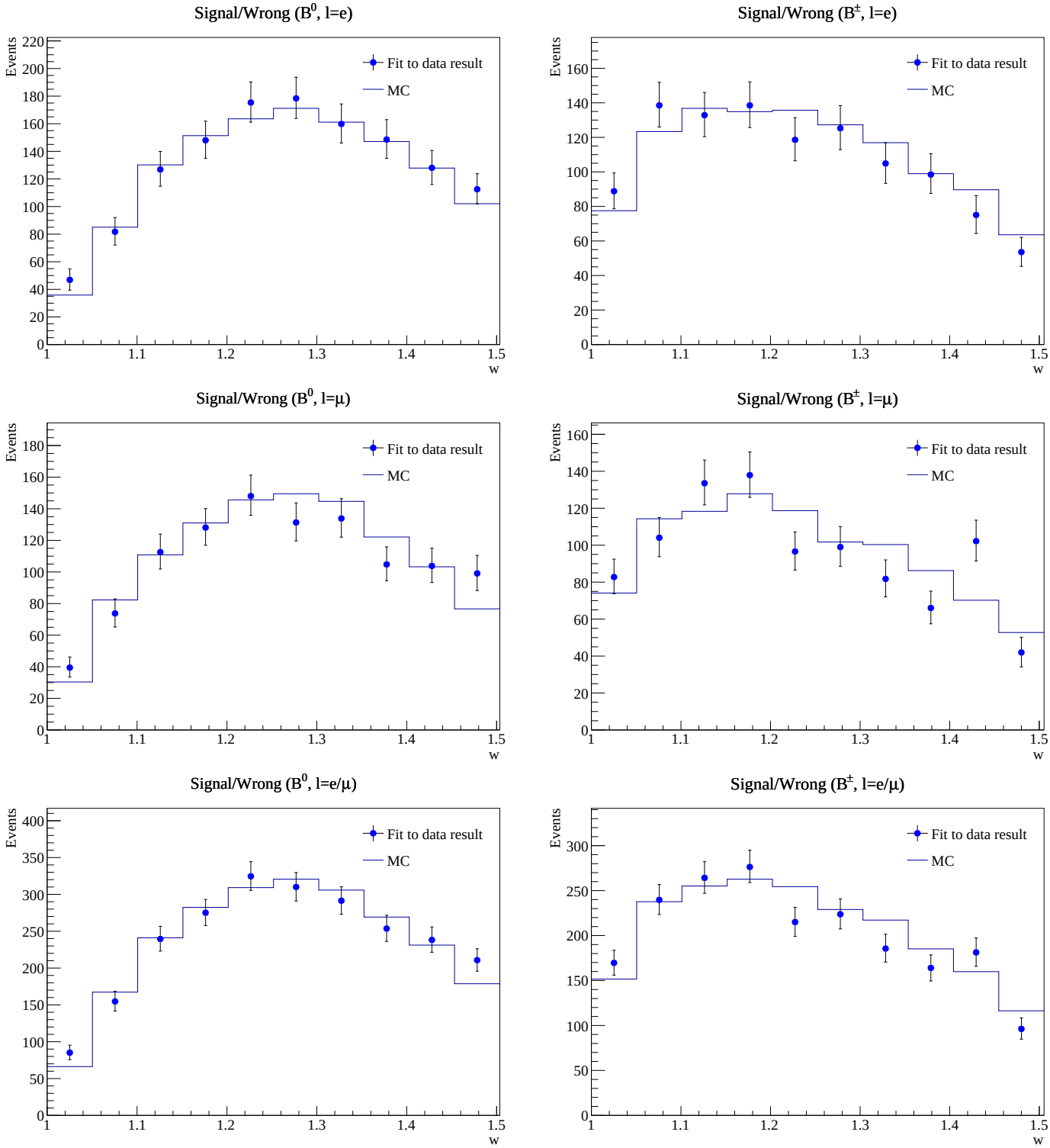


Figure C.1: Fit results for binning in w.

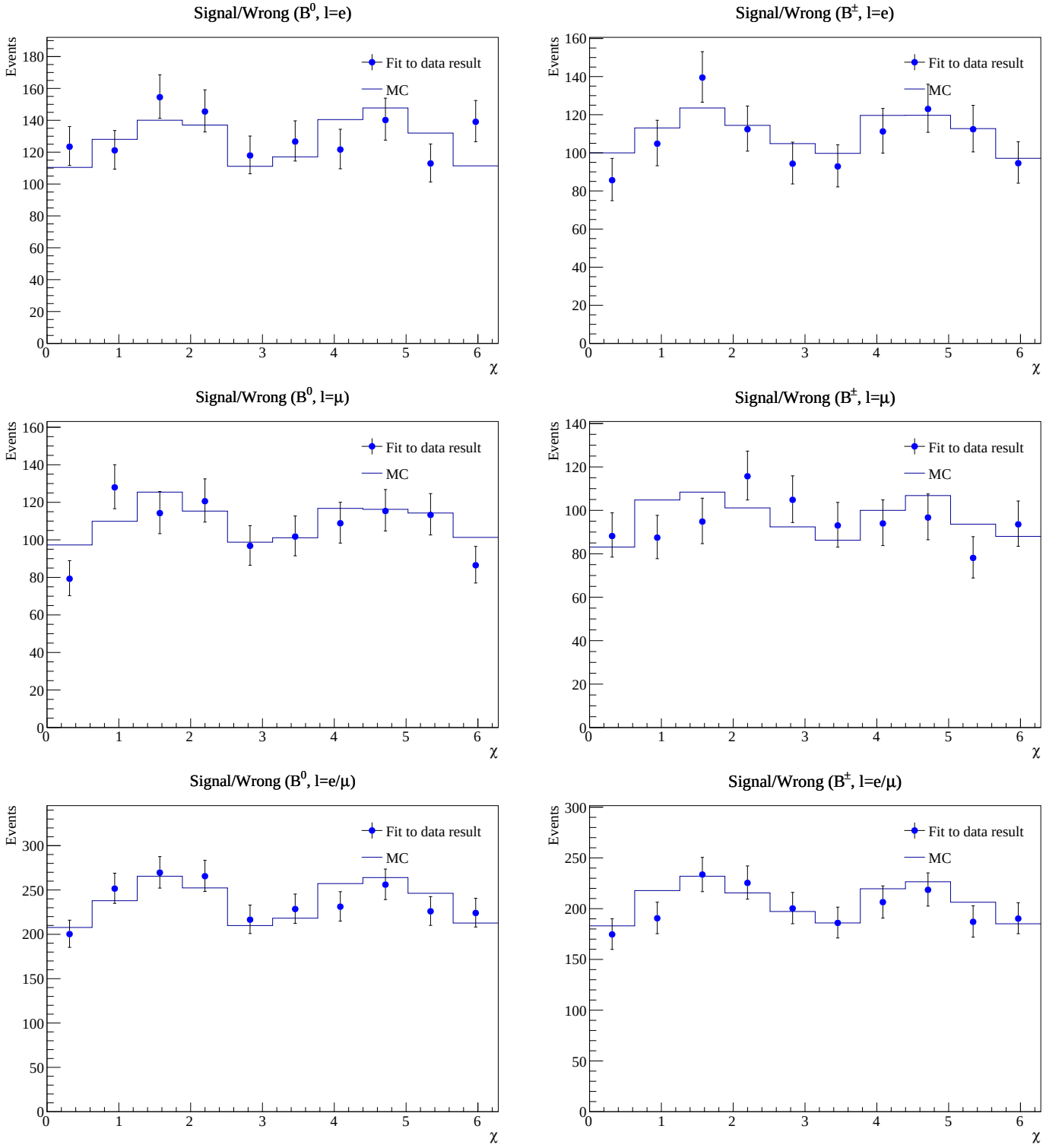


Figure C.2: Fit results for binning in χ.

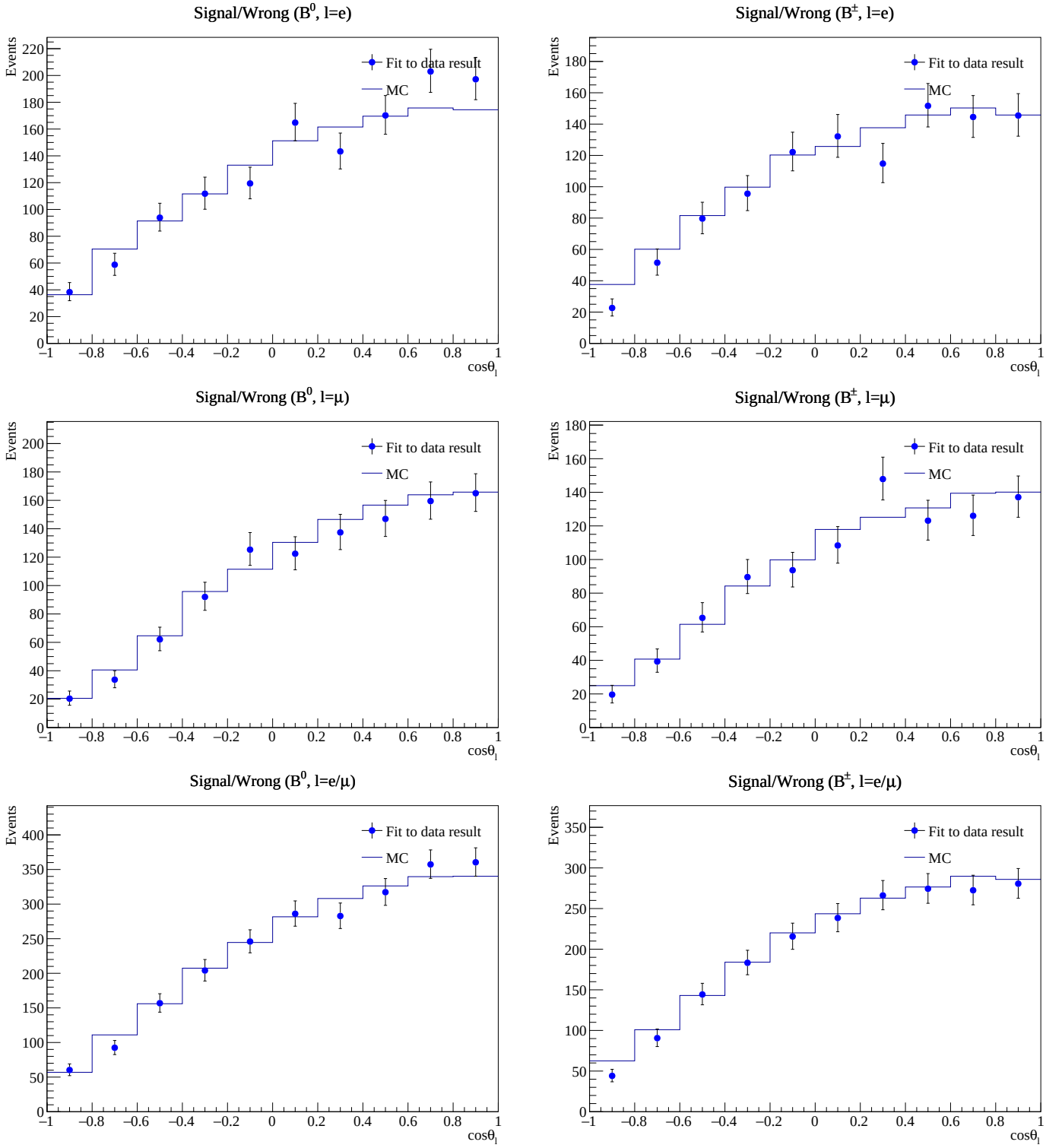


Figure C.3: Fit results for binning in $\cos\theta_\ell$.

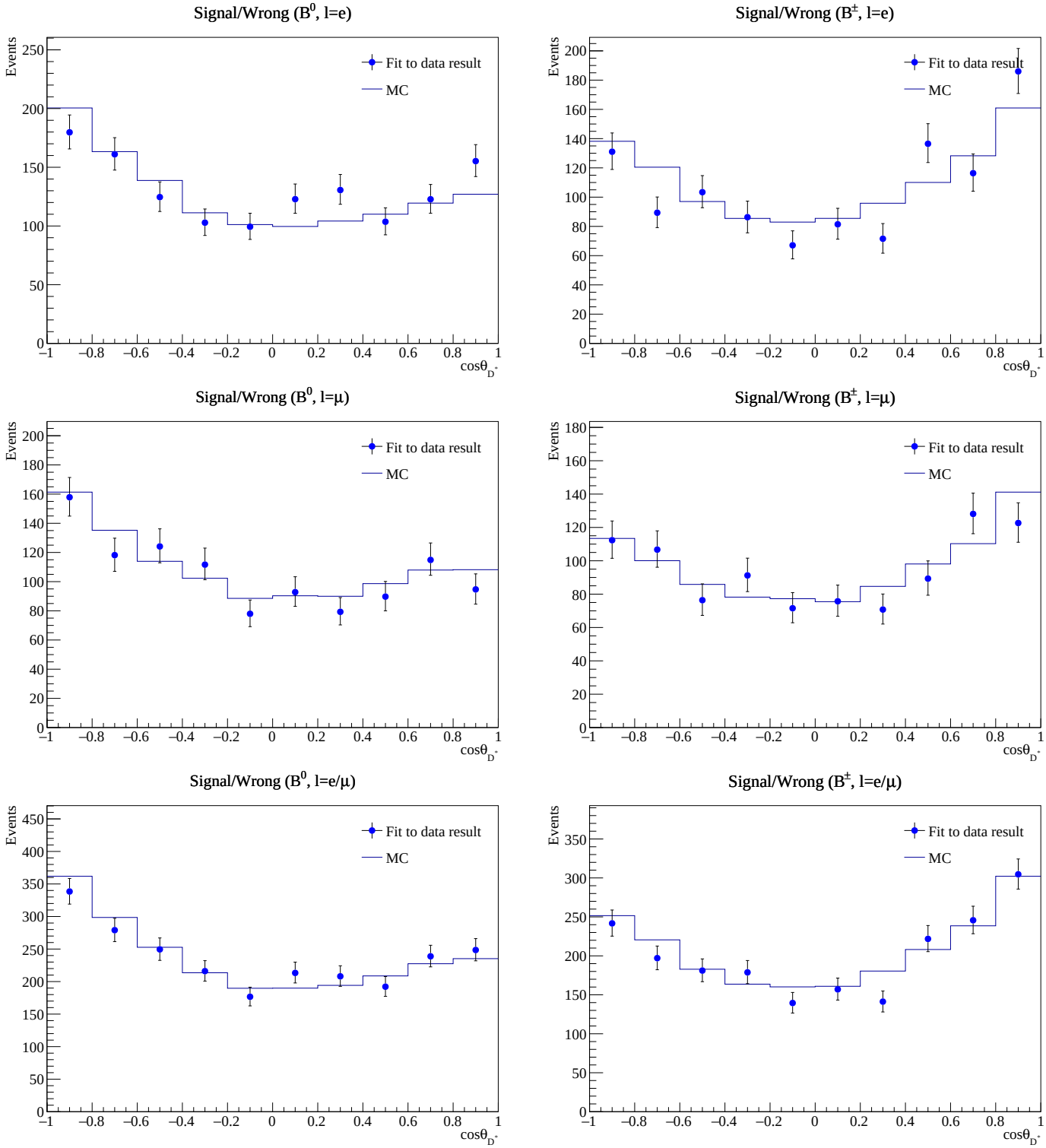


Figure C.4: Fit results for binning in $\cos\theta_{D^*}$.

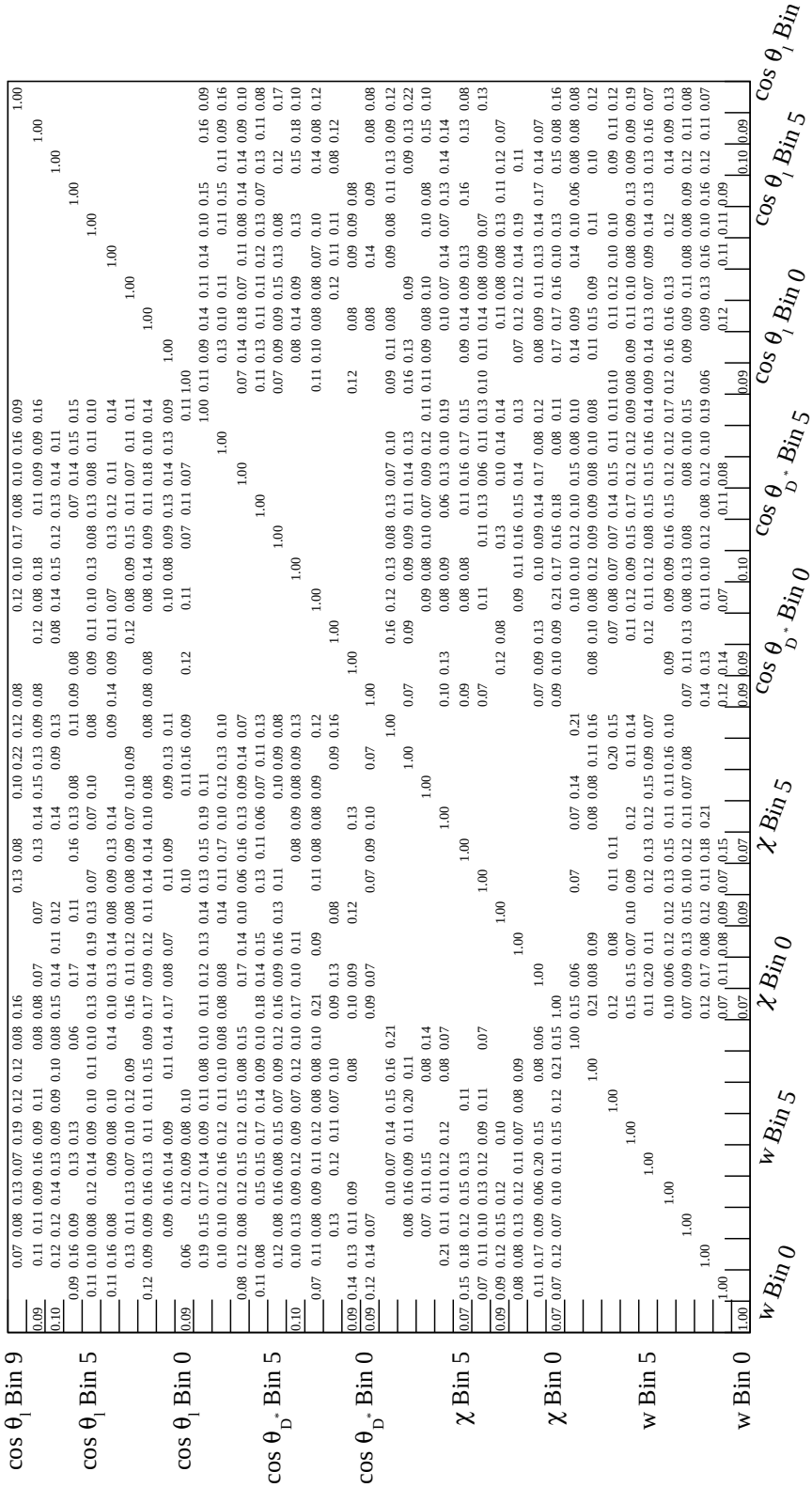


Figure C.5: Correlation matrix for (pre-unfolding) statistical errors of B^0 for the combined fit $\ell = e/\mu$. Note that the 10×10 submatrices $\mathrm{Corr}(E^{v,i}_{\mathrm{stat}}, E^{v,i'}_{\mathrm{stat}})$ are constrained to unity and correlations smaller than the doubled error on the correlation (around $2 \cdot 3\%$) are set to zero (as well as negative correlations).

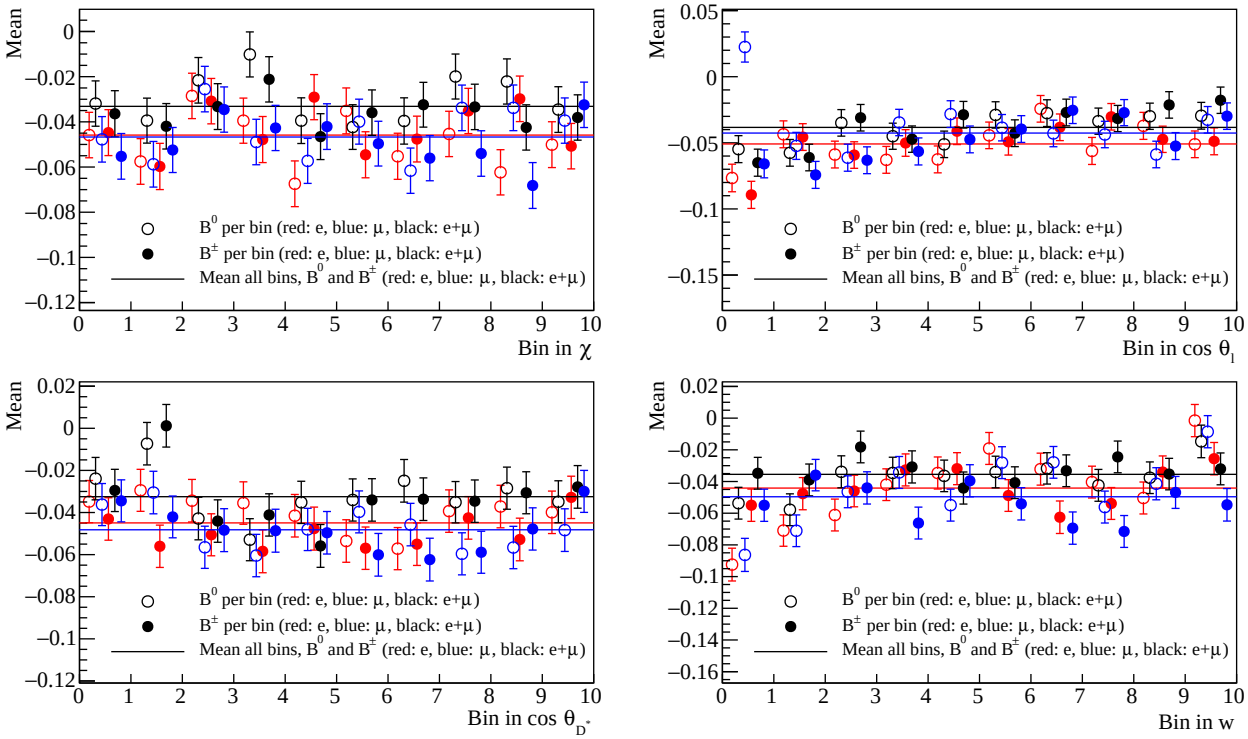


Figure C.6: Overview of the mean of the pull distributions for the toys. Note: The markers for electrons, muons and the combined fit have been slightly shifted to enhance readability. A total of 1000 toys were generated for this closure test.

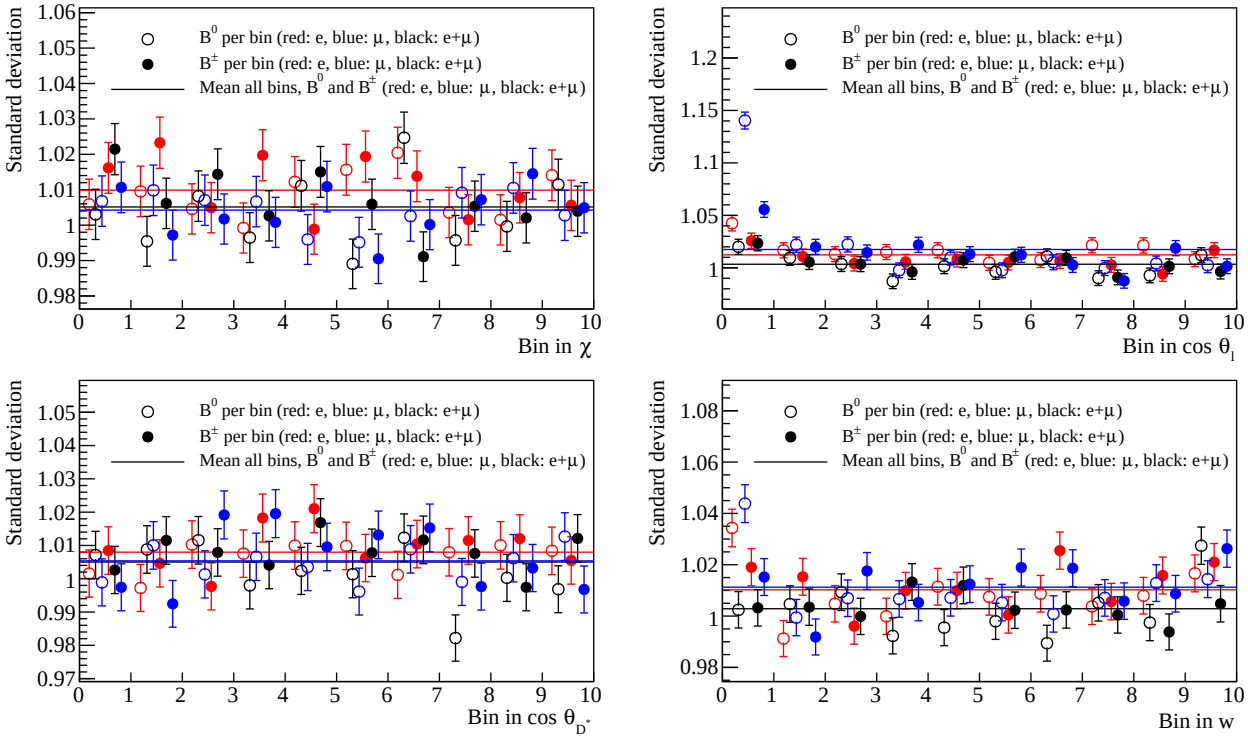


Figure C.7: Overview of the standard deviation of the pull distributions for the toys. Additional notes are given in figure C.6.

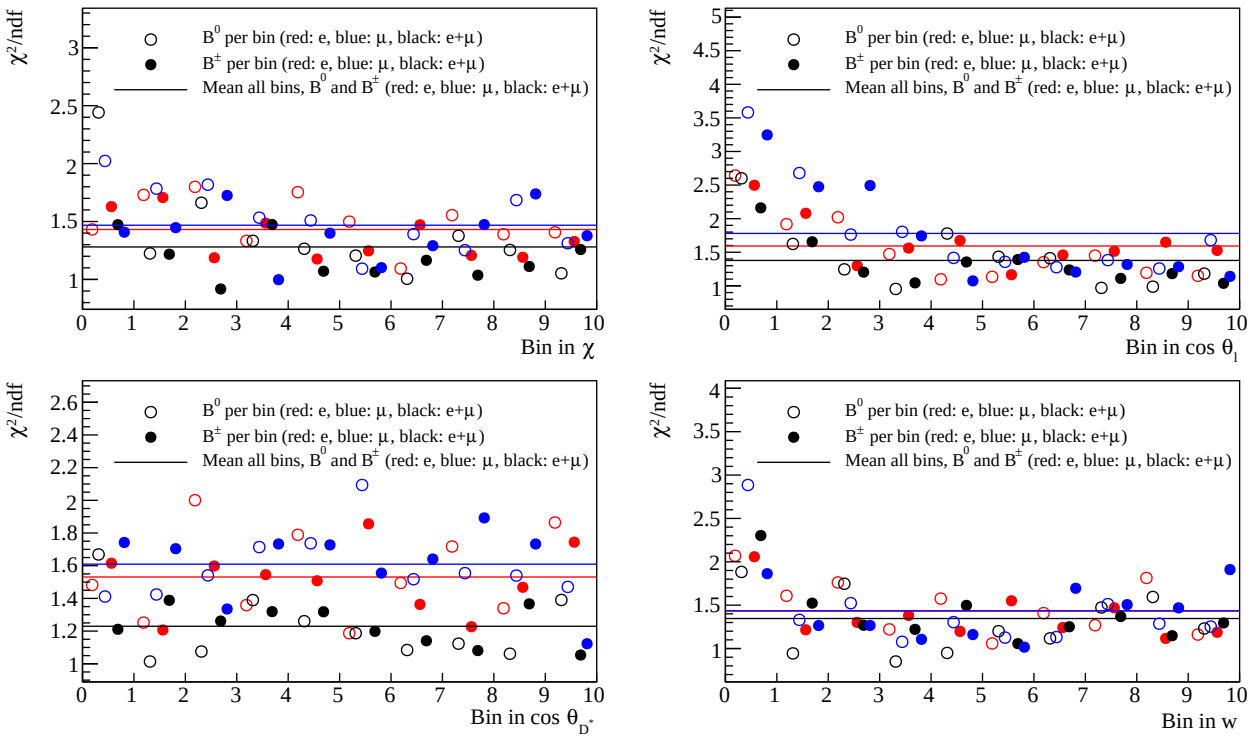


Figure C.8: Overview of the χ^2/n.d.f value of the fit of the histograms to Gaussians. Additional notes are given in figure C.6.

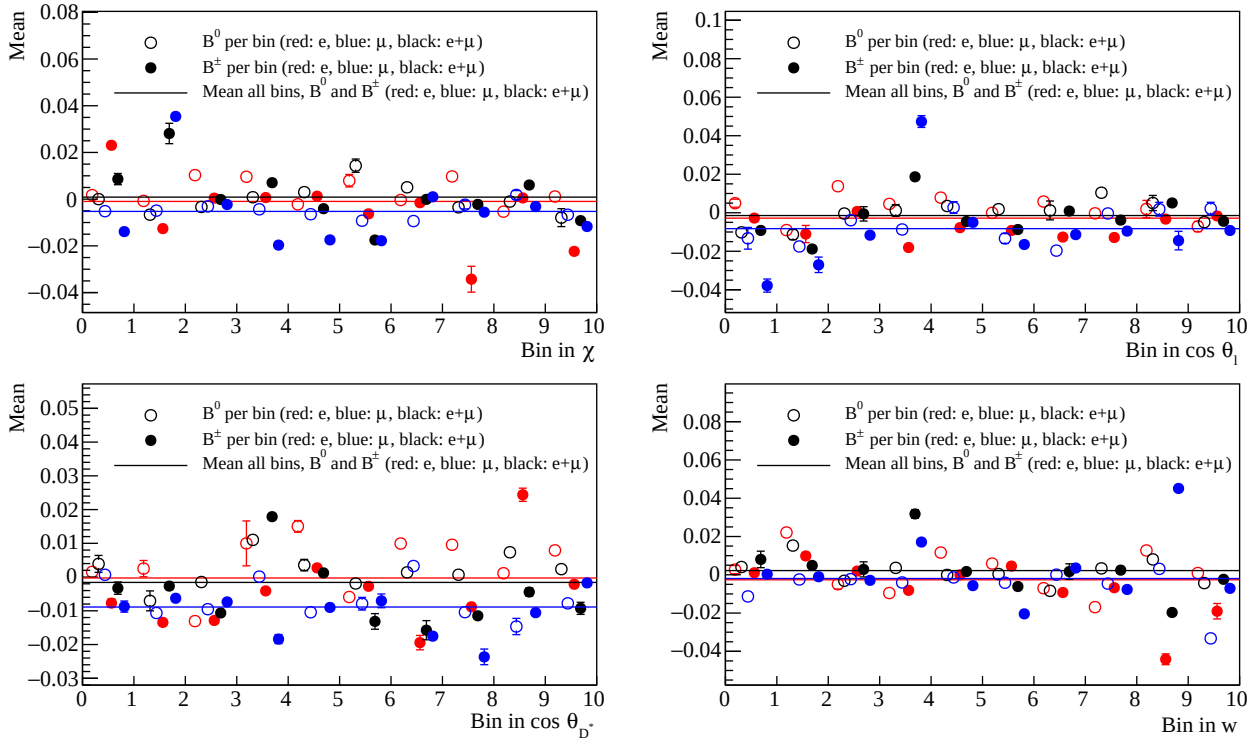


Figure C.9: Overview of the mean of the pull distributions for toys with varied MC (MC where one of three $B \longrightarrow D^*$ form factor parameters is varied by $\pm 1\sigma$, leading to six results per bin, B type and lepton type). Note: The markers for electrons muons and the combined fit results have been slightly shifted to enhance readability. Also note that the errors shown do not take the error of the single Gaussian fit into account, but only the deviations for the six different results for the MC variations. A total of 1000 toys were generated for the closure test for each MC variation.

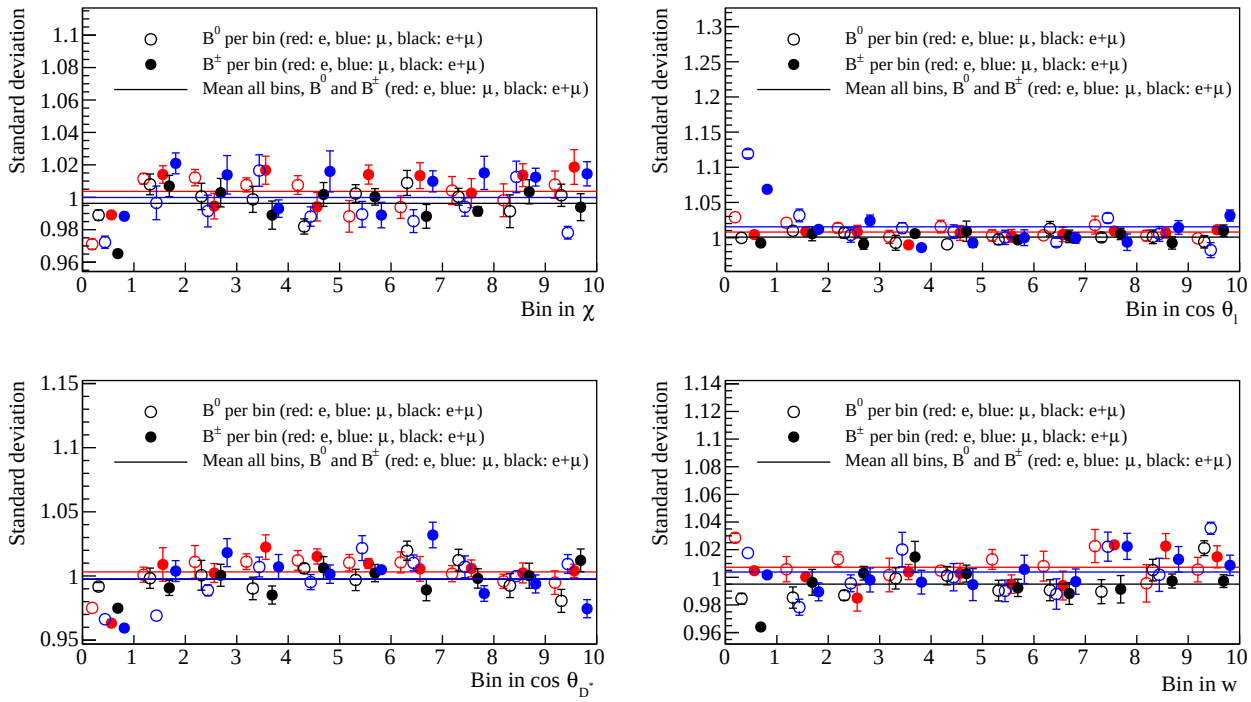


Figure C.10: Overview of the standard deviation of the pull distributions for the toys with varied MC. Additional notes are given in figure C.9.

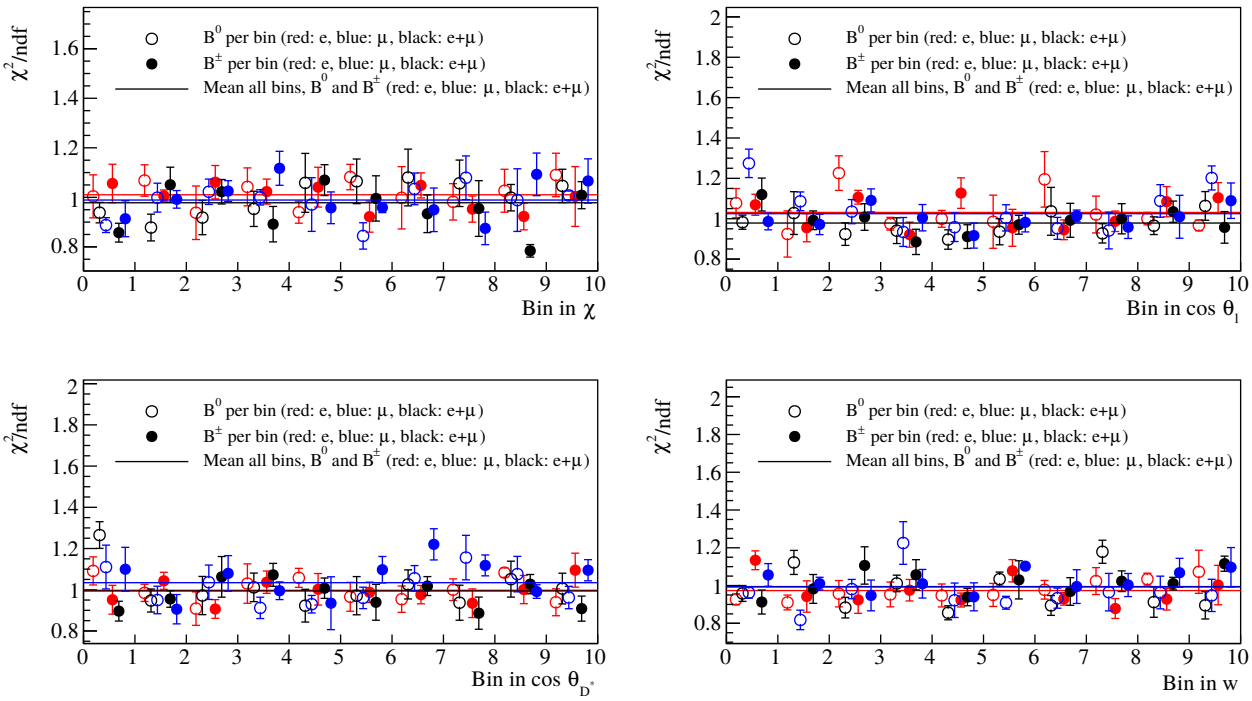


Figure C.11: Overview of the χ^2/n.d.f value of the fit of the histograms to Gaussians with varied MC. Additional notes are given in figure C.9.

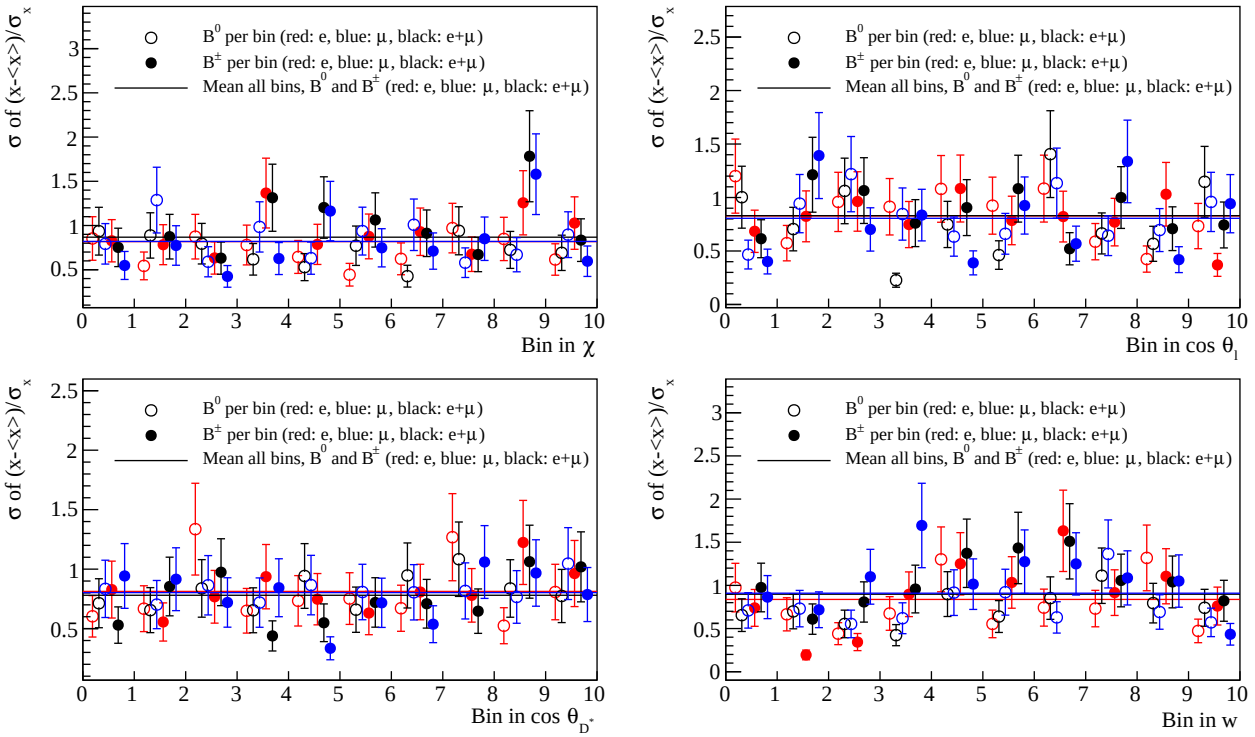


Figure C.12: Standard deviations of the six results in closure test 1.

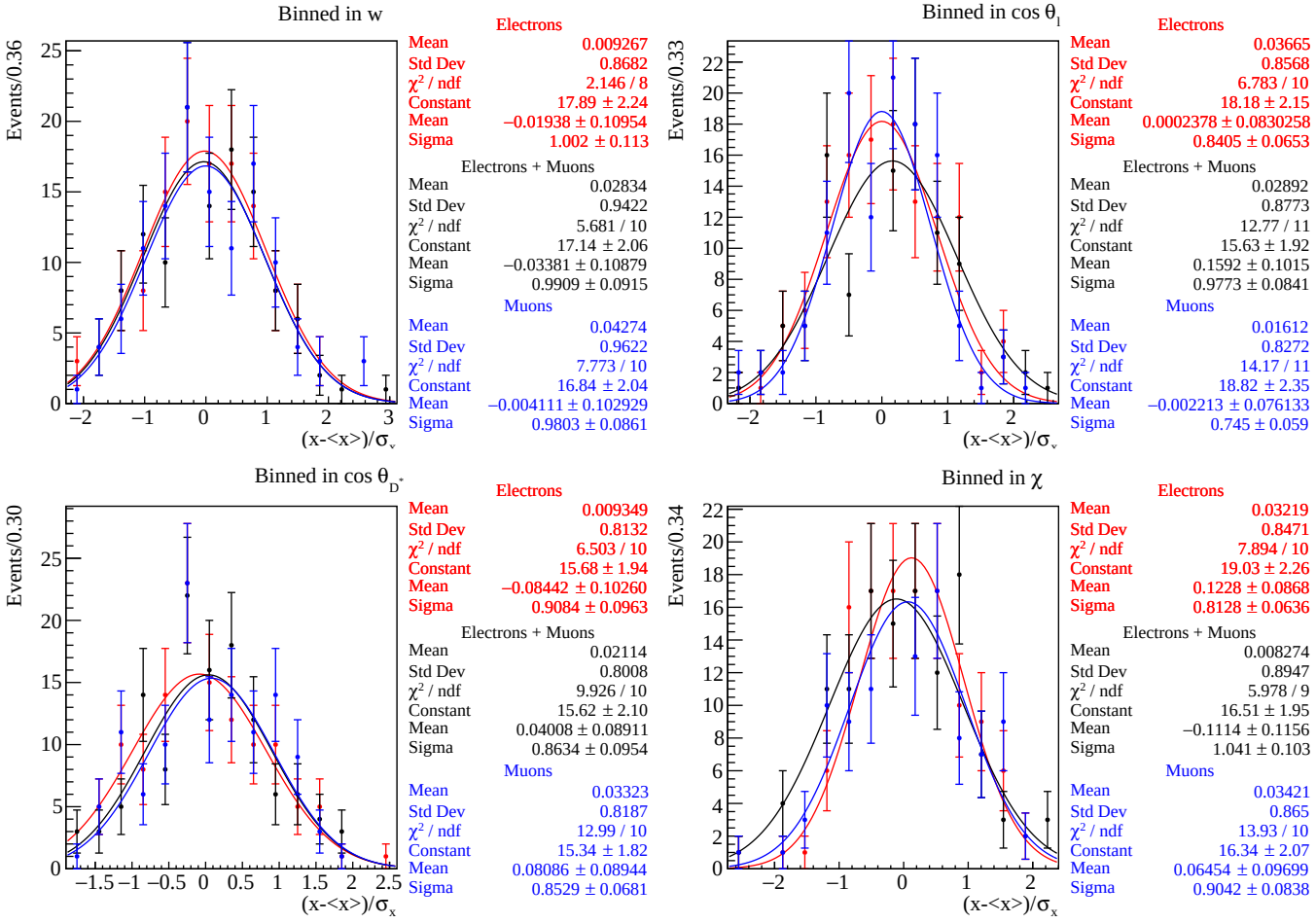


Figure C.13: The pull distributions of closure test 1 with different B types and bins in the kinematic variable considered together in one histogram.

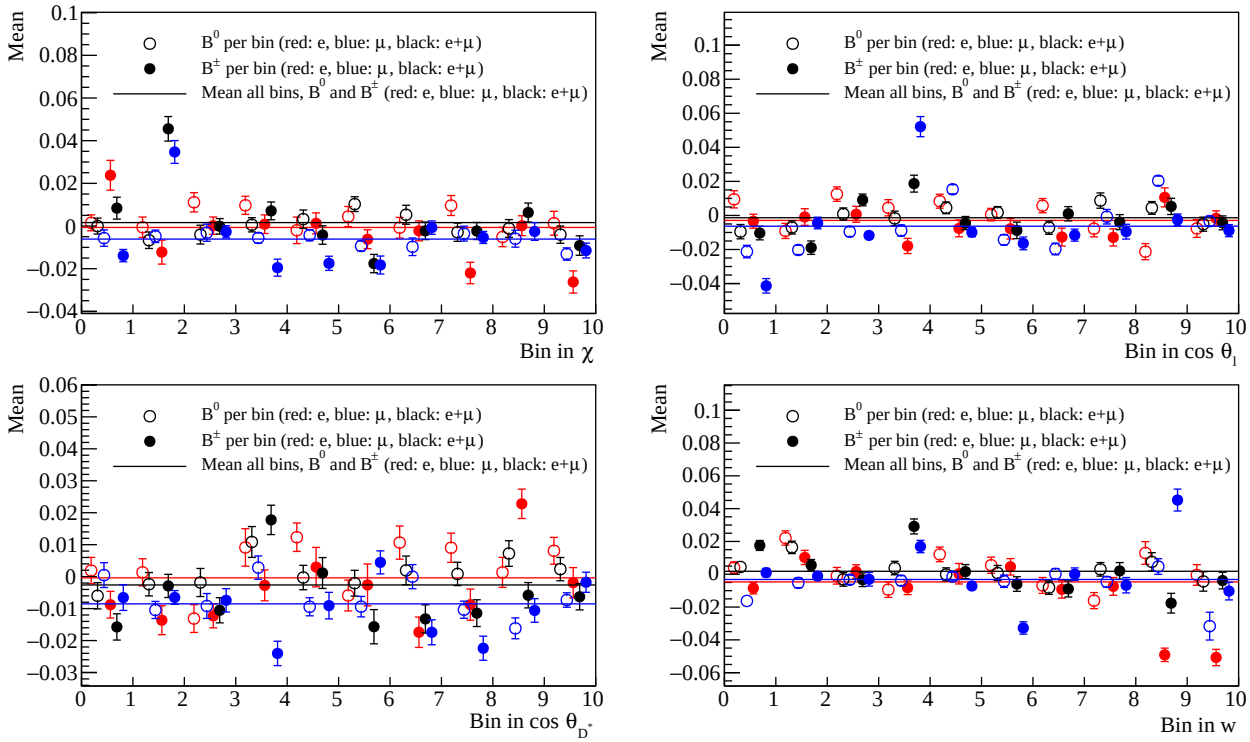


Figure C.14: Overview of the mean of the pull distributions for closure test 2 (data stream 0). Note: The markers for electrons, muons and the combined fit have been slightly shifted to enhance readability. A total of 1000 toys were generated for this closure test.

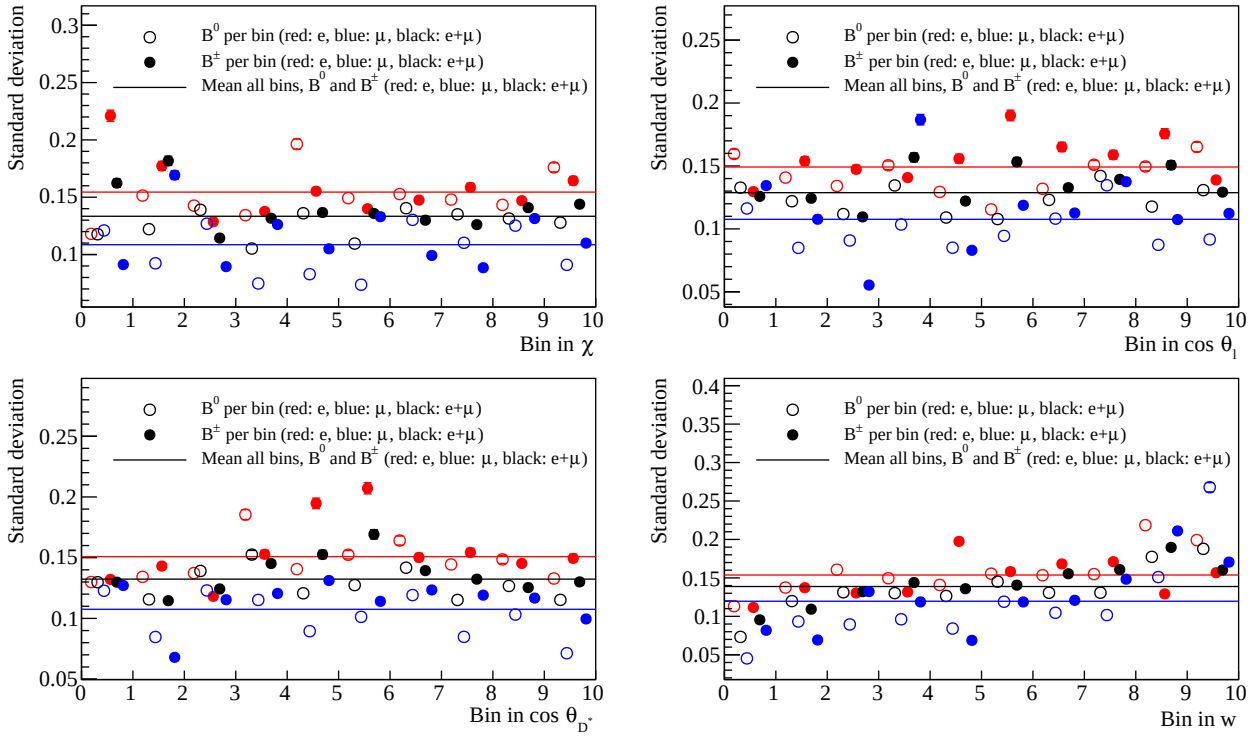


Figure C.15: Overview of the standard deviation of the pull distributions for closure test 2 (data stream 0). Additional notes are given in figure C.14.

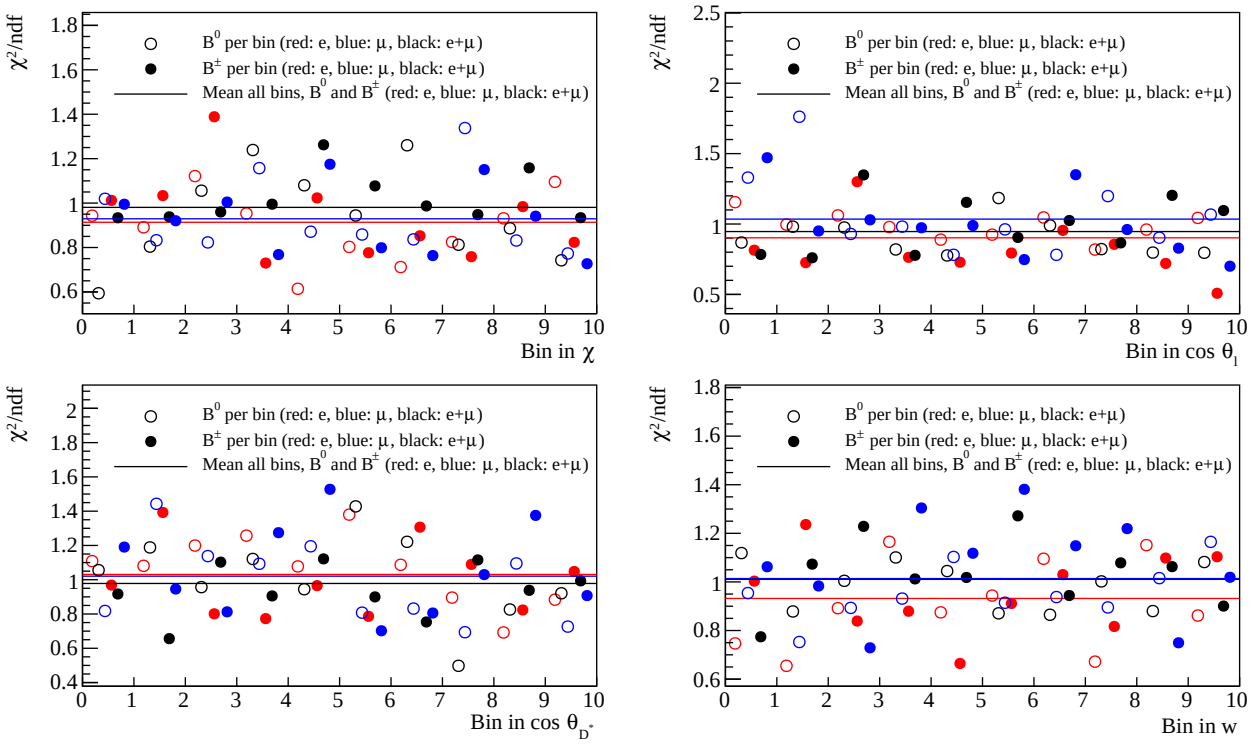


Figure C.16: Overview of the χ^2/n.d.f value of the fit of the pull histograms of closure test 2 (data stream 0) to Gaussians. Additional notes are given in figure C.14.

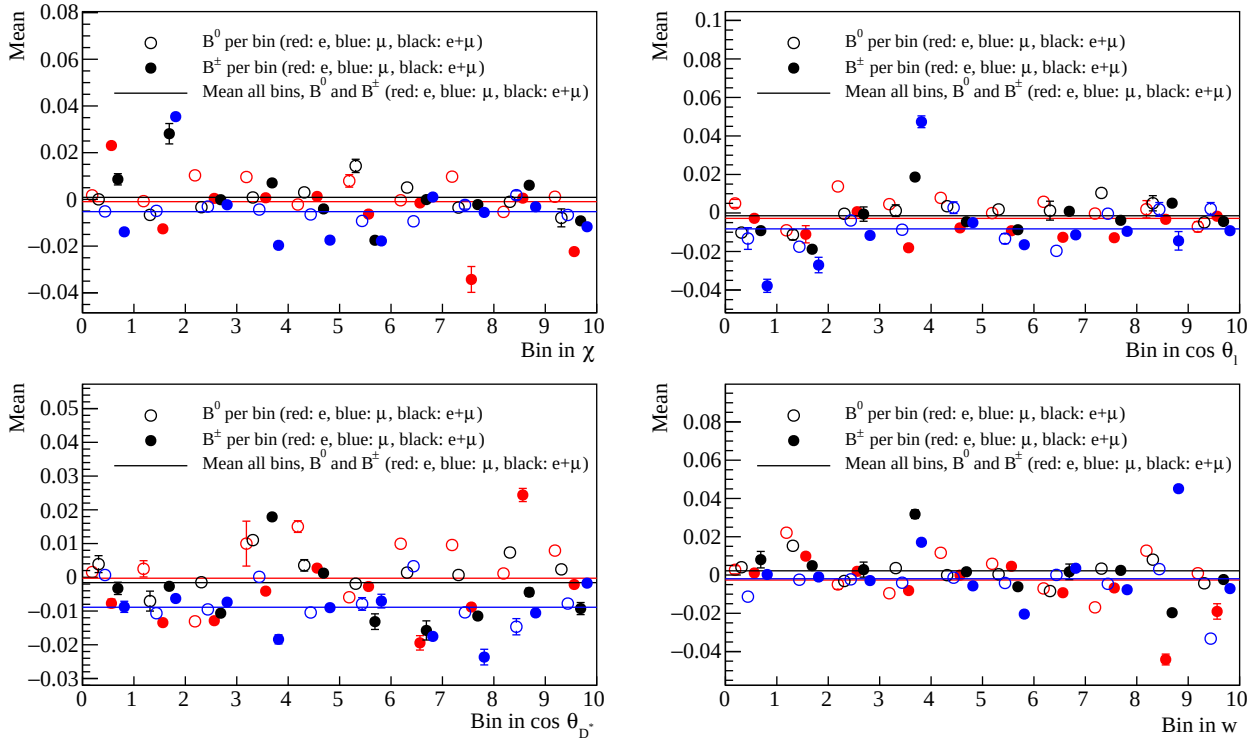


Figure C.17: Overview of the mean of the pull distributions for the toys with varied MC (MC where one of three $B \longrightarrow D^*$ form factor parameters is varied by $\pm 1\sigma$, leading to six results per bin, B type and lepton type). Note: The markers for electrons muons and the combined fit results have been slightly shifted to enhance readability. Also note that the errors shown do not take the error of the single Gaussian fit into account, but only the deviations for the six different results for the MC variations. A total of 1000 toys were generated for each MC stream for this closure test.

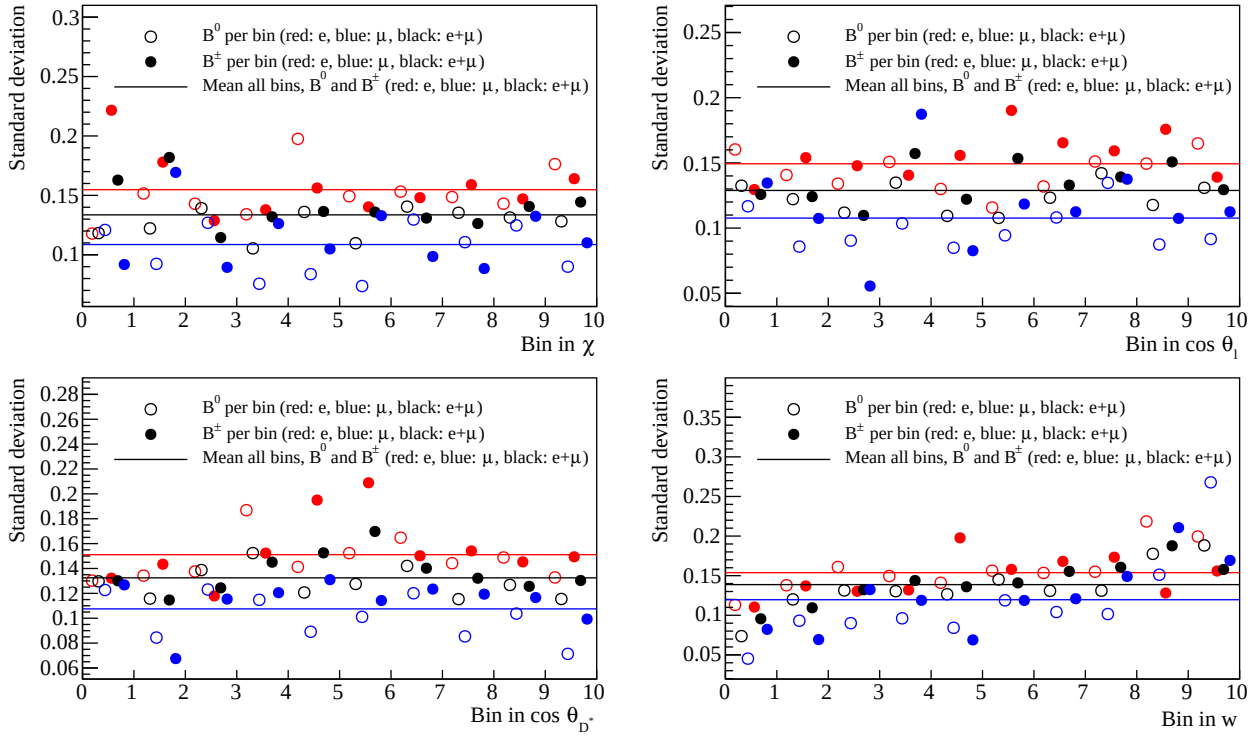


Figure C.18: Overview of the standard deviation of the pull distributions for closure test 2 (data stream 0) with varied MC. Additional notes are given in figure C.17.

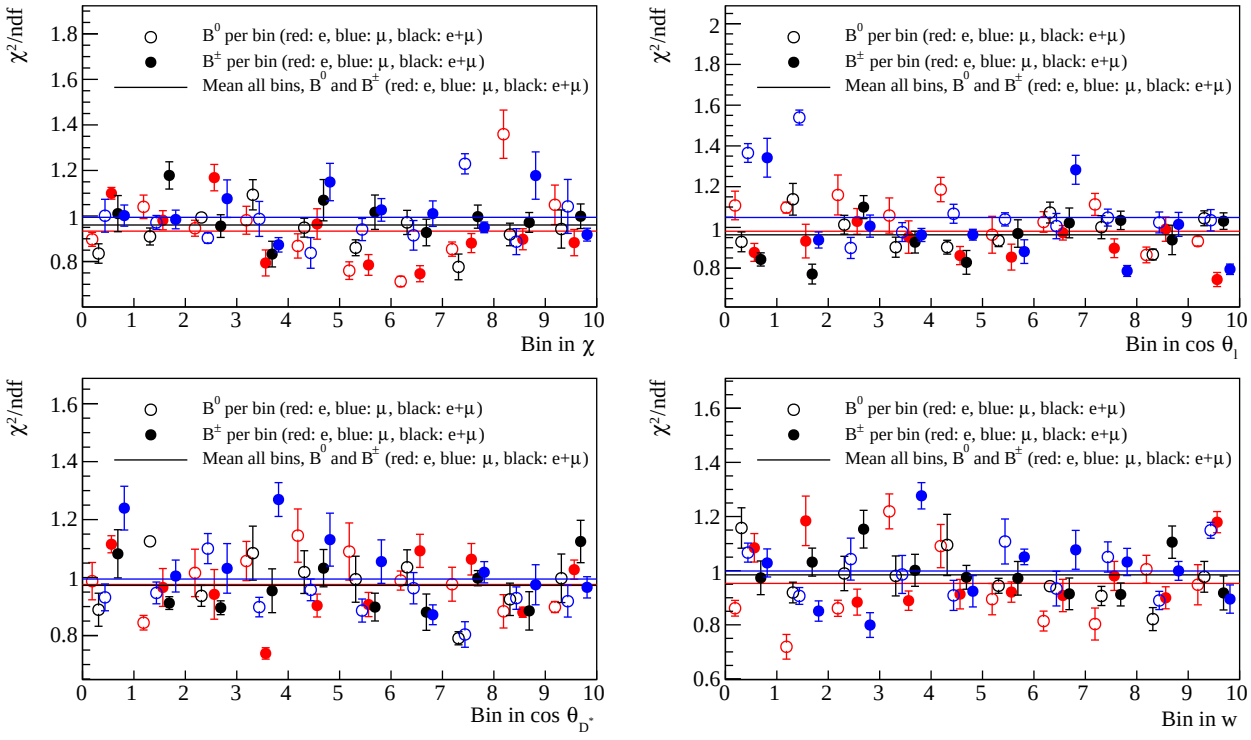


Figure C.19: Overview of the χ^2/n.d.f value of the fit of the pull histograms for closure test 2 (data stream 0) to Gaussians with varied MC. Additional notes are given in figure C.17.

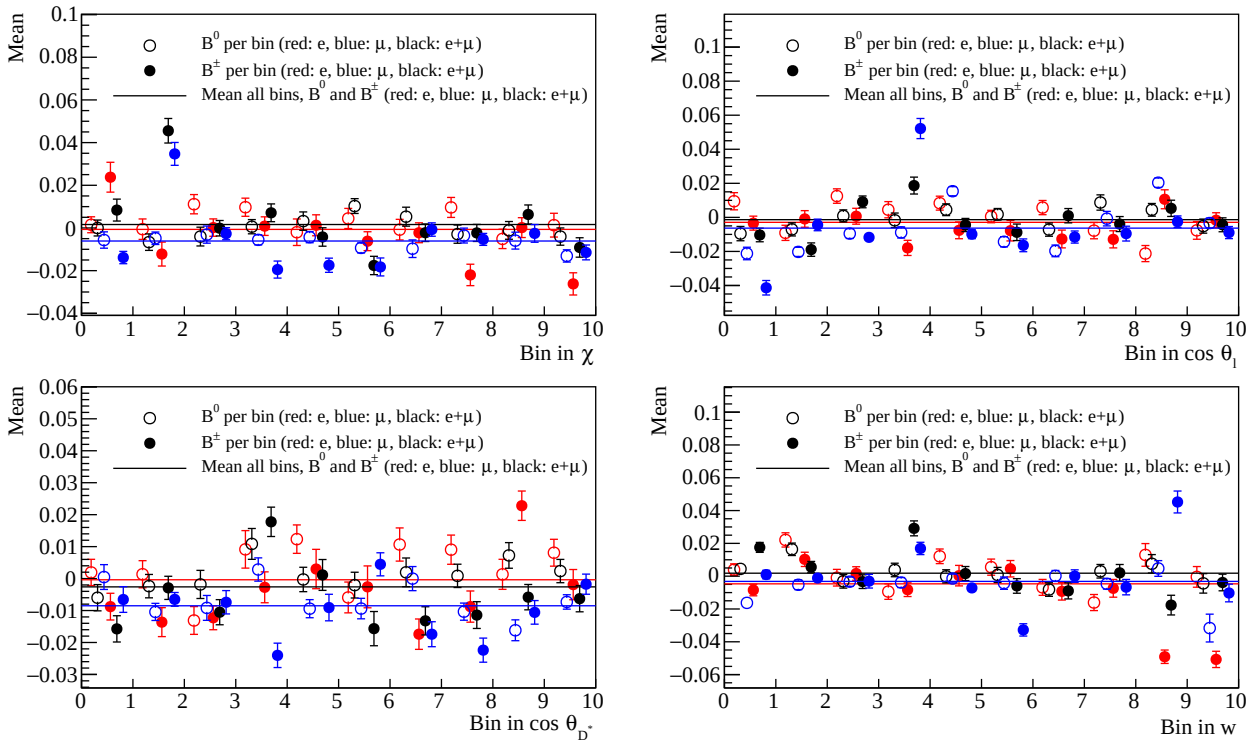


Figure C.20: Overview of the mean of the pull distributions for closure test 2 with real data (used to determine the PDF shape systematic error). Note: The markers for electrons, muons and the combined fit have been slightly shifted to enhance readability. A total of 1000 toys were generated for this.

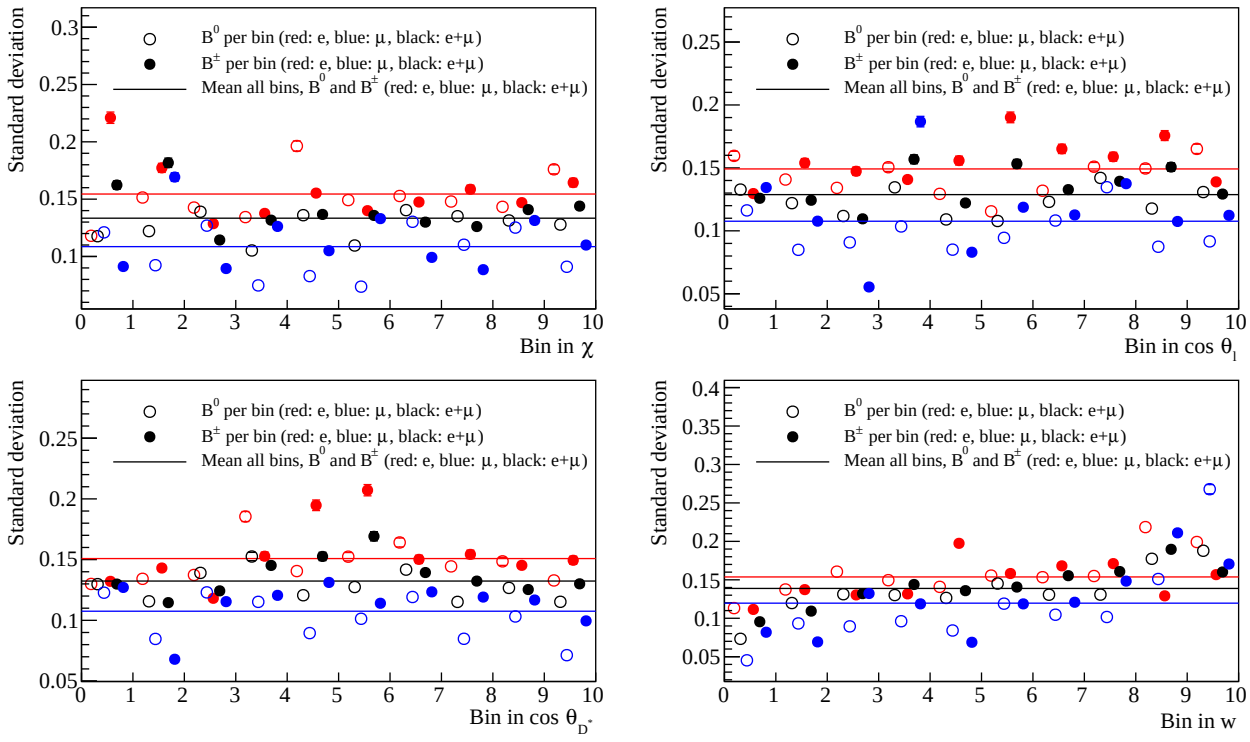


Figure C.21: Overview of the standard deviation of the pull distributions for closure test 2 with real data (used to determine the PDF shape systematic error). Additional notes are given in figure C.20.

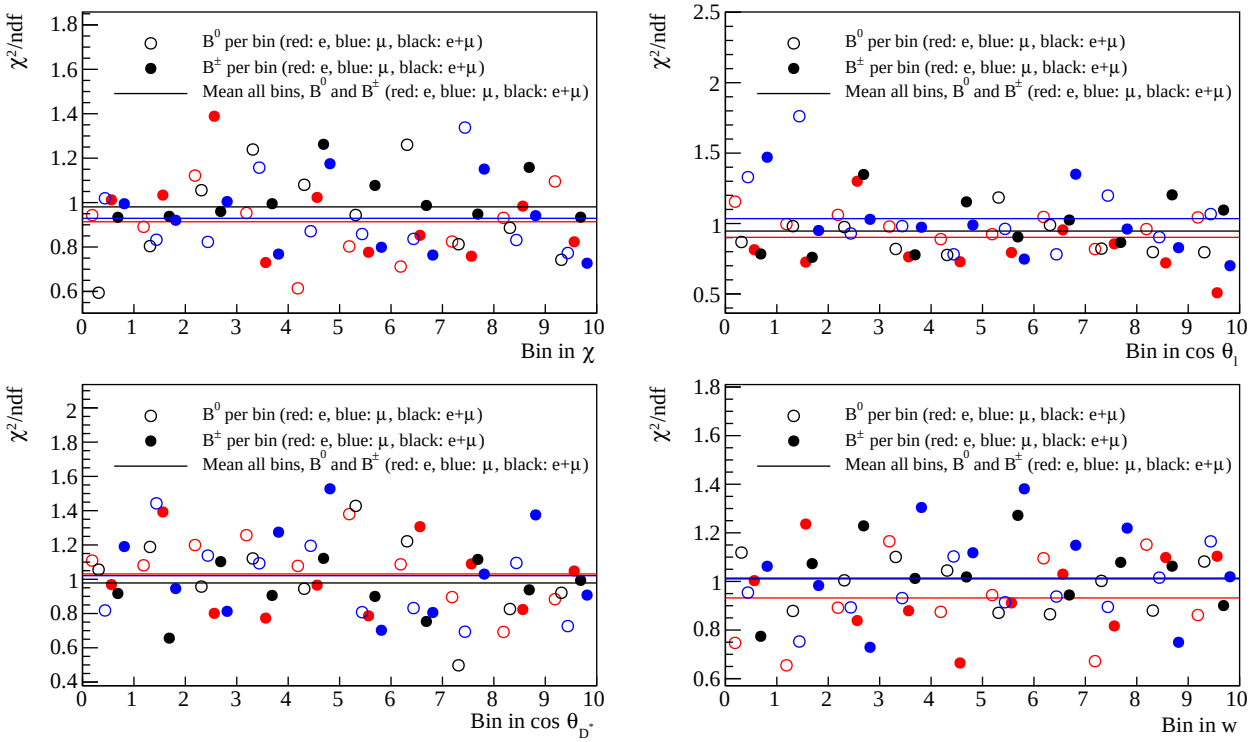


Figure C.22: Overview of the χ^2/n.d.f value of the fit of the pull histograms closure test 2 with real data (used to determine the PDF shape systematic error) to Gaussians. Additional notes are given in figure C.20.

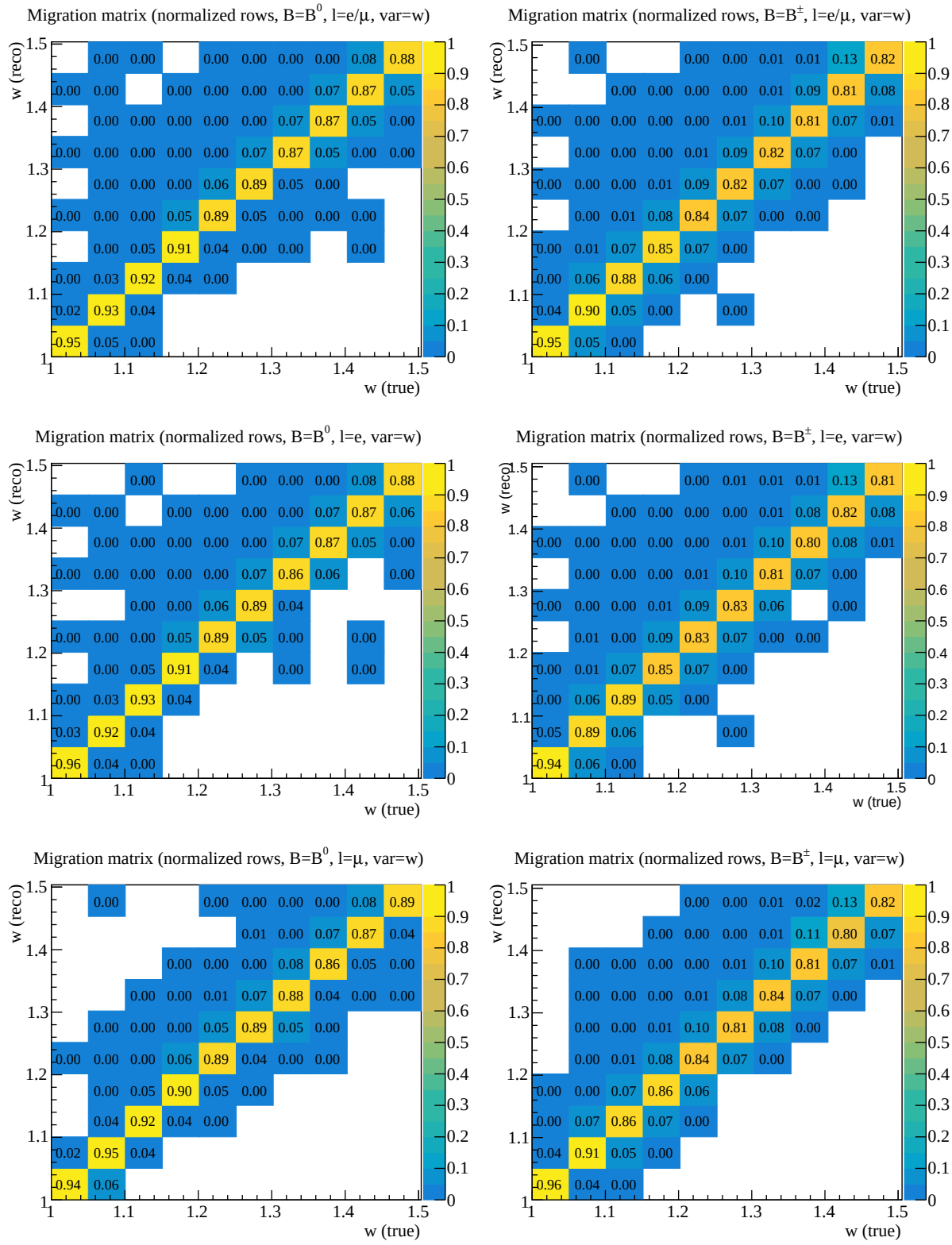


Figure C.23: Migration matrices for the binning in w.

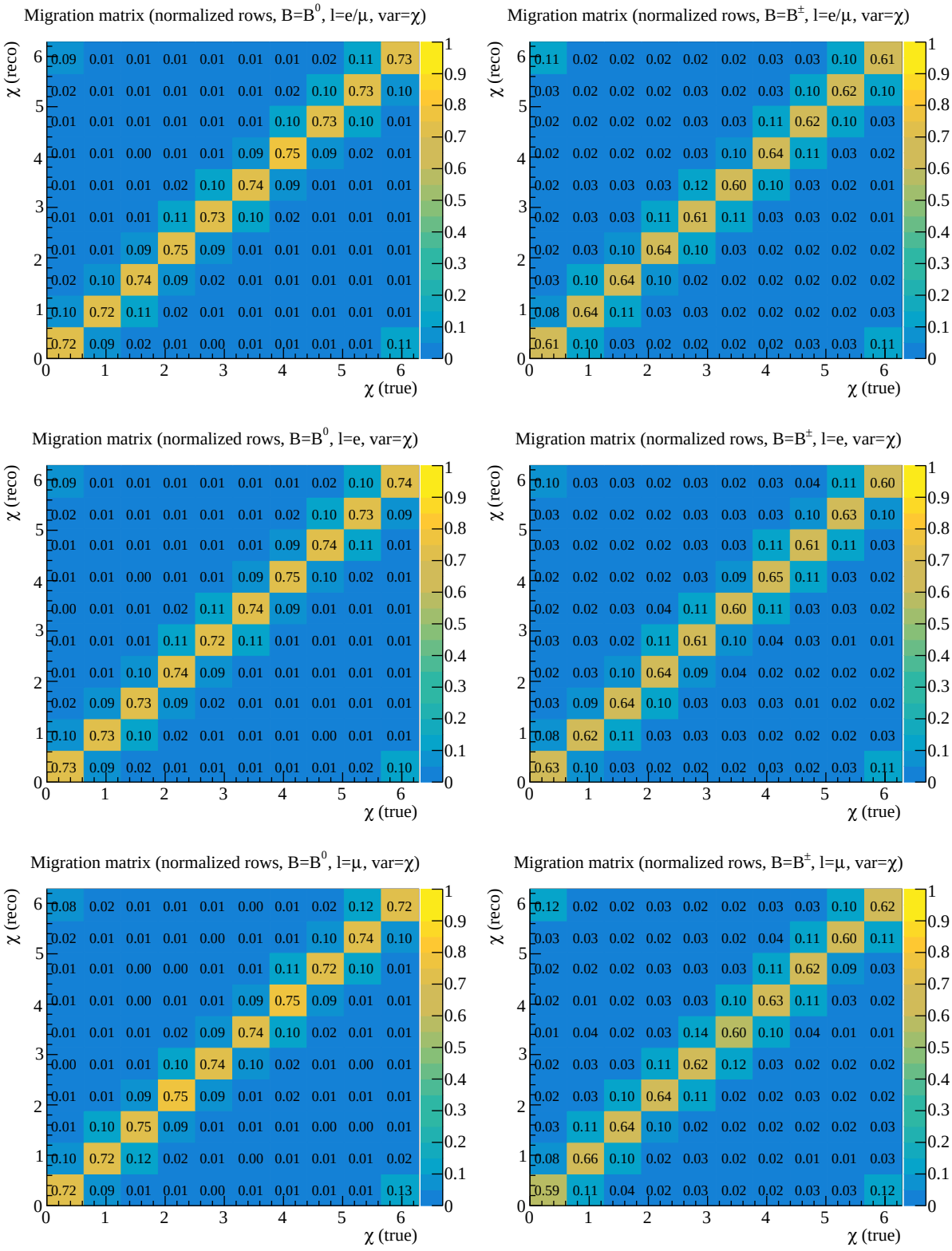


Figure C.24: Migration matrices for the binning in χ.

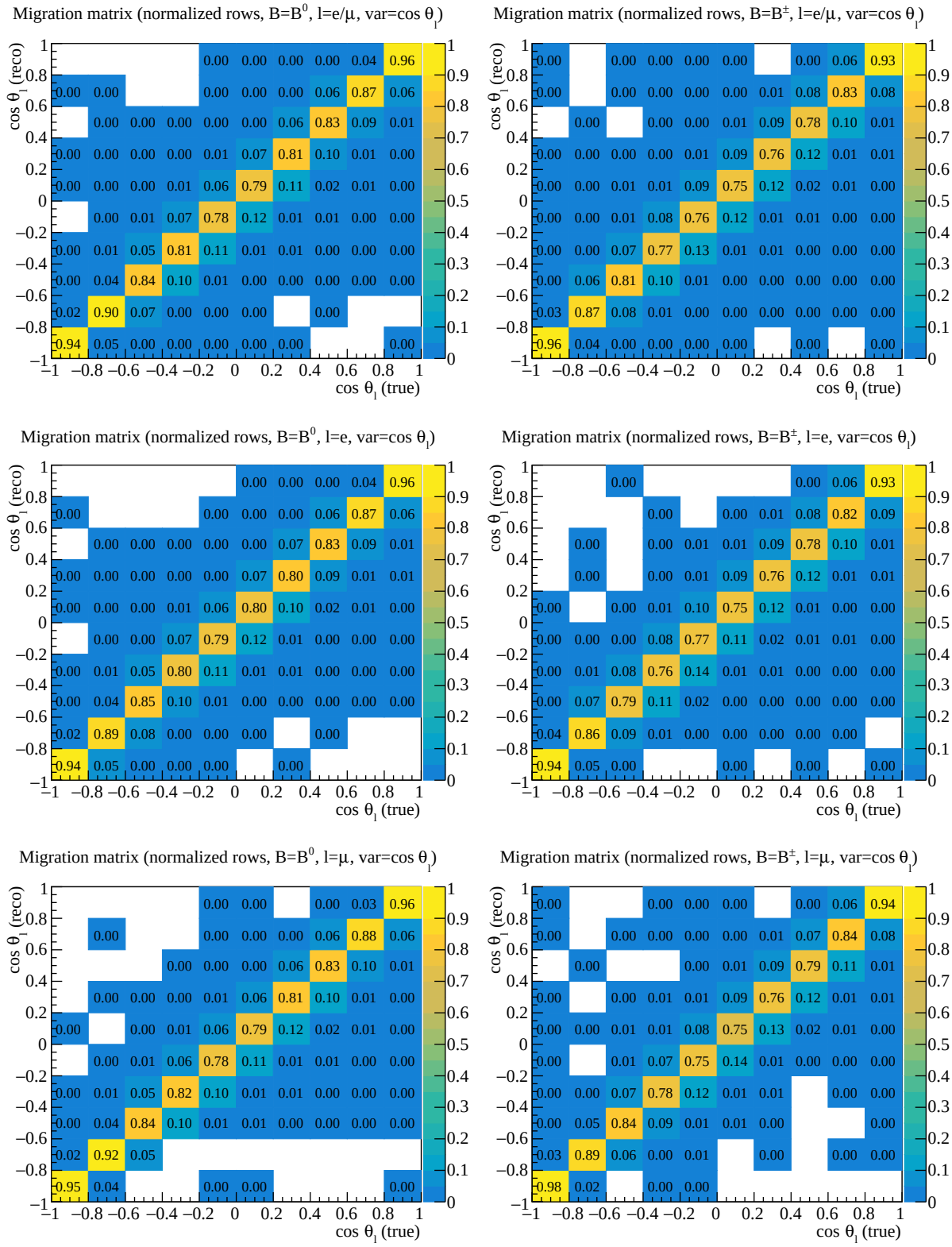


Figure C.25: Migration matrices for the binning in $\cos\theta_\ell$.

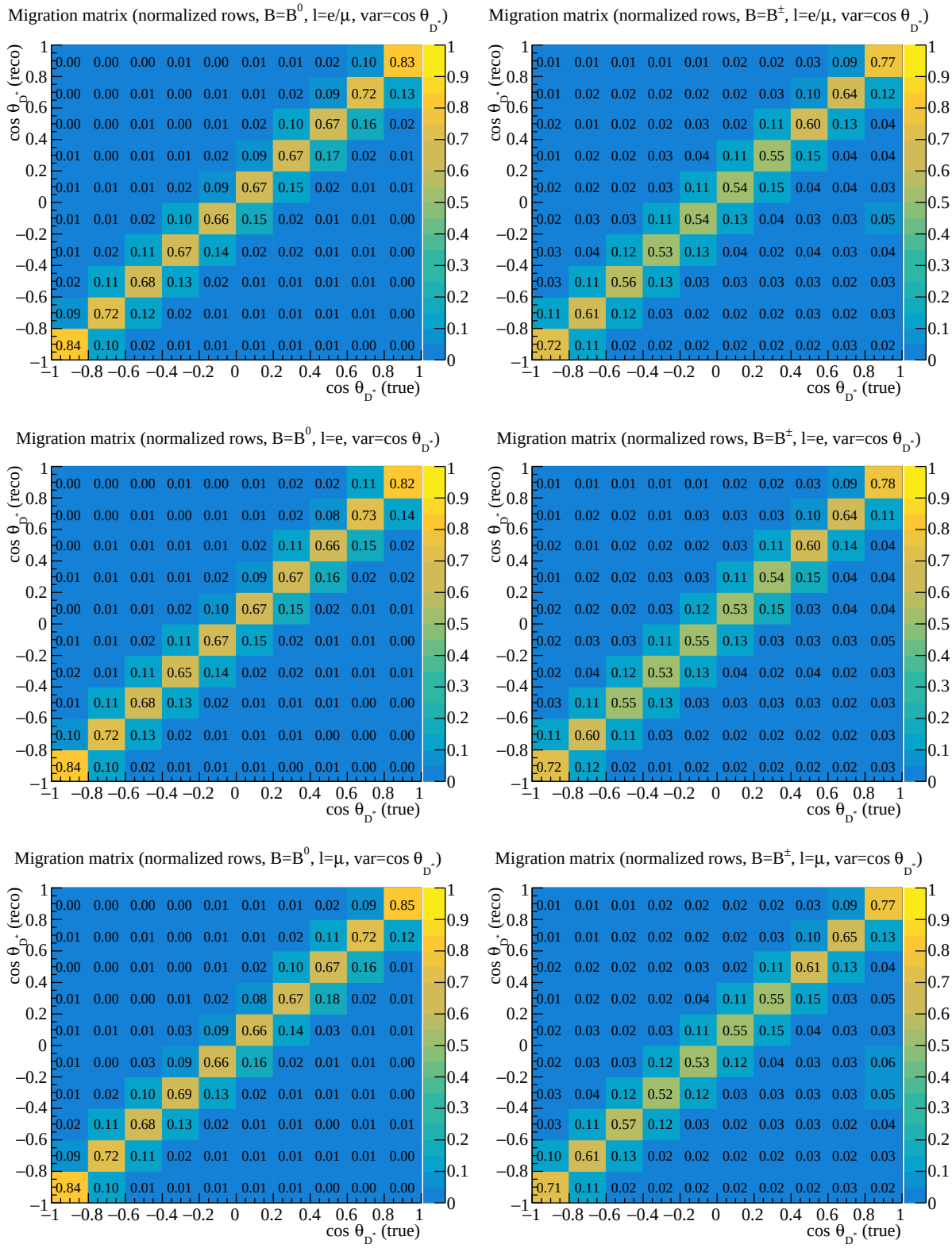


Figure C.26: Migration matrices for the binning in $\cos\theta_{D^*}$.

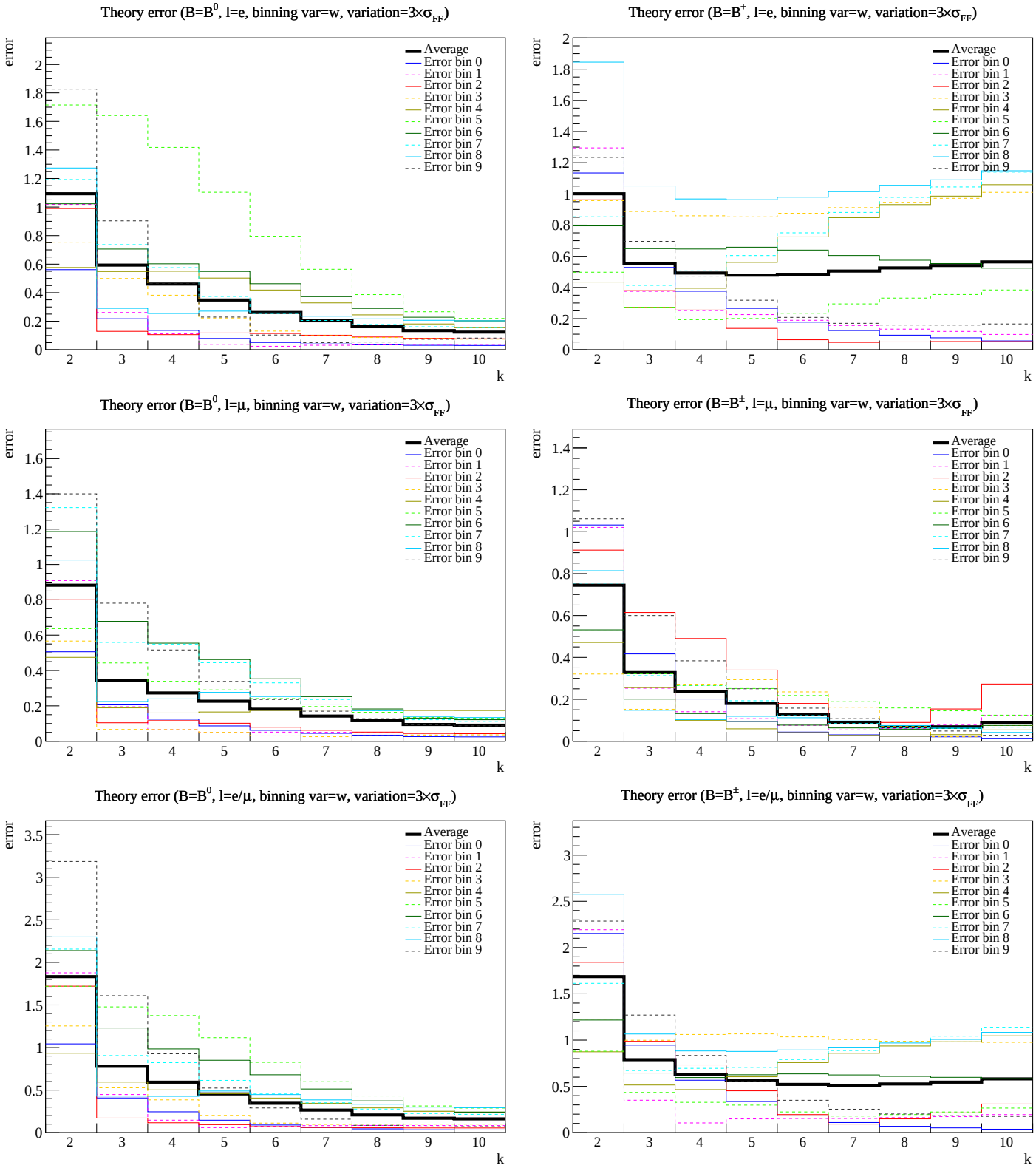


Figure C.27: Model uncertainties for the binning in w as described in section 3.8.2.

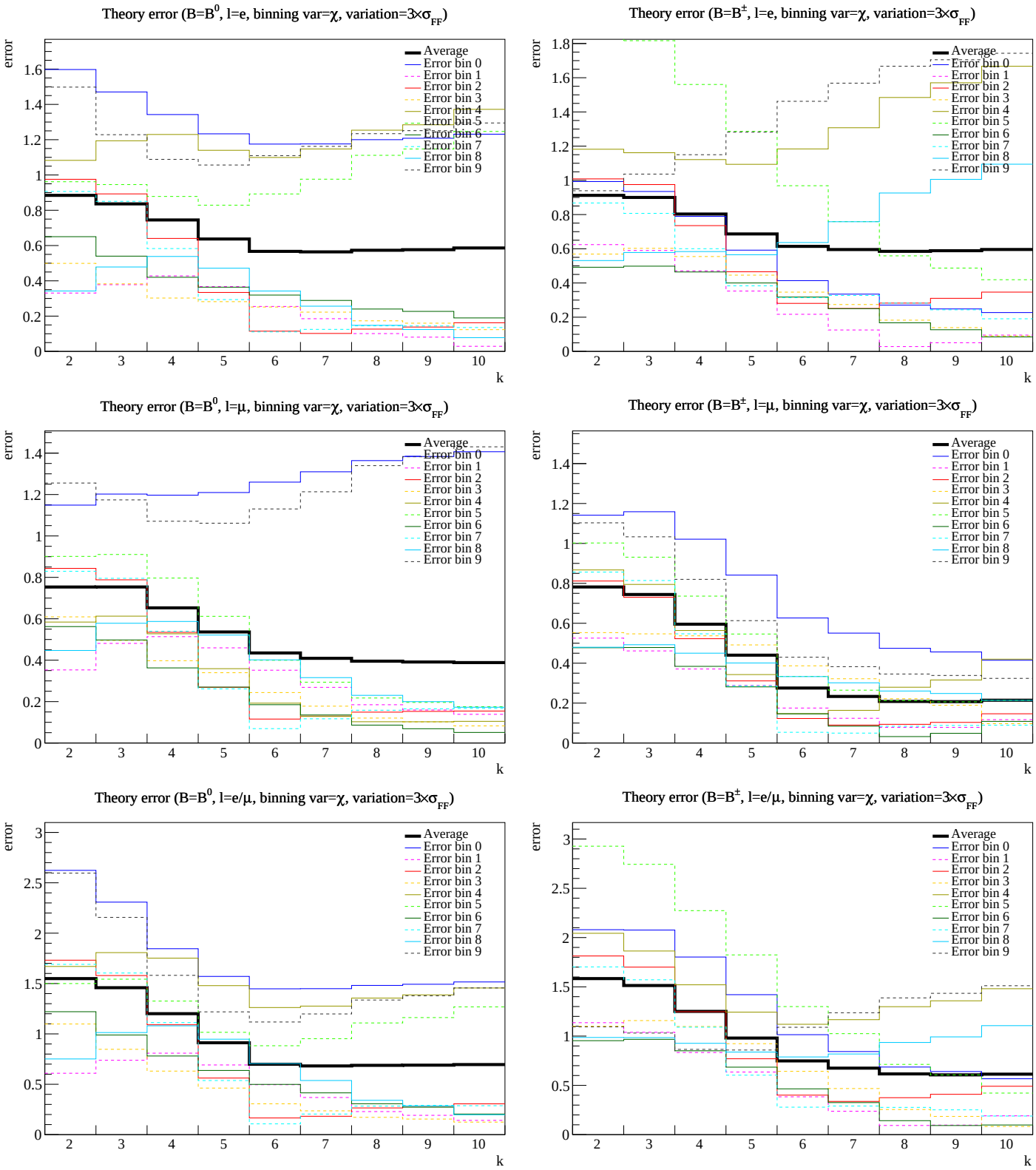


Figure C.28: Model uncertainties for the binning in χ.

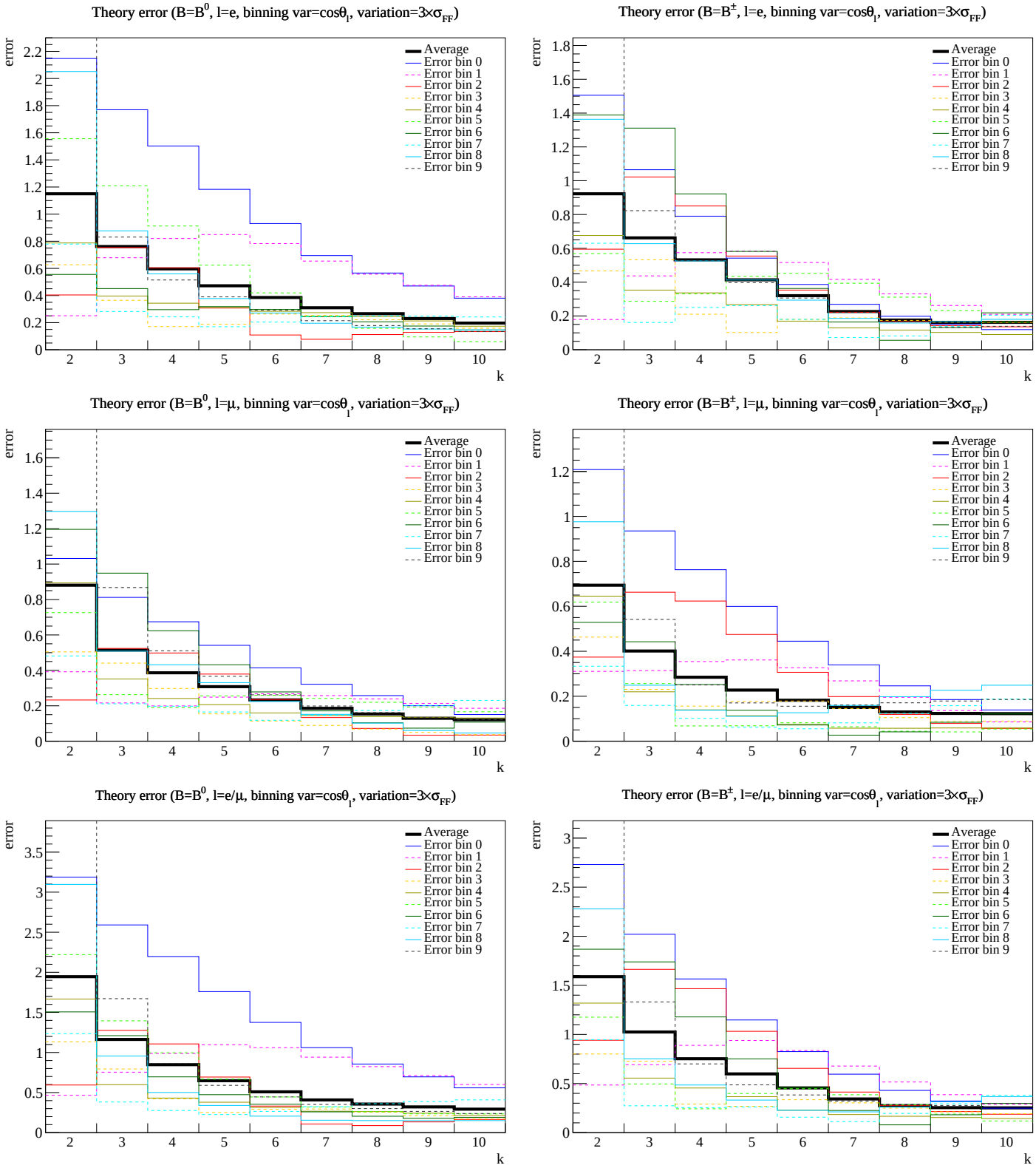


Figure C.29: Model uncertainties for the binning in $\cos\theta_\ell$.

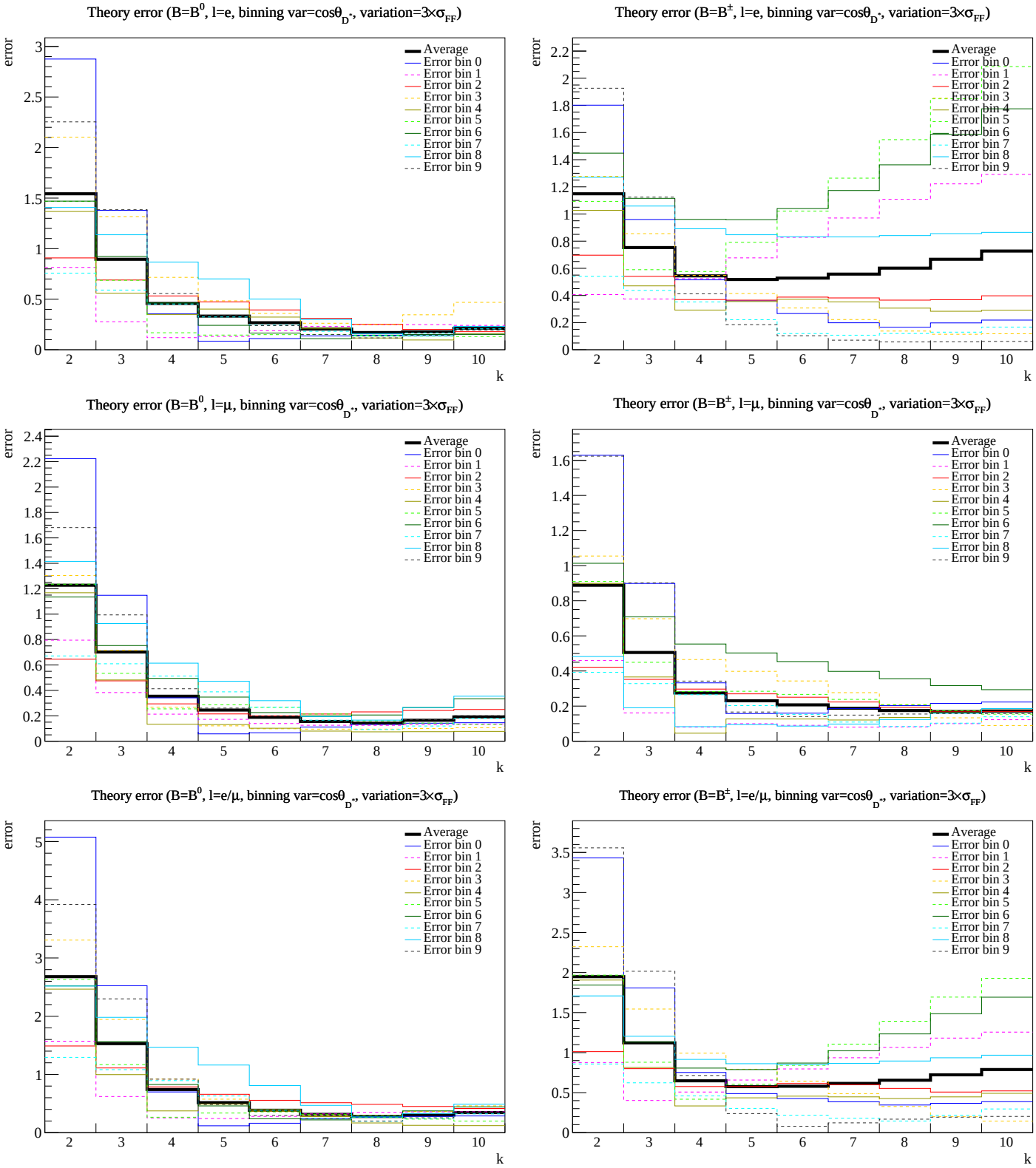


Figure C.30: Model uncertainties for the binning in $\cos\theta_{D^*}$.

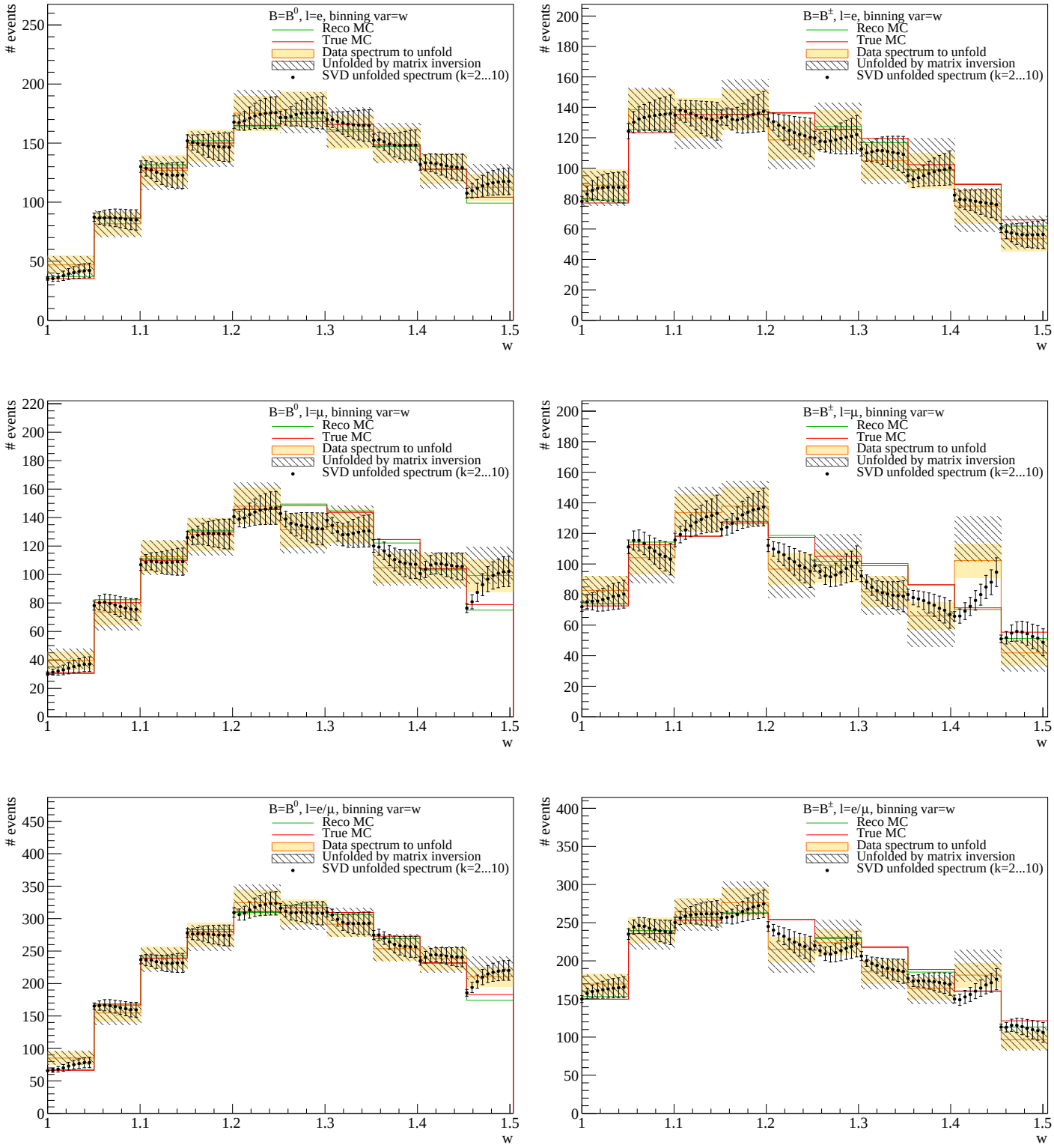


Figure C.31: Unfolding of real data binned in w using matrix inversion and SVD unfolding for multiple values of the regularization parameter k as discussed in section 3.8.5. Results unfolded using matrix inversion are shown as hatched rectangles, results obtained with the SVD unfolding algorithm (for several different regularization parameters k) are shown as black data points with error bars. The input data to be unfolded is shown as a golden histogram with golden shaded rectangles for its errors. For comparison, reco MC data (green histogram) and truth MC data (red histogram) are shown as well.

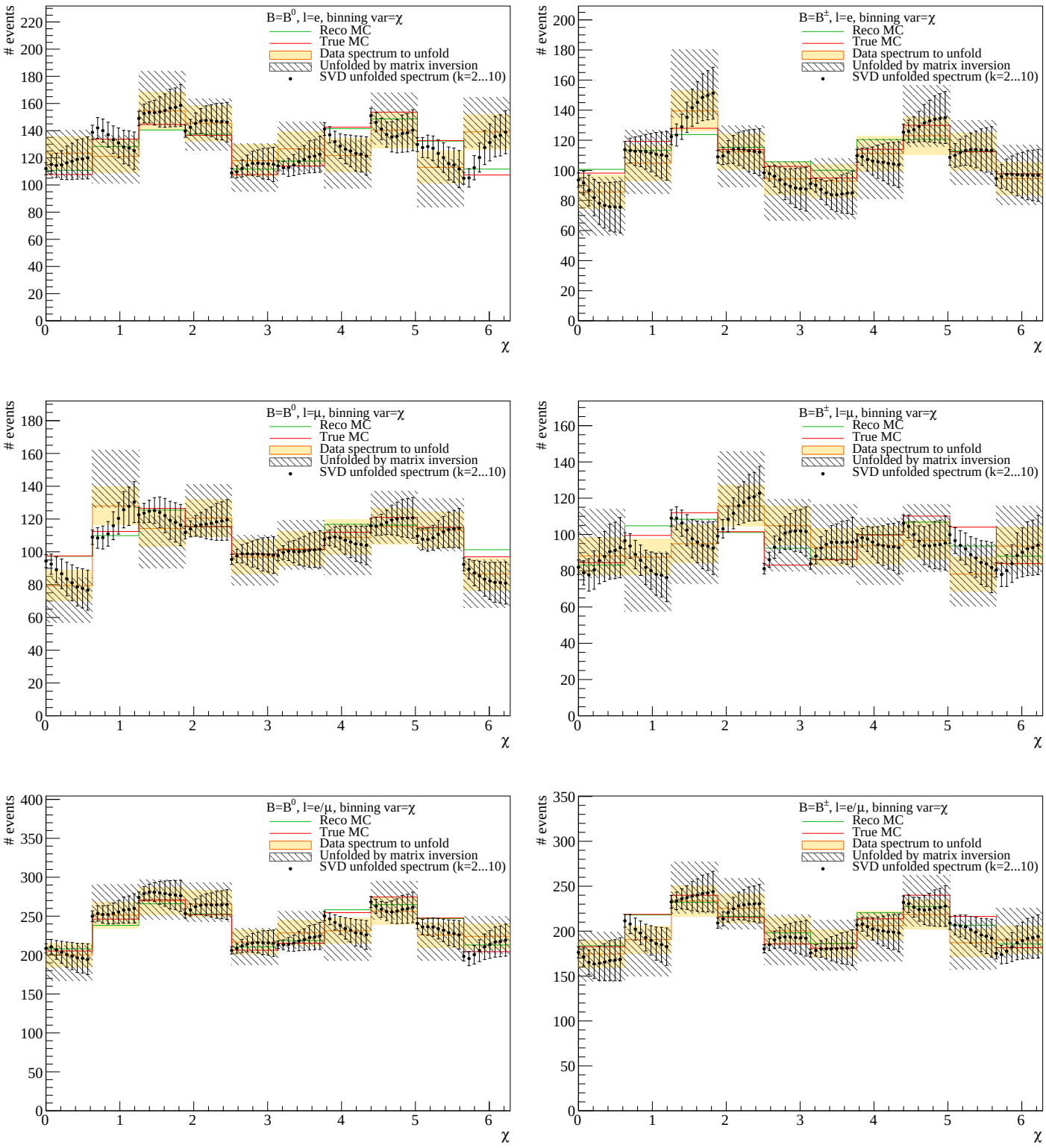


Figure C.32: Unfolding of real data binned in χ using matrix inversion and SVD unfolding for multiple values of the regularization parameter k. See figure C.31 for additional notes.

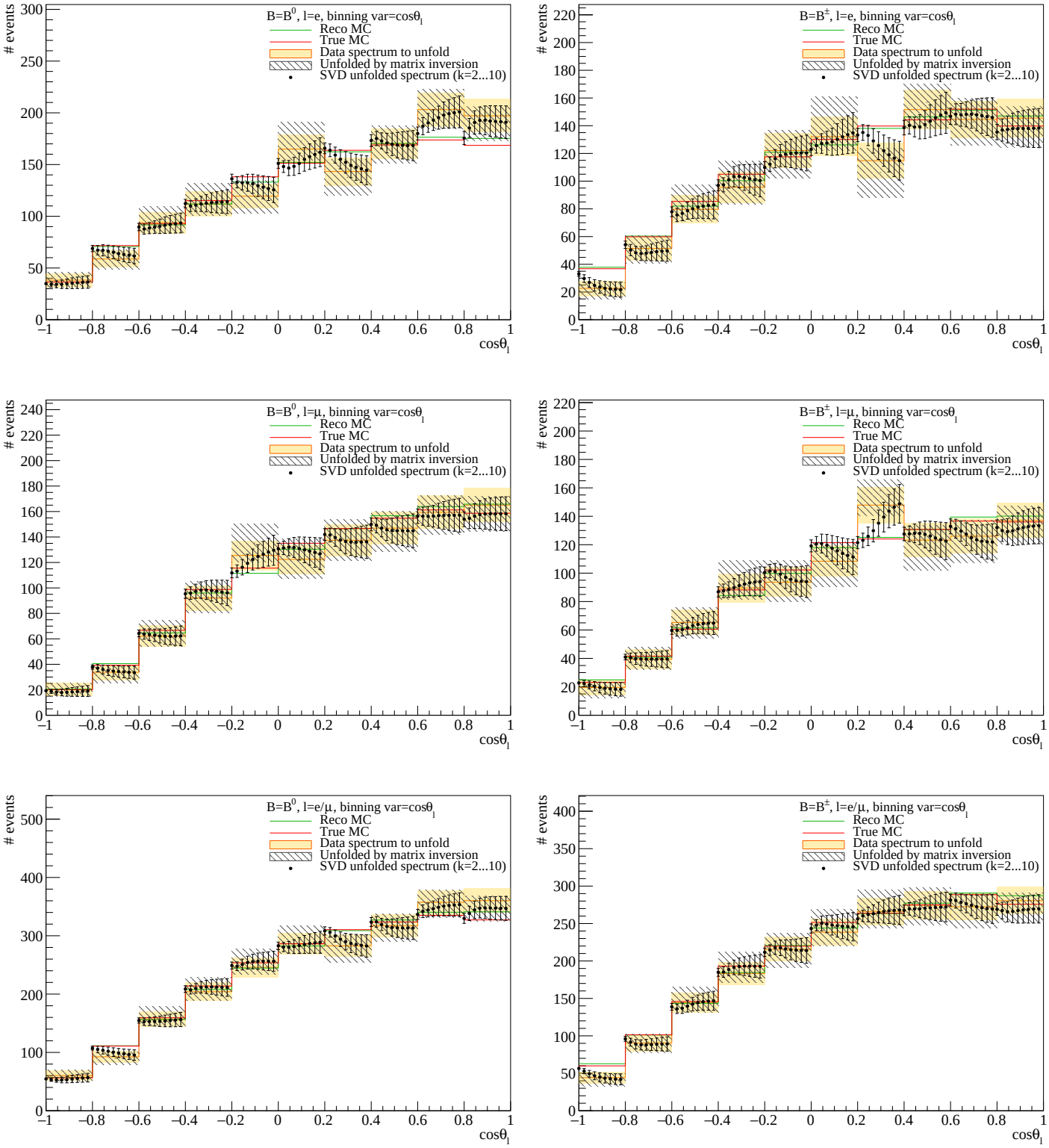


Figure C.33: Unfolding of real data binned in $\cos\theta_\ell$ using matrix inversion and SVD unfolding for multiple values of the regularization parameter k. See figure C.31 for additional notes.

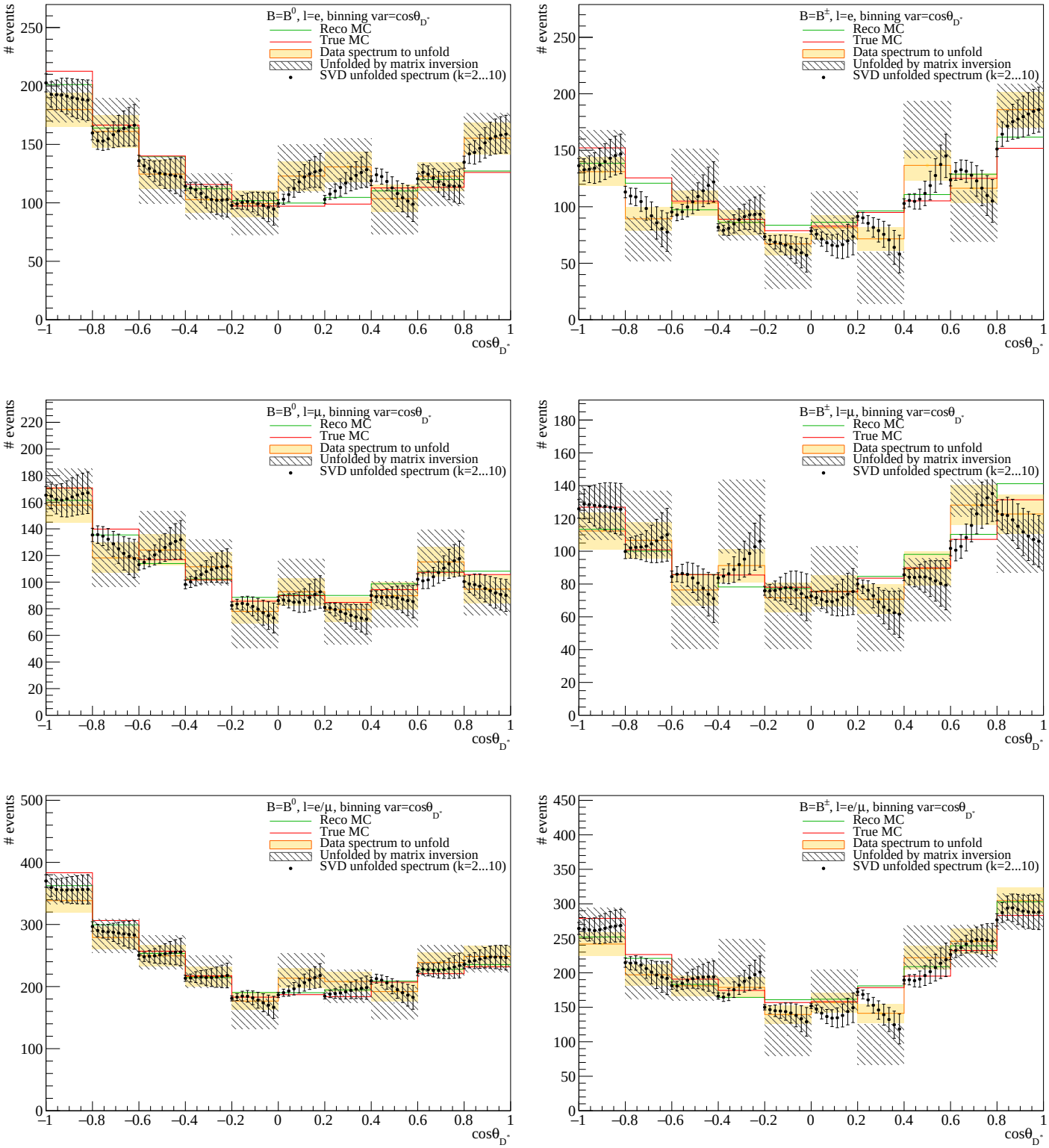


Figure C.34: Unfolding of real data binned in $\cos\theta_{D^*}$ using matrix inversion and SVD unfolding for multiple values of the regularization parameter k. See figure C.31 for additional notes.

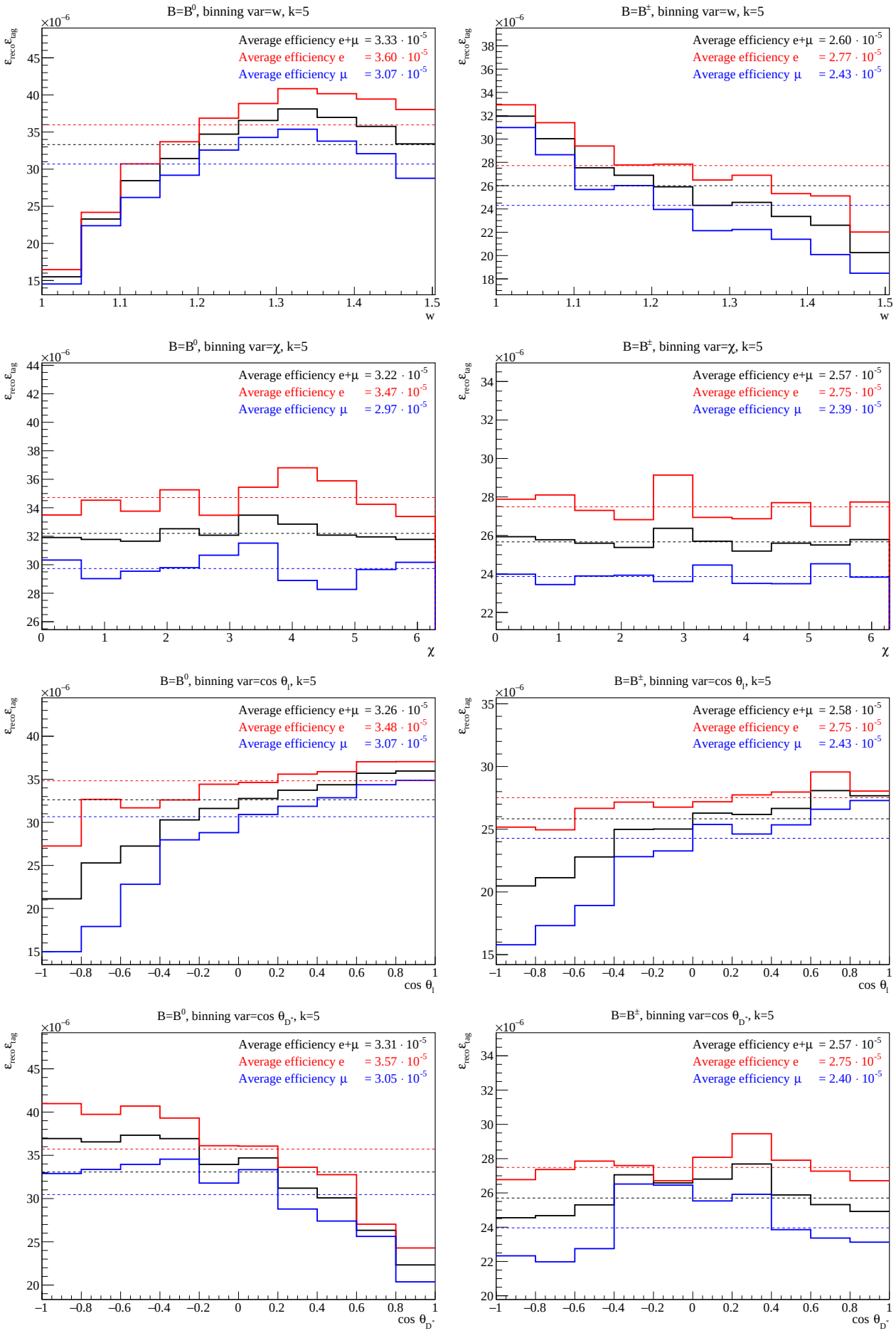


Figure C.35: Efficiency $\epsilon_{\text{reco}}\epsilon_{\text{tag}}$ per bin for electrons (red), muons (blue) and the combined fit (black). The dashed lines correspond to the average efficiency (calculated with event yields as weights).

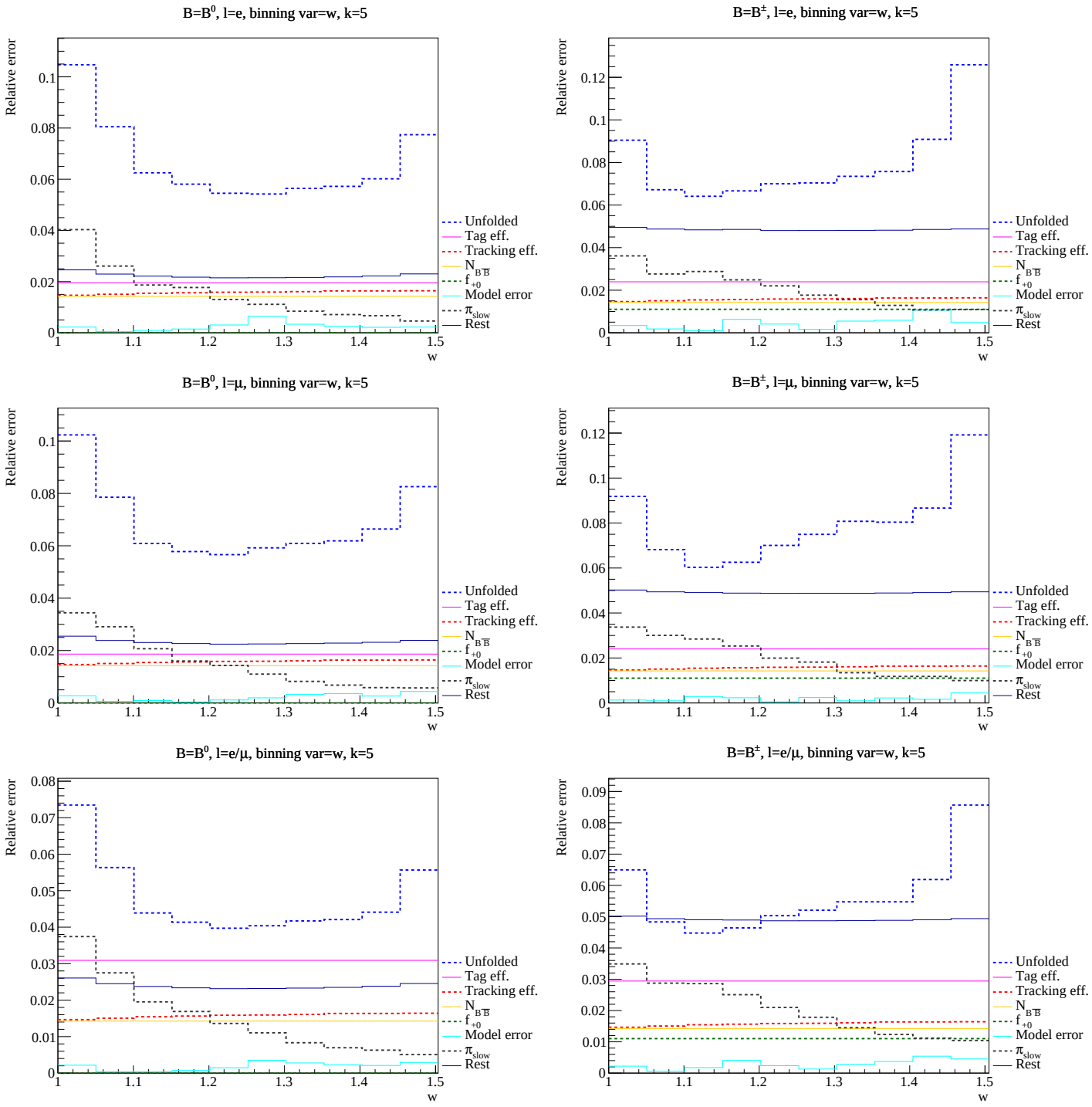


Figure C.36: Comparison of relative errors per bin in w. Note that the errors will be added squared, thus the relative impact of large errors is (even) bigger than this figure suggests.

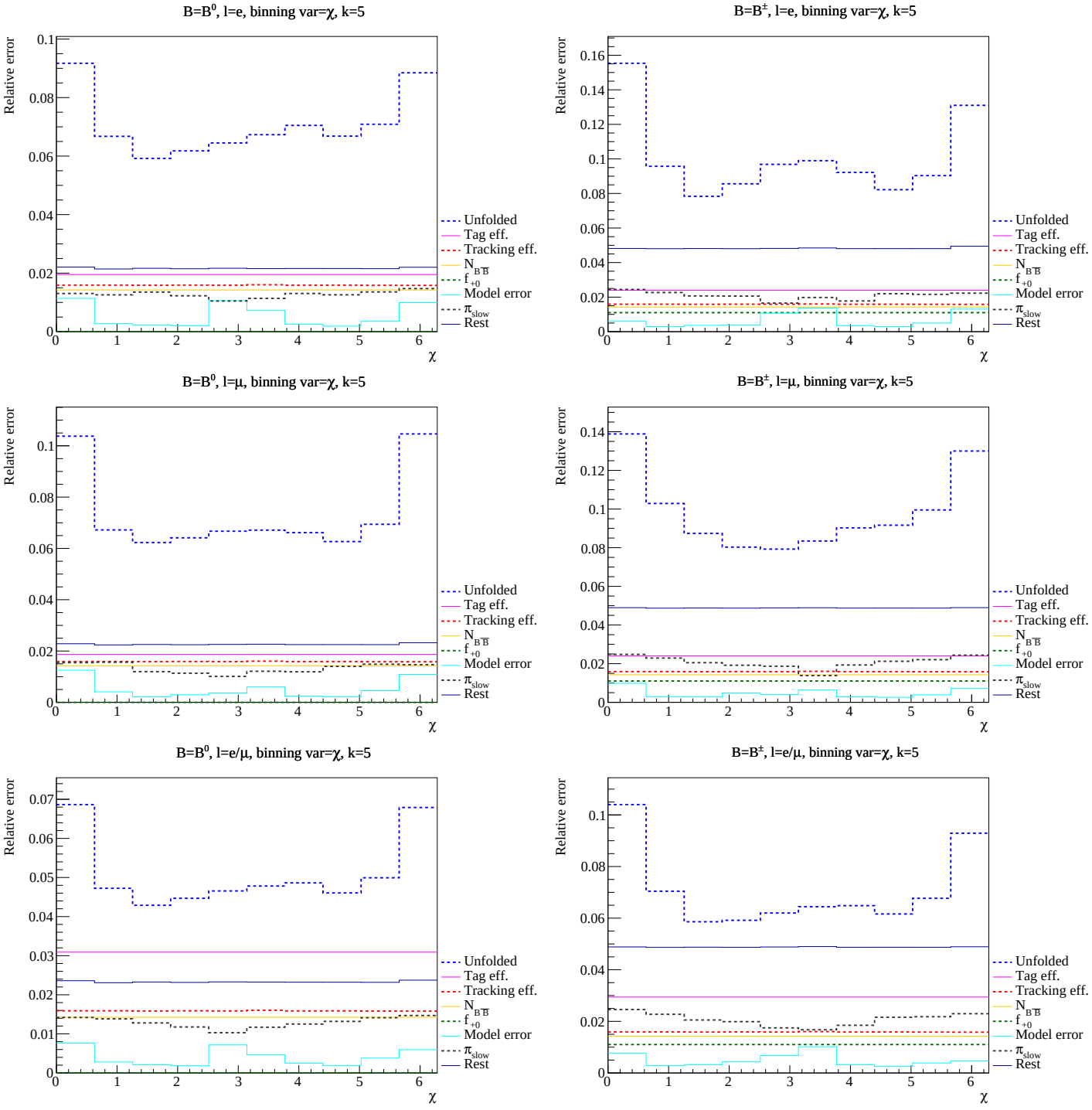


Figure C.37: Comparison of relative errors per bin in χ. Additional notes are given in figure C.36.

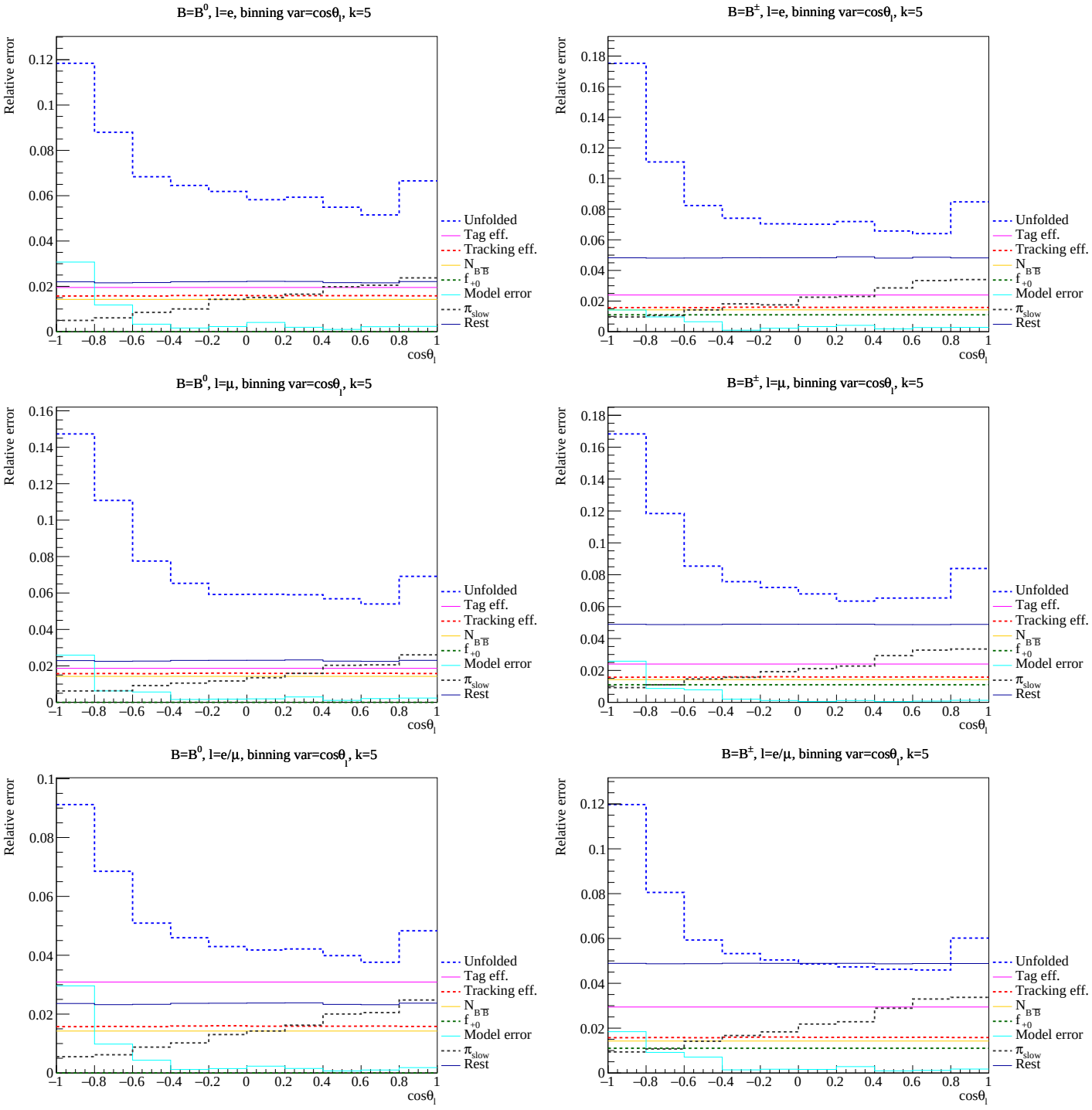


Figure C.38: Comparison of relative errors per bin in $\cos\theta_\ell$. Additional notes are given in figure C.36.

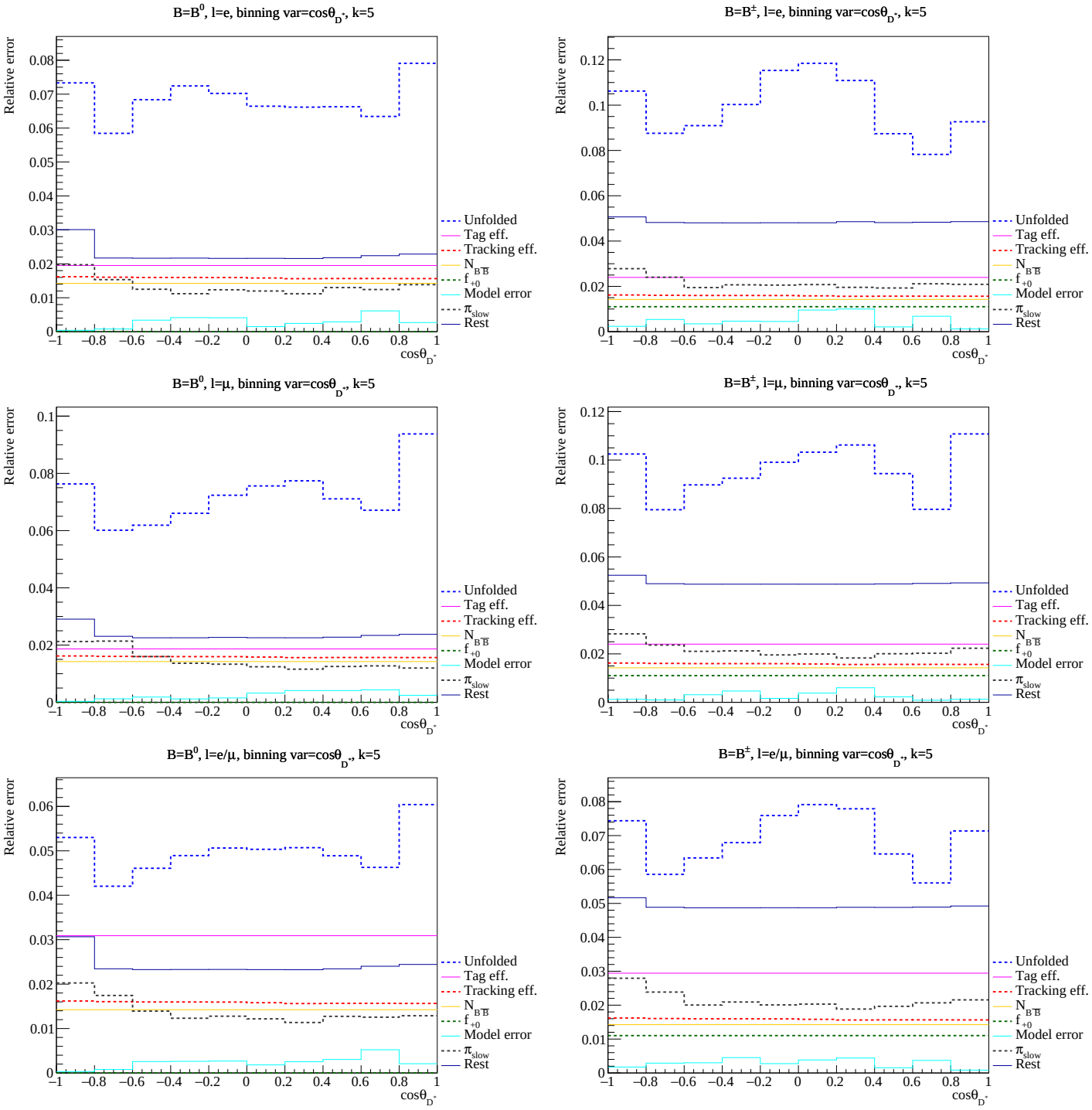


Figure C.39: Comparison of relative errors per bin in $\cos\theta_{D^*}$. Additional notes are given in figure C.36.

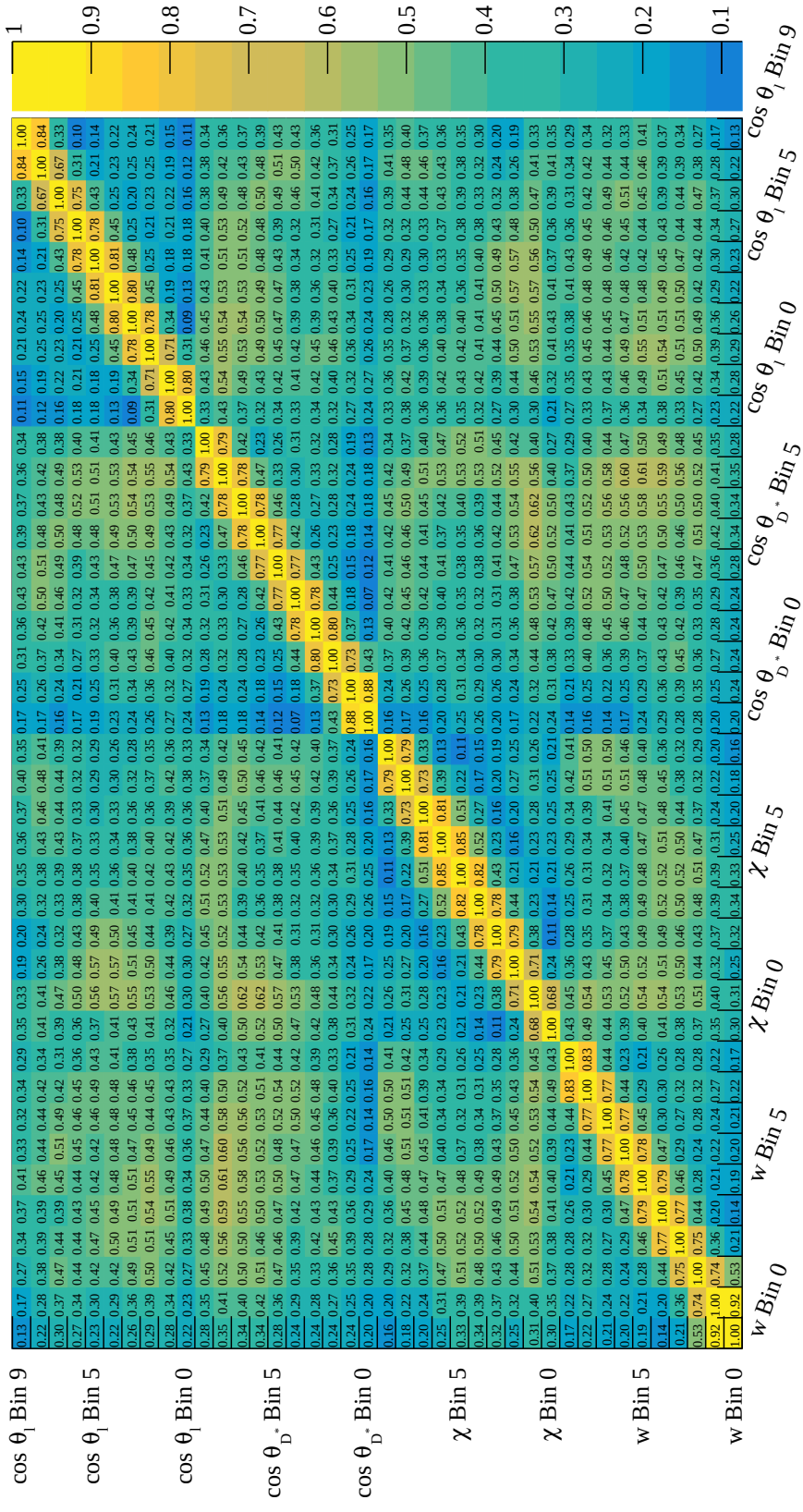


Figure C.40: Correlation matrix for $B = B^0$, $\ell = e$.

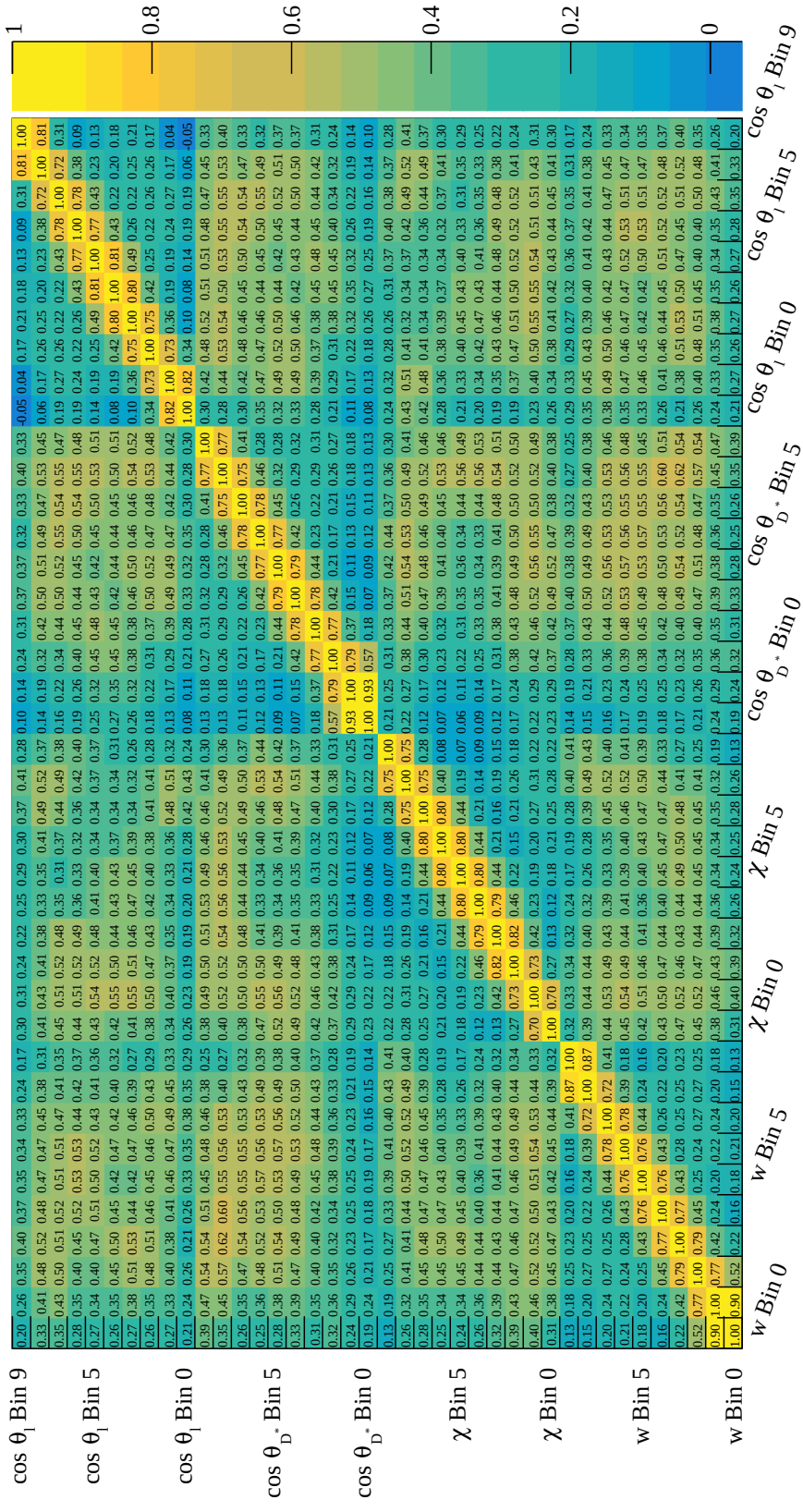


Figure C.41: Correlation matrix for $B = B^0$, $\ell = \mu$.

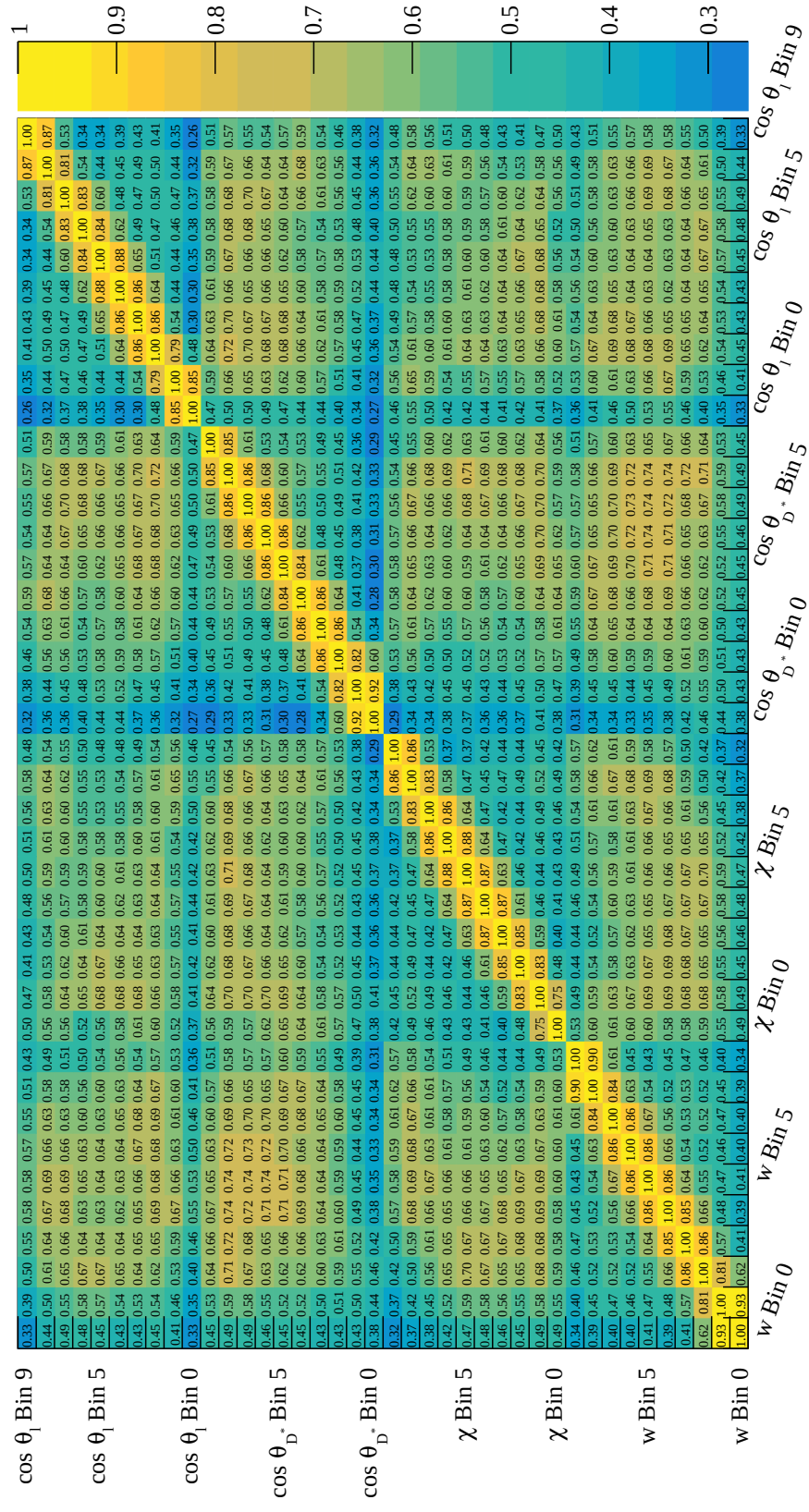


Figure C.42: Correlation matrix for $B = B^0$, $\ell = e + \mu$.

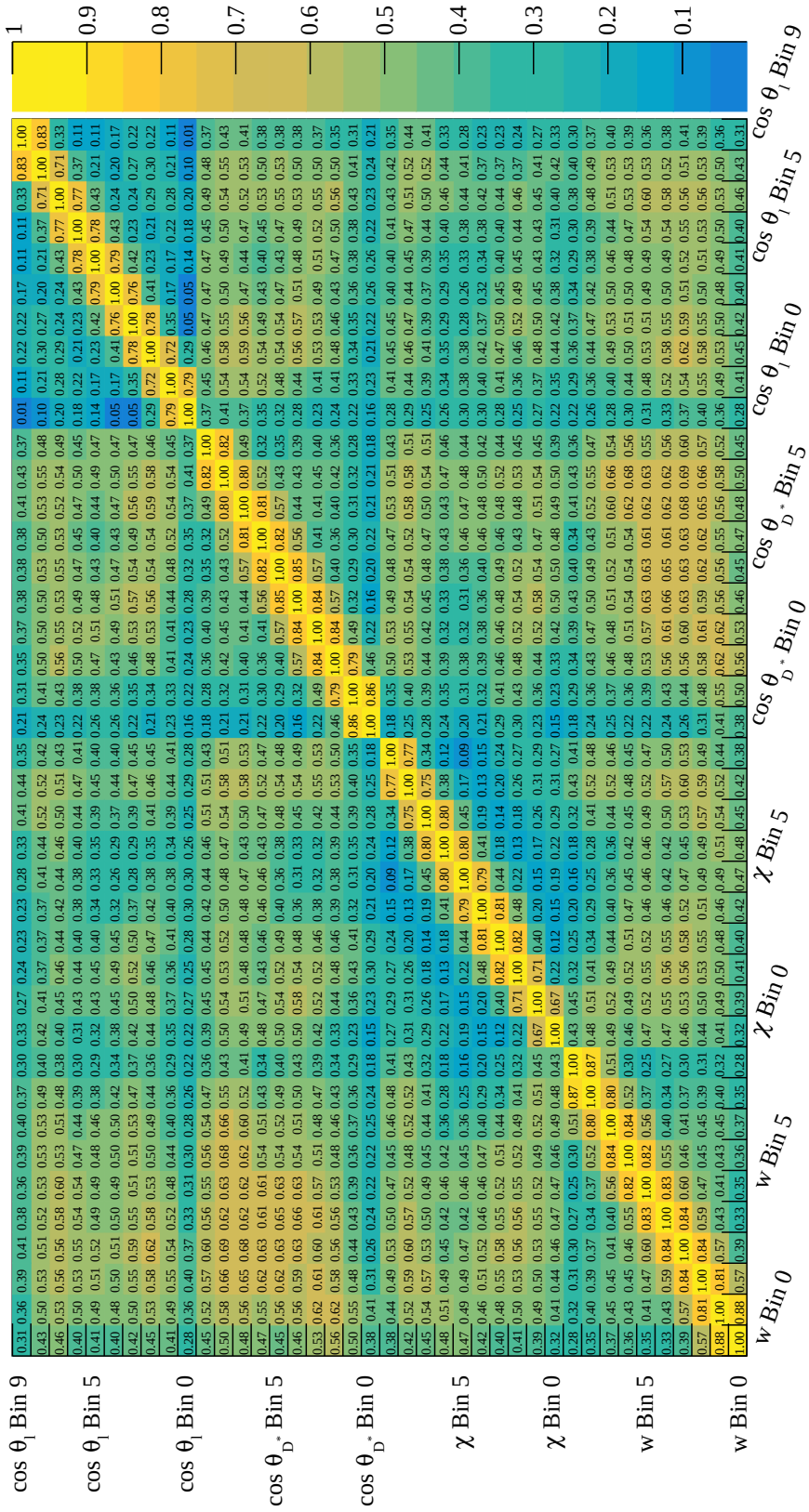


Figure C.43: Correlation matrix for $B = B^{\pm}$, $\ell = e$.

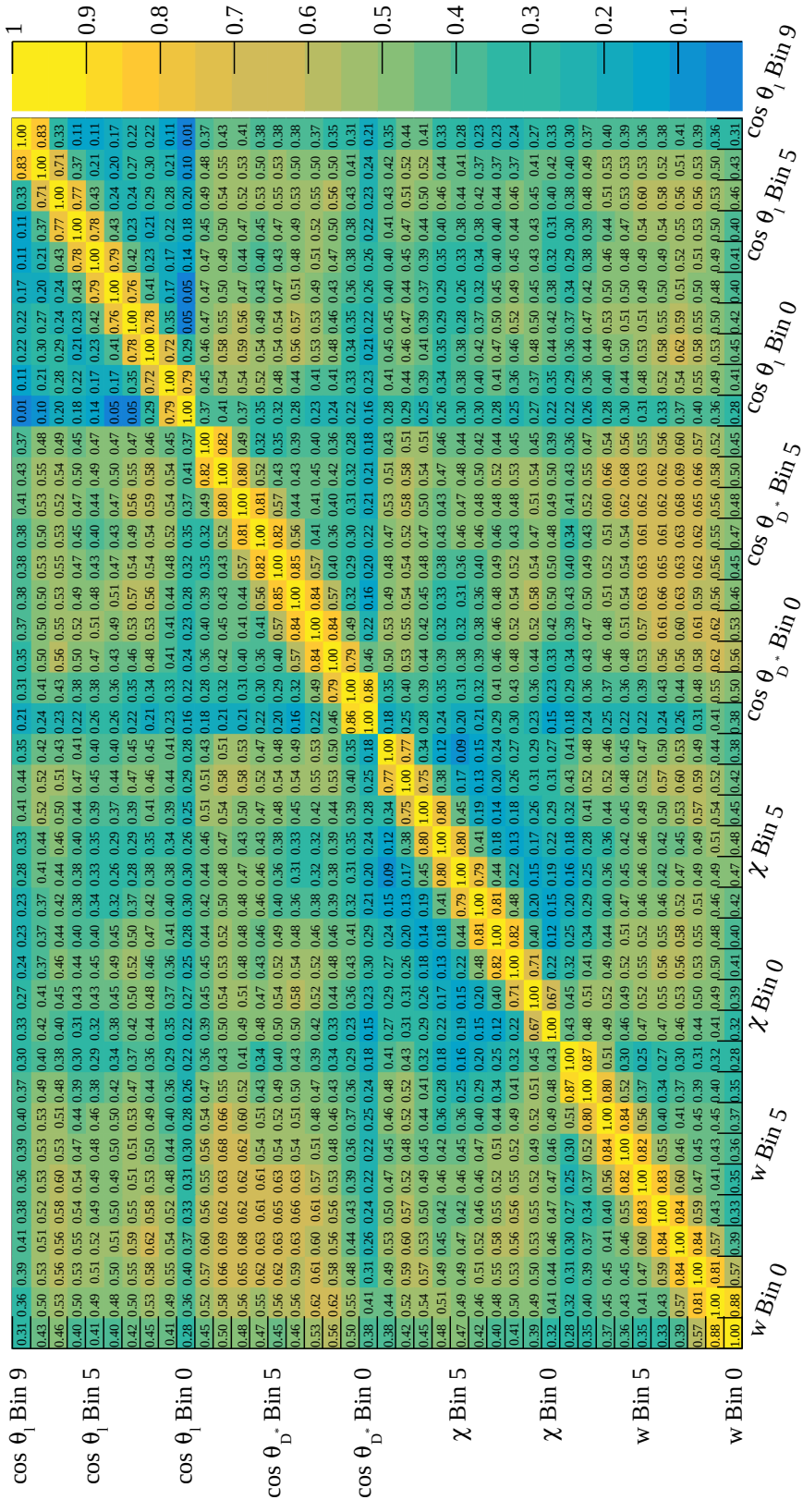


Figure C.44: Correlation matrix for $B = B^{\pm}$, $\ell = \mu$.

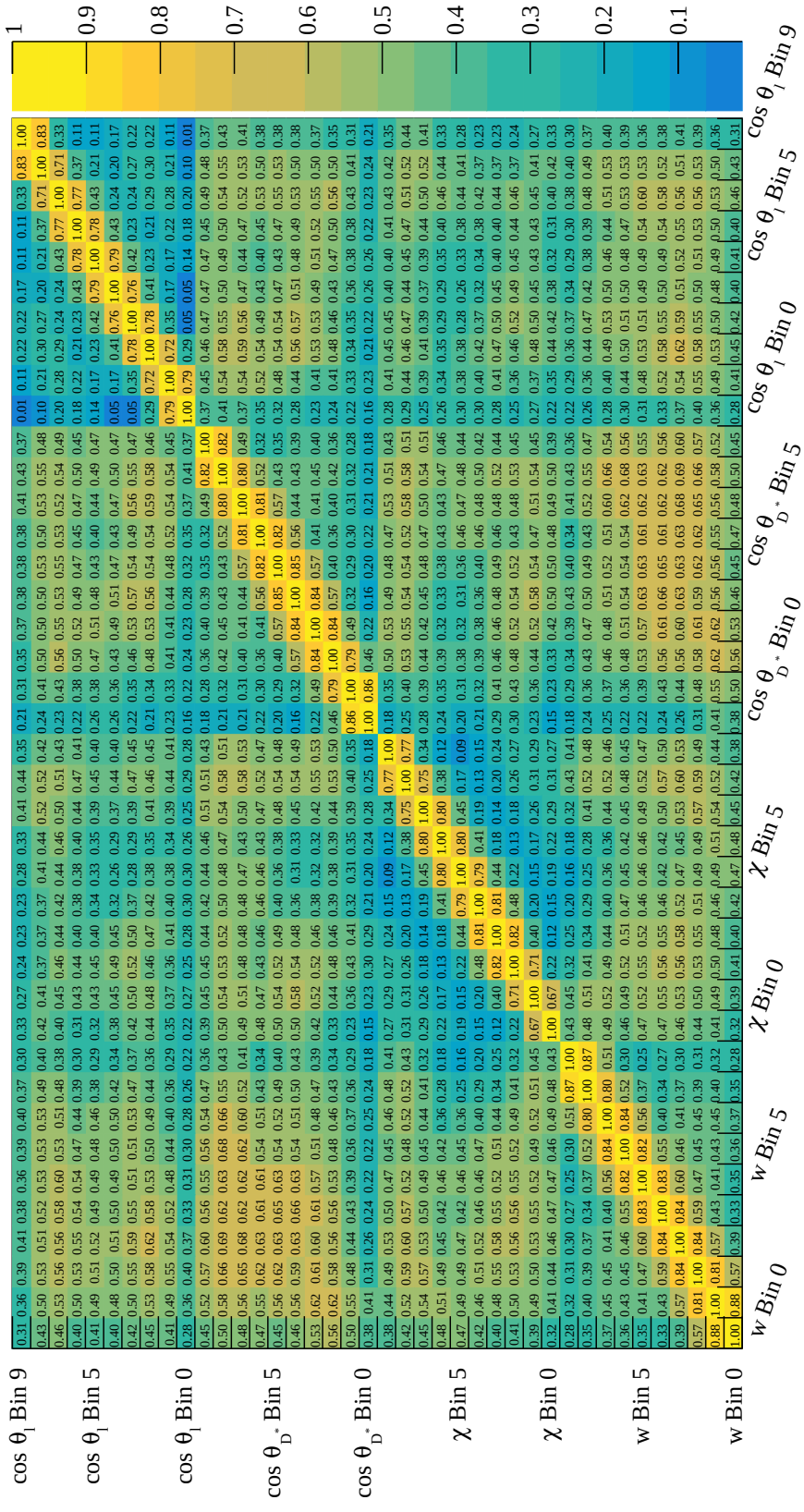


Figure C.45: Correlation matrix for $B = B^{\pm}$, $\ell = e + \mu$.

Acknowledgements

First of all, I would like to thank Prof. Dr. Thomas Kuhr for taking me into his group and his continuing support. I am deeply grateful that you gave me the freedom to explore both theoretical and experimental aspects of my topic and I am looking forward to spending the next years as a PhD student under your supervision as well.

Chapter 2 was driven by and would not have been possible without Dr. Martin Jung, who proposed the project and pointed me in the right directions along the way. Having feedback on chapter 3 from a theorist's perspective was also very helpful.

I am very grateful to Saskia Falke, who started the work on the analysis presented in chapter 2 and was kind enough to share her analysis code with me. Many thanks also go to Prof. Dr. Florian Bernlocher for walking me through the analysis and clearing up many of my misconceptions.

Special thanks goes to Prof. Dr. Gerhard Buchalla for being the second referee for this thesis.

I am very grateful to Prof. Dr. Hiroaki Aihara for hosting me at his group at the University of Tokyo for three months and would also like to thank Aki Naritomi, Kuniko Kono and Prof. Dr. Hikaru Tsuji for making this stay possible. I am also indebted to The University of Tokyo for providing me with a short term scholarship for the duration of the stay.

I would like to thank the Deutsche Forschungsgemeindschaft and the Excellence Cluster Universe for their generous travel support. Furthermore I am greatly indebted to the Studienstiftung des Deutschen Volkes and the Max-Weber Programm who supported me throughout most of my studies.

Finally, let me thank my family and my friends who had my back during all of this. You were there for me when things seemed overwhelming and I could not have done this without you.

List of Tables

List of Figures

Bibliography

[1] S. Weinberg. *The Quantum theory of fields. Vol. 1: Foundations*. Cambridge University Press, 2005. ISBN: 9780521670531, 9780511252044.

[2] S. Weinberg. *The quantum theory of fields. Vol. 2: Modern applications*. Cambridge University Press, 2013. ISBN: 9781139632478, 9780521670548, 9780521550024.

[3] M. Li, X.-D. Li, S. Wang, and Y. Wang. "Dark Energy: A Brief Review". In: *Front. Phys.(Beijing)* 8 (2013), pp. 828–846. DOI: 10.1007/s11467-013-0300-5. arXiv: 1209.0922 [astro-ph.CO].

[4] M. Drees and G. Gerbier. "Mini-Review of Dark Matter: 2012". In: (2012). arXiv: 1204.2373 [hep-ph].

[5] M. Tanabashi et al. "Review of Particle Physics". In: *Phys. Rev.* D 98.030001 (2018), p. 100001.

[6] S. Chatrchyan et al. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC". In: *Phys. Lett.* B716 (2012), pp. 30–61. DOI: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235 [hep-ex].

[7] G. Aad et al. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Phys. Lett.* B716 (2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex].

[8] B. Aubert et al. "The BaBar detector". In: *Nucl. Instrum. Meth.* A479 (2002), pp. 1–116. DOI: 10.1016/S0168-9002(01)02012-5. arXiv: hep-ex/0105044 [hep-ex].

[9] SLAC Collaboration. *Stanford Linear Accelerator Center*. Retrieved 04 Sep 2018. 2018. URL: https://www6.slac.stanford.edu/.

[10] A. Abashian et al. "The Belle Detector". In: *Nucl. Instrum. Meth.* A479 (2002), pp. 117–232. DOI: 10.1016/S0168-9002(01)02013-7.

[11] S. Kurokawa and E. Kikutani. "Overview of the KEKB accelerators". In: *Nucl. Instrum. Meth.* A499 (2003), pp. 1–7. DOI: 10.1016/S0168-9002(02)01771-0.

[12] A. A. Alves Jr. et al. "The LHCb Detector at the LHC". In: *JINST* 3 (2008), S08005. DOI: 10.1088/1748-0221/3/08/S08005.

[13] L. Evans and P. Bryant. "LHC Machine". In: *JINST* 3 (2008), S08001. DOI: 10.1088/1748-0221/3/08/S08001.

[14] J. P. Lees et al. "Measurement of an Excess of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ Decays and Implications for Charged Higgs Bosons". In: *Phys. Rev.* D88.7 (2013), p. 072012. DOI: 10.1103/PhysRevD.88.072012. arXiv: 1303.0571 [hep-ex].

[15] M. Huschle et al. "Measurement of the branching ratio of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ relative to $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ decays with hadronic tagging at Belle". In: *Phys. Rev.* D92.7 (2015), p. 072014. DOI: 10.1103/PhysRevD.92.072014. arXiv: 1507.03233 [hep-ex].

[16] Y. Sato et al. "Measurement of the branching ratio of $\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau$ relative to $\bar{B}^0 \to D^{*+}\ell^-\bar{\nu}_\ell$ decays with a semileptonic tagging method". In: *Phys. Rev.* D94.7 (2016), p. 072007. DOI: 10.1103/PhysRevD.94.072007. arXiv: 1607.07923 [hep-ex].

[17] R. Aaij et al. "Measurement of the ratio of branching fractions $\mathcal{B}(\bar{B}^0 \longrightarrow D^{*+}\tau^-\bar{\nu}_\tau)/\mathcal{B}(\bar{B}^0 \longrightarrow D^{*+}\mu^-\bar{\nu}_\mu)$". In: *Phys. Rev. Lett.* 115.11 (2015). [Erratum: Phys. Rev. Lett.115,no.15,159901(2015)], p. 111803. DOI: 10.1103/PhysRevLett.115.159901, 10.1103/PhysRevLett.115.111803. arXiv: 1506.08614 [hep-ex].

[18] S. Hirose et al. "Measurement of the τ lepton polarization and $R(D^*)$ in the decay $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ with one-prong hadronic τ decays at Belle". In: *Phys. Rev.* D97.1 (2018), p. 012004. DOI: 10.1103/PhysRevD.97.012004. arXiv: 1709.00129 [hep-ex].

[19] R. Aaij et al. "Measurement of the ratio of the $B^0 \to D^{*-}\tau^+\nu_\tau$ and $B^0 \to D^{*-}\mu^+\nu_\mu$ branching fractions using three-prong τ-lepton decays". In: *Phys. Rev. Lett.* 120.17 (2018), p. 171802. DOI: 10.1103/PhysRevLett.120.171802. arXiv: 1708.08856 [hep-ex].

[20] HFLAV Collaboration. *Average of R(D) and R(D*) for Summer 2018*. Retrieved 17 Aug 2018. 2018. URL: https://hflav-eos.web.cern.ch/hflav-eos/semi/summer18/RDRDs.html.

[21] S. Fajfer, J. F. Kamenik, I. Nisandzic, and J. Zupan. "Implications of Lepton Flavor Universality Violations in B Decays". In: *Phys. Rev. Lett.* 109 (2012), p. 161801. DOI: 10.1103/PhysRevLett.109.161801. arXiv: 1206.1872 [hep-ph].

[22] M. Freytsis, Z. Ligeti, and J. T. Ruderman. "Flavor models for $\bar{B} \to D^{(*)}\tau\bar{\nu}$". In: *Phys. Rev.* D92.5 (2015), p. 054018. DOI: 10.1103/PhysRevD.92.054018. arXiv: 1506.08896 [hep-ph].

[23] SLAC/BaBar Collaboration. *SLAC/BaBar Pictures Gallery*. Retrieved 04 Sep 2018. 2018. URL: http://www.slac.stanford.edu/xorg/BFLB/pictures.html.

[24] Belle II Collaboration. *Rollout of the Belle I Detector*. Retrieved 04 Sep 2018. URL: http://belle.kek.jp/~isamun/kaitai/rollout/img044.jpeg.html.

[25] LHCb Collaboration. *LHCb Collaboration Group Picture*. Retrieved 04 Sep 2018. 2018. URL: http://lhcb-public.web.cern.ch/lhcb-public/lhcb_collaboration.jpg.

[26] S. Collaboration. *Aerial View of SLAC*. Retrieved 04 Sep 2018. URL: http://home.slac.stanford.edu/pressreleases/images/SLACaerial-lg.jpg.

[27] Ichinoseki City. *Aerial View of KEKB*. Retrieved 04 Sep 2018. 2018. URL: https://www.city.ichinoseki.iwate.jp/images/content/20302/_kekbgaikan.jpg.

[28] C. Maximilien Brice. *Aerial View of the LHC*. Retrieved 04 Sep 2018. 2017. URL: https://home.cern/sites/home.web.cern.ch/files/image/update-for_the_public/2017/12/0807031_01-a4-at-144-dpi.jpg.

[29] M. Tanaka. "Charged Higgs effects on exclusive semitauonic B decays". In: *Z. Phys.* C67 (1995), pp. 321–326. DOI: 10.1007/BF01571294. arXiv: hep-ph/9411405 [hep-ph].

[30] I. Doršner, S. Fajfer, A. Greljo, J. F. Kamenik, and N. Košnik. "Physics of leptoquarks in precision experiments and at particle colliders". In: *Phys. Rept.* 641 (2016), pp. 1–68. DOI: 10.1016/j.physrep.2016.06.001. arXiv: 1603.04993 [hep-ph].

[31] E. Kou et al. "The Belle II Physics book". In: (2018). arXiv: 1808.10567 [hep-ex].

[32] Y. Ohnishi et al. "Accelerator design at SuperKEKB". In: *PTEP* 2013 (2013), 03A011. DOI: 10.1093/ptep/pts083.

[33] R. Aaij et al. "Test of lepton universality using $B^+ \to K^+\ell^+\ell^-$ decays". In: *Phys. Rev. Lett.* 113 (2014), p. 151601. DOI: 10.1103/PhysRevLett.113.151601. arXiv: 1406.6482 [hep-ex].

[34] R. Aaij et al. "Test of lepton universality with $B^0 \to K^{*0}\ell^+\ell^-$ decays". In: *JHEP* 08 (2017), p. 055. DOI: 10.1007/JHEP08(2017)055. arXiv: 1705.05802 [hep-ex].

[35] A. Pich. "The Standard model of electroweak interactions". In: *2004 European School of High-Energy Physics, Sant Feliu de Guixols, Spain, 30 May - 12 June 2004*. 2005, pp. 1–48. arXiv: hep-ph/0502010 [hep-ph]. URL: http://doc.cern.ch/yellowrep/2006/2006-003/p1.pdf.

[36] F. Englert and R. Brout. "Broken Symmetry and the Mass of Gauge Vector Mesons". In: *Phys. Rev. Lett.* 13 (1964). [,157(1964)], pp. 321–323. DOI: 10.1103/PhysRevLett.13.321.

[37] P. W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons". In: *Phys. Rev. Lett.* 13 (1964). [,160(1964)], pp. 508–509. DOI: 10.1103/PhysRevLett.13.508.

[38] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. "Global Conservation Laws and Massless Particles". In: *Phys. Rev. Lett.* 13 (1964). [,162(1964)], pp. 585–587. DOI: 10.1103/PhysRevLett.13.585.

[39] Z. Maki, M. Nakagawa, and S. Sakata. "Remarks on the unified model of elementary particles". In: *Prog. Theor. Phys.* 28 (1962). [,34(1962)], pp. 870–880. DOI: 10.1143/PTP.28.870.

[40] S. L. Glashow, J. Iliopoulos, and L. Maiani. "Weak Interactions with Lepton-Hadron Symmetry". In: *Phys. Rev.* D2 (1970), pp. 1285–1292. DOI: 10.1103/PhysRevD.2.1285.

[41] B. Pontecorvo. "Inverse beta processes and nonconservation of lepton charge". In: *Sov. Phys. JETP* 7 (1958). [Zh. Eksp. Teor. Fiz.34,247(1957)], pp. 172–173.

[42] P. F. de Salas, D. V. Forero, C. A. Ternes, M. Tortola, and J. W. F. Valle. "Status of neutrino oscillations 2018: 3σ hint for normal mass ordering and improved CP sensitivity". In: *Phys. Lett.* B782 (2018), pp. 633–640. DOI: 10.1016/j.physletb.2018.06.019. arXiv: 1708.01186 [hep-ph].

[43] M. Dine and A. Kusenko. "The Origin of the matter - antimatter asymmetry". In: *Rev. Mod. Phys.* 76 (2003), p. 1. DOI: 10.1103/RevModPhys.76.1. arXiv: hep-ph/0303065 [hep-ph].

[44] T. Blum, A. Denig, I. Logashenko, E. de Rafael, B. Lee Roberts, T. Teubner, and G. Venanzoni. "The Muon (g-2) Theory Value: Present and Future". In: (2013). arXiv: 1311.2198 [hep-ph].

[45] R. Pohl, R. Gilman, G. A. Miller, and K. Pachucki. "Muonic hydrogen and the proton radius puzzle". In: *Ann. Rev. Nucl. Part. Sci.* 63 (2013), pp. 175–204. DOI: 10.1146/annurev-nucl-102212-170627. arXiv: 1301.0905 [physics.atom-ph].

[46] G. Ciezarek, M. Franco Sevilla, B. Hamilton, R. Kowalewski, T. Kuhr, V. Lüth, and Y. Sato. "A Challenge to Lepton Universality in B Meson Decays". In: (2017). arXiv: 1703.01766 [hep-ex].

[47] CERN. *Large Electron-Positron storage ring: technical notebook*. Retrieved 04 Sep 2018. 1989. URL: http://cds.cern.ch/record/113994?ln=en.

[48] M. Acciarri et al. "Measurement of the branching ratios $b \longrightarrow e\nu X$, $\mu\nu X$, $\tau\nu X$ and νX". In: *Z. Phys.* C71 (1996), pp. 379–390. DOI: 10.1007/s002880050184.

[49] A. Matyja et al. "Observation of $B^0 \longrightarrow D^{*-}\tau^+\nu_\tau$ decay at Belle". In: *Phys. Rev. Lett.* 99 (2007), p. 191807. DOI: 10.1103/PhysRevLett.99.191807. arXiv: 0706.4429 [hep-ex].

[50] B. Aubert et al. "Observation of the semileptonic decays $B \to D^*\tau^-\bar{\nu}_\tau$ and evidence for $B \to D\tau^-\bar{\nu}_\tau$". In: *Phys. Rev. Lett.* 100 (2008), p. 021801. DOI: 10.1103/PhysRevLett.100.021801. arXiv: 0709.1698 [hep-ex].

[51] D. Bigi and P. Gambino. "Revisiting $B \to D\ell\nu$". In: *Phys. Rev.* D94.9 (2016), p. 094008. DOI: 10.1103/PhysRevD.94.094008. arXiv: 1606.08030 [hep-ph].

[52] D. Bigi, P. Gambino, and S. Schacht. "$R(D^*)$, $|V_{cb}|$, and the Heavy Quark Symmetry relations between form factors". In: *JHEP* 11 (2017), p. 061. DOI: 10.1007/JHEP11(2017)061. arXiv: 1707.09509 [hep-ph].

[53] F. U. Bernlochner, Z. Ligeti, M. Papucci, and D. J. Robinson. "Combined analysis of semileptonic B decays to D and D^*: $R(D^{(*)})$, $|V_{cb}|$, and new physics". In: *Phys. Rev.* D95.11 (2017), p. 115008. DOI: 10.1103/PhysRevD.95.115008. arXiv: 1703.05330 [hep-ph].

[54] S. Jaiswal, S. Nandi, and S. K. Patra. "Extraction of $|V_{cb}|$ from $B \to D^{(*)}\ell\nu_\ell$ and the Standard Model predictions of $R(D^{(*)})$". In: *JHEP* 12 (2017), p. 060. DOI: 10.1007/JHEP12(2017)060. arXiv: 1707.09977 [hep-ph].

[55] J. A. Bailey et al. "$B \longrightarrow D\ell\nu$ form factors at nonzero recoil and V_{cb}| from 2+1-flavor lattice QCD". In: *Phys. Rev.* D92.3 (2015), p. 034506. DOI: 10.1103/PhysRevD.92.034506. arXiv: 1503.07237 [hep-lat].

[56] H. Na, C. M. Bouchard, G. P. Lepage, C. Monahan, and J. Shigemitsu. "$B \to Dl\nu$ form factors at nonzero recoil and extraction of $|V_{cb}|$". In: *Phys. Rev.* D92.5 (2015). [Erratum: Phys. Rev.D93,no.11,119906(2016)], p. 054510. DOI: 10.1103/PhysRevD.93.119906,10.1103/PhysRevD.92.054510. arXiv: 1505.03925 [hep-lat].

[57] B. Aubert et al. "Measurement of —V(cb)— and the Form-Factor Slope in anti-B —¿ D l- anti-nu Decays in Events Tagged by a Fully Reconstructed B Meson". In: *Phys. Rev. Lett.* 104 (2010), p. 011802. DOI: 10.1103/PhysRevLett.104.011802. arXiv: 0904.4063 [hep-ex].

[58] R. Glattauer et al. "Measurement of the decay $B \to D\ell\nu_\ell$ in fully reconstructed events and determination of the Cabibbo-Kobayashi-Maskawa matrix element $|V_{cb}|$". In: *Phys. Rev.* D93.3 (2016), p. 032006. DOI: 10.1103/PhysRevD.93.032006. arXiv: 1510.03657 [hep-ex].

[59] S. Fajfer, J. F. Kamenik, and I. Nisandzic. "On the $B \to D^*\tau\bar{\nu}_\tau$ Sensitivity to New Physics". In: *Phys. Rev.* D85 (2012), p. 094025. DOI: 10.1103/PhysRevD.85.094025. arXiv: 1203.2654 [hep-ph].

[60] I. Caprini, L. Lellouch, and M. Neubert. "Dispersive bounds on the shape of $\bar{B} \longrightarrow D^{(*)}$ lepton anti-neutrino form-factors". In: *Nucl. Phys.* B530 (1998), pp. 153–181. DOI: 10.1016/S0550-3213(98)00350-2. arXiv: hep-ph/9712417 [hep-ph].

[61] A. Abdesselam et al. "Precise determination of the CKM matrix element $|V_{cb}|$ with $\bar{B}^0 \to D^{*+}\ell^-\bar{\nu}_\ell$ decays with hadronic tagging at Belle". In: (2017). arXiv: 1702.01521 [hep-ex].

[62] C. G. Boyd, B. Grinstein, and R. F. Lebed. "Constraints on form-factors for exclusive semileptonic heavy to light meson decays". In: *Phys. Rev. Lett.* 74 (1995), pp. 4603–4606. DOI: 10.1103/PhysRevLett.74.4603. arXiv: hep-ph/9412324 [hep-ph].

[63] C. G. Boyd, B. Grinstein, and R. F. Lebed. "Precision corrections to dispersive bounds on form-factors". In: *Phys. Rev.* D56 (1997), pp. 6895–6911. DOI: 10.1103/PhysRevD.56.6895. arXiv: hep-ph/9705252 [hep-ph].

[64] C. G. Boyd, B. Grinstein, and R. F. Lebed. "Model independent determinations of anti-B —¿ D (lepton), D* (lepton) anti-neutrino form-factors". In: *Nucl. Phys.* B461 (1996), pp. 493–511. DOI: 10.1016/0550-3213(95)00653-2. arXiv: hep-ph/9508211 [hep-ph].

[65] B. Aubert et al. "Measurements of the Semileptonic Decays anti-B —¿ D l anti-nu and anti-B —¿ D* l anti-nu Using a Global Fit to D X l anti-nu Final States". In: *Phys. Rev.* D79 (2009), p. 012002. DOI: 10.1103/PhysRevD.79.012002. arXiv: 0809.0828 [hep-ex].

[66] W. Dungel et al. "Measurement of the form factors of the decay B0 -¿ D*- ell+ nu and determination of the CKM matrix element —Vcb—". In: *Phys. Rev.* D82 (2010), p. 112007. DOI: 10.1103/PhysRevD.82.112007. arXiv: 1010.5620 [hep-ex].

[67] J. P. Lees et al. "Evidence of $B^+ \to \tau^+\nu$ decays with hadronic B tags". In: *Phys. Rev.* D88.3 (2013), p. 031102. DOI: 10.1103/PhysRevD.88.031102. arXiv: 1207.0698 [hep-ex].

[68] B. Kronenbitter et al. “Measurement of the branching fraction of $B^+ \longrightarrow \tau^+\nu_\tau$ decays with the semileptonic tagging method”. In: *Phys. Rev.* D92.5 (2015), p. 051102. DOI: 10.1103/PhysRevD.92.051102. arXiv: 1503.05613 [hep-ex].

[69] HFLAV collaboration. *HFLAV average of* $\mathrm{BR}(B \longrightarrow \tau\nu)$ *as of May 2018*. Retrieved 19 Aug 2018. 2018. URL: http://www.slac.stanford.edu/xorg/hflav/rare/May2018/radll/OUTPUT/HTML/tauPLUSnu3.html.

[70] G. Hiller and F. Kruger. “More model-independent analysis of $b \to s$ processes”. In: *Phys. Rev.* D69 (2004), p. 074020. DOI: 10.1103/PhysRevD.69.074020. arXiv: hep-ph/0310219 [hep-ph].

[71] S. Wehle et al. “Lepton-Flavor-Dependent Angular Analysis of $B \to K^*\ell^+\ell^-$”. In: *Phys. Rev. Lett.* 118.11 (2017), p. 111801. DOI: 10.1103/PhysRevLett.118.111801. arXiv: 1612.05014 [hep-ex].

[72] CMS Collaboration. “Measurement of the P_1 and P_5' angular parameters of the decay $\mathrm{B}^0 \to \mathrm{K}^{*0}\mu^+\mu^-$ in proton-proton collisions at $\sqrt{s} = 8$ TeV”. In: (2017).

[73] The ATLAS collaboration. “Angular analysis of $B_d^0 \to K^*\mu^+\mu^-$ decays in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector”. In: (2017).

[74] W. Altmannshofer, P. Stangl, and D. M. Straub. “Interpreting Hints for Lepton Flavor Universality Violation”. In: *Phys. Rev.* D96.5 (2017), p. 055008. DOI: 10.1103/PhysRevD.96.055008. arXiv: 1704.05435 [hep-ph].

[75] B. Grzadkowski, M. Iskrzynski, M. Misiak, and J. Rosiek. “Dimension-Six Terms in the Standard Model Lagrangian”. In: *JHEP* 10 (2010), p. 085. DOI: 10.1007/JHEP10(2010)085. arXiv: 1008.4884 [hep-ph].

[76] T. Appelquist and J. Carazzone. “Infrared Singularities and Massive Fields”. In: *Phys. Rev.* D11 (1975), p. 2856. DOI: 10.1103/PhysRevD.11.2856.

[77] D. Blaschke, M. A. Ivanov, and T. Mannel, eds. *Heavy quark physics*. Vol. 647. 2004. DOI: 10.1007/b97728.

[78] C. Grojean, E. Salvioni, and R. Torre. “A weakly constrained W’ at the early LHC”. In: *JHEP* 07 (2011), p. 002. DOI: 10.1007/JHEP07(2011)002. arXiv: 1103.2761 [hep-ph].

[79] C. Patrignani et al. “Review of Particle Physics: Leptoquark Quantum Numbers”. In: *Chin. Phys.* C40.10 (2016), p. 100001. DOI: 10.1088/1674-1137/40/10/100001.

[80] F. del Aguila, J. de Blas, and M. Perez-Victoria. “Electroweak Limits on General New Vector Bosons”. In: *JHEP* 09 (2010), p. 033. DOI: 10.1007/JHEP09(2010)033. arXiv: 1005.3998 [hep-ph].

[81] G. C. Branco, P. M. Ferreira, L. Lavoura, M. N. Rebelo, M. Sher, and J. P. Silva. “Theory and phenomenology of two-Higgs-doublet models”. In: *Phys. Rept.* 516 (2012), pp. 1–102. DOI: 10.1016/j.physrep.2012.02.002. arXiv: 1106.0034 [hep-ph].

[82] W. Buchmuller and C. Ludeling. “Field Theory and Standard Model”. In: *High-energy physics. Proceedings, European School, Kitzbuehel, Austria, August 21-September, 2005*. 2006. arXiv: hep-ph/0609174 [hep-ph]. URL: http://inspirehep.net/record/726283/files/arXiv:hep-ph_0609174.pdf.

[83] C. P. et. al. (Particle Data Group). *Electroweak Model and Constraints on New Physics (PDG Review, 2017 version)*. Retrieved 08 Aug 2018. 2017. URL: http://pdg.lbl.gov/2017/reviews/rpp2017-rev-standard-model.pdf.

[84] F. Mahmoudi and O. Stal. “Flavor constraints on the two-Higgs-doublet model with general Yukawa couplings”. In: *Phys. Rev.* D81 (2010), p. 035016. DOI: 10.1103/PhysRevD.81.035016. arXiv: 0907.1791 [hep-ph].

[85] M. M. Mühlleitner. *Beyond the Standard Model Physics*. Lecture at Karlsruhe Institute of Technology (KIT). Retrieved May 2017. 2015. URL: https://www.itp.kit.edu/~rauch/Teaching/WS1415_BSMHiggs/bsm.pdf.

[86] S. L. Glashow and S. Weinberg. "Natural Conservation Laws for Neutral Currents". In: *Phys. Rev.* D15 (1977), p. 1958. DOI: 10.1103/PhysRevD.15.1958.

[87] E. A. Paschos. "Diagonal Neutral Currents". In: *Phys. Rev.* D15 (1977), p. 1966. DOI: 10.1103/PhysRevD.15.1966.

[88] H. Na, C. M. Bouchard, G. P. Lepage, C. Monahan, and J. Shigemitsu. "$B \to Dl\nu$ form factors at nonzero recoil and extraction of $|V_{cb}|$". In: *Phys. Rev.* D92.5 (2015). [Erratum: Phys. Rev.D93,no.11,119906(2016)], p. 054510. DOI: 10.1103/PhysRevD.93.119906,10.1103/PhysRevD.92.054510. arXiv: 1505.03925 [hep-lat].

[89] M. Misiak et al. "Estimate of $\mathcal{B}(\bar{B} \to X_s\gamma)$ at $O(\alpha_s^2)$". In: *Phys. Rev. Lett.* 98 (2007), p. 022002. DOI: 10.1103/PhysRevLett.98.022002. arXiv: hep-ph/0609232 [hep-ph].

[90] A. Pich and P. Tuzon. "Yukawa Alignment in the Two-Higgs-Doublet Model". In: *Phys. Rev.* D80 (2009), p. 091702. DOI: 10.1103/PhysRevD.80.091702. arXiv: 0908.1554 [hep-ph].

[91] M. Jung, A. Pich, and P. Tuzon. "Charged-Higgs phenomenology in the Aligned two-Higgs-doublet model". In: *JHEP* 11 (2010), p. 003. DOI: 10.1007/JHEP11(2010)003. arXiv: 1006.0470 [hep-ph].

[92] T. P. Cheng and M. Sher. "Mass Matrix Ansatz and Flavor Nonconservation in Models with Multiple Higgs Doublets". In: *Phys. Rev.* D35 (1987), p. 3484. DOI: 10.1103/PhysRevD.35.3484.

[93] I. P. Ivanov. "Building and testing models with extended Higgs sectors". In: (2017). arXiv: 1702.03776 [hep-ph].

[94] A. Celis, M. Jung, X.-Q. Li, and A. Pich. "Scalar contributions to $b \to c(u)\tau\nu$ transitions". In: (2016). arXiv: 1612.07757 [hep-ph].

[95] A. Celis, M. Jung, X.-Q. Li, and A. Pich. "Sensitivity to charged scalars in $\boldsymbol{B \to D^{(*)}\tau\nu_\tau}$ and $\boldsymbol{B \to \tau\nu_\tau}$ decays". In: *JHEP* 01 (2013), p. 054. DOI: 10.1007/JHEP01(2013)054. arXiv: 1210.8443 [hep-ph].

[96] Z. Sullivan. "Fully differential W' production and decay at next-to-leading order in QCD". In: *Phys. Rev.* D66 (2002), p. 075011. DOI: 10.1103/PhysRevD.66.075011. arXiv: hep-ph/0207290 [hep-ph].

[97] B. A. Dobrescu and Z. Liu. "W' Boson near 2 TeV: Predictions for Run 2 of the LHC". In: *Phys. Rev. Lett.* 115.21 (2015), p. 211802. DOI: 10.1103/PhysRevLett.115.211802. arXiv: 1506.06736 [hep-ph].

[98] C. S. Aulakh, A. Melfo, and G. Senjanovic. "Minimal supersymmetric left-right model". In: *Phys. Rev.* D57 (1998), pp. 4174–4178. DOI: 10.1103/PhysRevD.57.4174. arXiv: hep-ph/9707256 [hep-ph].

[99] J. C. Pati and A. Salam. "Lepton Number as the Fourth Color". In: *Phys. Rev.* D10 (1974). [Erratum: Phys. Rev.D11,703(1975)], pp. 275–289. DOI: 10.1103/PhysRevD.10.275,10.1103/PhysRevD.11.703.2.

[100] R. N. Mohapatra and J. C. Pati. "Left-Right Gauge Symmetry and an Isoconjugate Model of CP Violation". In: *Phys. Rev.* D11 (1975), pp. 566–571. DOI: 10.1103/PhysRevD.11.566.

[101] R. N. Mohapatra and J. C. Pati. "A Natural Left-Right Symmetry". In: *Phys. Rev.* D11 (1975), p. 2558. DOI: 10.1103/PhysRevD.11.2558.

[102] G. Senjanovic and R. N. Mohapatra. "Exact Left-Right Symmetry and Spontaneous Violation of Parity". In: *Phys. Rev.* D12 (1975), p. 1502. DOI: 10.1103/PhysRevD.12.1502.

[103] P. Cox, A. Kusenko, O. Sumensari, and T. T. Yanagida. "SU(5) Unification with TeV-scale Leptoquarks". In: *JHEP* 03 (2017), p. 035. DOI: 10.1007/JHEP03(2017)035. arXiv: 1612.03923 [hep-ph].

[104] H. Georgi and S. L. Glashow. "Unity of All Elementary Particle Forces". In: *Phys. Rev. Lett.* 32 (1974), pp. 438–441. DOI: 10.1103/PhysRevLett.32.438.

[105] H. Fritzsch and P. Minkowski. "Unified Interactions of Leptons and Hadrons". In: *Annals Phys.* 93 (1975), pp. 193–266. DOI: 10.1016/0003-4916(75)90211-0.

[106] P. H. Frampton and B.-H. Lee. "SU15 Grand Unification". In: *Phys. Rev. Lett.* 64 (1990), p. 619. DOI: 10.1103/PhysRevLett.64.619.

[107] E. Farhi and L. Susskind. "Technicolor". In: *Phys. Rept.* 74 (1981), p. 277. DOI: 10.1016/0370-1573(81)90173-3.

[108] B. Schrempp and F. Schrempp. "Light Leptoquarks". In: *Phys. Lett.* 153B (1985), pp. 101–107. DOI: 10.1016/0370-2693(85)91450-9.

[109] W. Buchmuller, R. Ruckl, and D. Wyler. "Leptoquarks in Lepton - Quark Collisions". In: *Phys. Lett.* B191 (1987). [Erratum: Phys. Lett.B448,320(1999)], pp. 442–448. DOI: 10.1016/S0370-2693(99)00014-3,10.1016/0370-2693(87)90637-X.

[110] J. L. Hewett and T. G. Rizzo. "Don't stop thinking about leptoquarks: Constructing new models". In: *Phys. Rev.* D58 (1998), p. 055005. DOI: 10.1103/PhysRevD.58.055005. arXiv: hep-ph/9708419 [hep-ph].

[111] J. M. Arnold, B. Fornal, and M. B. Wise. "Simplified models with baryon number violation but no proton decay". In: *Phys. Rev.* D87 (2013), p. 075004. DOI: 10.1103/PhysRevD.87.075004. arXiv: 1212.4556 [hep-ph].

[112] M. Leurer. "A Comprehensive study of leptoquark bounds". In: *Phys. Rev.* D49 (1994), pp. 333–342. DOI: 10.1103/PhysRevD.49.333. arXiv: hep-ph/9309266 [hep-ph].

[113] M. Leurer. "New bounds on leptoquarks". In: *Phys. Rev. Lett.* 71 (1993), pp. 1324–1327. DOI: 10.1103/PhysRevLett.71.1324. arXiv: hep-ph/9304211 [hep-ph].

[114] A. Djouadi, T. Kohler, M. Spira, and J. Tutas. "$(e\ b)$, $(e\ t)$ type leptoquarks at e^-e^+ colliders". In: *Z. Phys.* C46 (1990), pp. 679–686. DOI: 10.1007/BF01560270.

[115] I. Doršner, S. Fajfer, A. Greljo, J. F. Kamenik, and N. Košnik. "Physics of leptoquarks in precision experiments and at particle colliders". In: *Phys. Rept.* 641 (2016), pp. 1–68. DOI: 10.1016/j.physrep.2016.06.001. arXiv: 1603.04993 [hep-ph].

[116] M. Duraisamy, P. Sharma, and A. Datta. "Azimuthal $B \to D^*\tau^-\bar{\nu_\tau}$ angular distribution with tensor operators". In: *Phys. Rev.* D90.7 (2014), p. 074013. DOI: 10.1103/PhysRevD.90.074013. arXiv: 1405.3719 [hep-ph].

[117] D. Becirevic, S. Fajfer, I. Nisandzic, and A. Tayduganov. "Angular distributions of $\bar{B} \to D^{(*)}\ell\bar{\nu}_\ell$ decays and search of New Physics". In: (2016). arXiv: 1602.03030 [hep-ph].

[118] Y. Sakaki, M. Tanaka, A. Tayduganov, and R. Watanabe. "Testing leptoquark models in $\bar{B} \to D^{(*)}\tau\bar{\nu}$". In: *Phys. Rev.* D88.9 (2013), p. 094012. DOI: 10.1103/PhysRevD.88.094012. arXiv: 1309.0301 [hep-ph].

[119] M. Tanaka and R. Watanabe. "New physics in the weak interaction of $\bar{B} \to D^{(*)}\tau\bar{\nu}$". In: *Phys. Rev.* D87.3 (2013), p. 034028. DOI: 10.1103/PhysRevD.87.034028. arXiv: 1212.1878 [hep-ph].

[120] Y. Sakaki, M. Tanaka, A. Tayduganov, and R. Watanabe. "Probing New Physics with q^2 distributions in $\bar{B} \to D^{(*)}\tau\bar{\nu}$". In: *Phys. Rev.* D91.11 (2015), p. 114028. DOI: 10.1103/PhysRevD.91.114028. arXiv: 1412.3761 [hep-ph].

[121] T. D. Cohen, H. Lamm, and R. F. Lebed. "Tests of the Standard Model in $B \to D\ell\nu_\ell$, $B \to D^*\ell\nu_\ell$ and $B_c \to J/\psi\,\ell\nu_\ell$". In: (2018). arXiv: 1807.00256 [hep-ph].

[122] P. Biancofiore, P. Colangelo, and F. De Fazio. "On the anomalous enhancement observed in $B \to D^{(*)}\tau\bar{\nu}_\tau$ decays". In: *Phys. Rev.* D87.7 (2013), p. 074010. DOI: 10.1103/PhysRevD.87.074010. arXiv: 1302.1042 [hep-ph].

[123] A. Celis, M. Jung, X.-Q. Li, and A. Pich. "$B \longrightarrow D(*)\tau\nu_\tau$ decays in two-Higgs-doublet models". In: *J. Phys. Conf. Ser.* 447 (2013), p. 012058. DOI: 10.1088/1742-6596/447/1/012058. arXiv: 1302.5992 [hep-ph].

[124] M. Duraisamy and A. Datta. "The Full $B \to D^*\tau^-\bar{\nu_\tau}$ Angular Distribution and CP violating Triple Products". In: *JHEP* 09 (2013), p. 059. DOI: 10.1007/JHEP09(2013)059. arXiv: 1302.7031 [hep-ph].

[125] C.-H. Chen and C.-Q. Geng. "Lepton angular asymmetries in semileptonic charmful B decays". In: *Phys. Rev.* D71 (2005), p. 077501. DOI: 10.1103/PhysRevD.71.077501. arXiv: hep-ph/0503123 [hep-ph].

[126] A. Datta, M. Duraisamy, and D. Ghosh. "Diagnosing New Physics in $b \to c\,\tau\,\nu_\tau$ decays in the light of the recent BaBar result". In: *Phys. Rev.* D86 (2012), p. 034027. DOI: 10.1103/PhysRevD.86.034027. arXiv: 1206.3760 [hep-ph].

[127] M. A. Ivanov, J. G. Körner, and C.-T. Tran. "Analyzing new physics in the decays $\bar{B}^0 \to D^{(*)}\tau^-\bar{\nu}_\tau$ with form factors obtained from the covariant quark model". In: *Phys. Rev.* D94.9 (2016), p. 094028. DOI: 10.1103/PhysRevD.94.094028. arXiv: 1607.02932 [hep-ph].

[128] W. Dungel et al. "Measurement of the form factors of the decay $B^0 \longrightarrow D^{*-}\ell^+\nu$ and determination of the CKM matrix element $|V_{cb}|$". In: *Phys. Rev.* D82 (2010), p. 112007. DOI: 10.1103/PhysRevD.82.112007. arXiv: 1010.5620 [hep-ex].

[129] S. Mönig/Falke, F. Bernlocher, J. Hasenbusch, M. Rüger, and J. Dingfelder. *Belle Note 1419: Determination of the CKM matrix element $|V_{cb}|$ and search for new physics using semileptonic $B \to D^*\ell\nu$ decays with hadronic tagging*. Retrieved 16 Mar 2018. 2016. URL: http://belle.kek.jp/secured/belle_note/gn1419/.

[130] A. J. Bevan et al. "The Physics of the B Factories". In: *Eur. Phys. J.* C74 (2014), p. 3026. DOI: 10.1140/epjc/s10052-014-3026-9. arXiv: 1406.6311 [hep-ex].

[131] I. Abe et al. "The KEKB injector linac". In: *Nucl. Instrum. Meth.* A499 (2003), pp. 167–190. DOI: 10.1016/S0168-9002(02)01787-4.

[132] Belle collaboration. *Belle Luminosity Records*. Retrieved July 30 2018Retrieved 30 Jul 2018. 2009. URL: http://www-acc.kek.jp/kekb/commissioning/record/luminosity_record.html.

[133] T. Abe et al. *Belle II Technical Design Report*. 2010. arXiv: 1011.0352 [physics.ins-det].

[134] R. Abe et al. "The new beampipe for the Belle experiment". In: *Nucl. Instrum. Meth.* A535 (2004), pp. 558–561. DOI: 10.1016/j.nima.2004.07.260.

[135] Z. Natkaniec et al. "Status of the Belle silicon vertex detector". In: *Nucl. Instrum. Meth.* A560 (2006), pp. 1–4. DOI: 10.1016/j.nima.2005.11.228.

[136] M. Feindt, F. Keller, M. Kreps, T. Kuhr, S. Neubauer, D. Zander, and A. Zupanc. "A Hierarchical NeuroBayes-based Algorithm for Full Reconstruction of B Mesons at B Factories". In: *Nucl. Instrum. Meth.* A654 (2011), pp. 432–440. DOI: 10.1016/j.nima.2011.06.008. arXiv: 1102.3876 [hep-ex].

[137] M. Feindt. "A Neural Bayesian Estimator for Conditional Probability Densities". In: (2004). arXiv: physics/0402093 [physics.data-an].

[138] D. J. Lange. "The EvtGen particle decay simulation package". In: *Nucl. Instrum. Meth.* A462 (2001), pp. 152–155. DOI: 10.1016/S0168-9002(01)00089-4.

[139] R. Brun, F. Bruyant, M. Maire, A. C. McPherson, and P. Zanarini. "GEANT3". In: (1987).

[140] E. Barberio and Z. Was. "PHOTOS: A Universal Monte Carlo for QED radiative corrections. Version 2.0". In: *Comput. Phys. Commun.* 79 (1994), pp. 291–308. DOI: 10.1016/0010-4655(94)90074-4.

[141] C. Patrignani et al. "Review of Particle Physics". In: *Chin. Phys.* C40.10 (2016), p. 100001. DOI: 10.1088/1674-1137/40/10/100001.

[142] K. A. Olive et al. "Review of Particle Physics". In: *Chin. Phys.* C38 (2014), p. 090001. DOI: 10.1088/1674-1137/38/9/090001.

[143] Y. Amhis et al. "Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2014". In: (2014). arXiv: 1412.7515 [hep-ex].

[144] W. Dungel. *Systematic investigation of the reconstruction effiency in low momentum $\pi^{\pm}$ and π^0*. Belle Note 1176 (Belle internal)Retrieved 30 Jul 2018. 2011. URL: http://belle.kek.jp/secured/belle_note/gn1176/bn1176.pdf.

[145] M. Rüger. *Korrektur von b-tagging Effizienzen für eine $B \longrightarrow D^*\ell\nu$-Analyse am Belle Experiment*. Bachelor Thesis. 2016.

[146] R. Glattauer et al. "Measurement of the decay $B \to D\ell\nu_\ell$ in fully reconstructed events and determination of the Cabibbo-Kobayashi-Maskawa matrix element $|V_{cb}|$". In: *Phys. Rev.* D93.3 (2016), p. 032006. DOI: 10.1103/PhysRevD.93.032006. arXiv: 1510.03657 [hep-ex].

[147] W. Verkerke and D. P. Kirkby. "The RooFit toolkit for data modeling". In: *eConf* C0303241 (2003). [,186(2003)], MOLT007. arXiv: physics/0306116 [physics].

[148] K. S. Cranmer. "Kernel estimation in high-energy physics". In: *Comput. Phys. Commun.* 136 (2001), pp. 198–207. DOI: 10.1016/S0010-4655(00)00243-5. arXiv: hep-ex/0011057 [hep-ex].

[149] ROOT documentation. *PyROOT: Python bindings for the ROOT framework*. Retrieved 28 Aug 2018. URL: https://root.cern.ch/pyroot.

[150] A. Hocker and V. Kartvelishvili. "SVD approach to data unfolding". In: *Nucl. Instrum. Meth.* A372 (1996), pp. 469–481. DOI: 10.1016/0168-9002(95)01478-0. arXiv: hep-ph/9509307 [hep-ph].

[151] H. L. Kerstin Tackmann Andreas Hoecker. *TSVDUnfold Class Reference*. Retrieved 27 Jul 2018. 2010. URL: https://root.cern.ch/doc/master/classTSVDUnfold.html.

[152] Y. Amhis et al. "Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2016". In: (2016). arXiv: 1612.07233 [hep-ex].

[153] J. A. Bailey et al. "Update of $|V_{cb}|$ from the $\bar{B} \to D^*\ell\bar{\nu}$ form factor at zero recoil with three-flavor lattice QCD". In: *Phys. Rev.* D89.11 (2014), p. 114504. DOI: 10.1103/PhysRevD.89.114504. arXiv: 1403.0635 [hep-lat].

[154] S. Aoki et al. "Review of lattice results concerning low-energy particle physics". In: *Eur. Phys. J.* C77.2 (2017), p. 112. DOI: 10.1140/epjc/s10052-016-4509-7. arXiv: 1607.00299 [hep-lat].

Selbstständigkeitserklärung

Hiermit versichere ich, die vorliegende Arbeit mit dem Titel

Construction of Angular Observables Sensitive to New Physics in $\bar{B}\longrightarrow D^*\tau^-\bar{\nu}_\tau$ Decays and Measurements of Differential Cross Sections of $\bar{B}\longrightarrow D^*\ell^-\bar{\nu}_\ell$ Decays with Hadronic Tagging at Belle

Konstruktion Winkelabhängiger Observablen Sensitiv auf Neue Physik in $\bar{B}\longrightarrow D^*\tau^-\bar{\nu}_\tau$ Zerfällen und Messung Differenzieller Wirkungsquerschnitte des $\bar{B}\longrightarrow D^*\ell^-\bar{\nu}_\ell$ Zerfalls mit Hadronischen Tags am Belle Experiment

selbstständig verfasst zu haben und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet zu haben.

Ort, Datum

Kilian Lieret

[Commit (None), compiled at Monday 23rd August, 2021 18:07]