Calibration of Machine Learning Based Hadronic Tagging in Preparation for a $|V_{cb}|$ Measurement and Clustering of Kinematic Distributions

Kilian Adriano Lieret

München 2022

Calibration of Machine Learning Based Hadronic Tagging in Preparation for a $|V_{cb}|$ Measurement and Clustering of Kinematic Distributions

Kilian Adriano Lieret

Dissertation
an der Fakultät für Physik
der Ludwig-Maximilians-Universität
München

vorgelegt von
Kilian Adriano Lieret
aus München

München, den 8. April 2022

Kalibration eines Machine-Learning-Algorithmus für hadronisches Tagging zur Vorbereitung einer $|V_{cb}|$-Messung und Clustering von kinematischen Verteilungen

Kilian Adriano Lieret

Dissertation
an der Fakultät für Physik
der Ludwig-Maximilians-Universität
München

vorgelegt von
Kilian Adriano Lieret
aus München

München, den 8. April 2022

Erstgutachter: Prof. Dr. Thomas Kuhr
Zweitgutachter: Prof. Dr. Florian Bernlochner
Tag der mündlichen Prüfung: 27. Mai 2022

Zusammenfassung

Messungen der Verhältnisse $\mathcal{R}(\mathrm{D}^{(*)})$ der Zerfallsraten von $\mathrm{B} \to \mathrm{D}^{(*)}\tau\nu_\tau$ und $\mathrm{B} \to \mathrm{D}^{(*)}\ell\nu_\ell$ ($\ell = \mathrm{e}, \mu$) zeigen erhebliche Abweichungen zu den Vorhersagen des Standardmodells der Teilchenphysik. Die vorliegende Arbeit präsentiert verschiedene Ergebnisse, die zu einem verbesserten Verständnis dieser Anomalien beitragen können.

Die Abweichungen können als Effekte von Physik jenseits des Standardmodells interpretiert werden. Der Parameterraum, der diese Effekte beschreibt, ist aufgrund seiner Größe jedoch eine Herausforderung für theoretische wie experimentelle Studien. Clusteranalysen können diesen Parameterraum in Teilmengen mit ähnlichen Zerfallskinematiken unterteilen, um entsprechende Untersuchungen zu vereinfachen. Ich stelle ein Open-Source-Softwarepaket vor, das derartige Methoden implementiert und wende es auf den $\mathrm{B} \to \mathrm{D}^{(*)}\tau\nu_\tau$-Zerfall an, der für die $\mathcal{R}(\mathrm{D}^{(*)})$-Anomalien verantwortlich gemacht wird.

Weiterhin zeige ich Vorbereitungen für eine Messung des Zerfalls $\mathrm{e}^-\mathrm{e}^+ \to \Upsilon(4\mathrm{S}) \to \mathrm{B}_{\mathrm{tag}}(\to \text{Hadronen})\,\mathrm{B}_{\mathrm{sig}}(\to \mathrm{D}^*\ell\nu)$ mit dem Datensatz des Belle-Experiments. Neben seiner Bedeutung als Normalisierungskanal der $\mathcal{R}(\mathrm{D}^{(*)})$-Observable erlaubt dieser Prozess die Messung des CKM-Matrix-Elements $|V_{\mathrm{cb}}|$ und die Extrapolation hadronischer Formfaktoren.

Eine der entscheidenden Verbesserungen dieser Messung gegenüber Vorgängerstudien ist die Verwendung der Full Event Interpretation (FEI). Dieser Machine-Learning-Algorithmus rekonstruiert $\mathrm{B}_{\mathrm{tag}}$-Mesonen in mehreren tausend Zerfallskanälen. Allerdings führen Ungenauigkeiten in der für die Analyse essenziellen Monte-Carlo-Simulation zu Effizienzunterschieden der FEI zwischen simulierten und gemessenen Daten.

Um diese Effizienzunterschiede zu korrigieren, rekonstruiere ich den Zerfall $\mathrm{e}^-\mathrm{e}^+ \to \Upsilon(4\mathrm{S}) \to \mathrm{B}_{\mathrm{tag}}(\to \text{Hadronen})\,\mathrm{B}_{\mathrm{sig}}(\to X\ell\nu)$. Unter der Annahme, dass die Rekonstruktionseffizienzen von $\mathrm{B} \to X\ell\nu_\ell$ ausreichend verstanden sind, lassen sich die Effizienzunterschiede zwischen Daten und Monte-Carlo-Simulation der FEI zuordnen, sodass eine Kalibrierung vorgenommen werden kann. Neben der Korrektur der Gesamteffizienz beeinflusst die Kalibrierung auch wichtige Observablen für die Signalextraktion in $|V_{\mathrm{cb}}|$- und $\mathcal{R}(\mathrm{D}^{(*)})$-Analysen.

Zur Validierung der Kalibrierung betrachte ich $\mathrm{B} \to \mathrm{D}\ell\nu_\ell$-Zerfälle und führe weitere Studien zur Bestätigung zentraler Annahmen der Kalibrierungsmethodik durch. Es zeigt sich, dass sich die Kalibrierung für $\mathrm{B}_{\mathrm{tag}}$-Mesonen mit korrekt und inkorrekt rekonstruiertem flavor wesentlich unterscheidet, was in früheren Analysen nicht berücksichtigt wurde. Für eine Korrektur dieses Effekts präsentiere ich mehrere Strategien, die ich erfolgreich auf einen vorläufigen Datensatz anwende. Aufgrund der zentralen Rolle der FEI für eine Vielzahl von Studien am Belle- und Belle II-Experiment sind diese Resultate von weitreichender Bedeutung.

Der Erfolg der FEI verdeutlicht auch die Bedeutung neuester Softwaretechniken für moderne Messungen. Um das volle Potenzial der gemessenen Daten auszuschöpfen, investieren große Kollaborationen zunehmend in die Softwarekenntnisse ihrer Mitglieder. Ich zeige von mir koordinierte Aktivitäten am Belle II-Experiment und in der High Energy Physics Software Foundation (HSF).

Abstract

Measurements of the decay rate ratios $\mathcal{R}(\mathrm{D}^{(*)})$ of $\mathrm{B} \to \mathrm{D}^{(*)}\tau\nu_\tau$ and $\mathrm{B} \to \mathrm{D}^{(*)}\ell\nu_\ell$ decays ($\ell = \mathrm{e}, \mu$) show a substantial tension with the predictions of the Standard Model of particle physics. This thesis explores multiple topics that can facilitate a better understanding of these so-called flavor anomalies.

These flavor anomalies may point to contributions of physics beyond the Standard Model, however the large, unexplored parameter space defining such contributions makes exploratory physics studies challenging. Clustering algorithms can divide this parameter space into subsets featuring similar decay kinematics, which helps to simplify studies. I present an open-source software package that makes such techniques accessible and demonstrate its application on the $\mathrm{B} \to \mathrm{D}^{(*)}\tau\nu_\tau$ decay, which is speculated to be responsible for the $\mathcal{R}(\mathrm{D}^{(*)})$ anomalies.

Furthermore, I show preparations for a study of the decay $\mathrm{e}^-\mathrm{e}^+ \to \Upsilon(4\mathrm{S}) \to \mathrm{B}_{\mathrm{tag}}(\to \text{hadrons})\,\mathrm{B}_{\mathrm{sig}}(\to \mathrm{D}^*\ell\nu)$ with the dataset of the Belle experiment. Besides being of interest as the normalization channel of the $\mathcal{R}(\mathrm{D}^*)$ ratio, this decay also allows for the measurement of the CKM matrix element $|V_{\mathrm{cb}}|$ and for fits to hadronic form factors.

One of the significant improvements of this study over previous studies at Belle is the use of the Full Event Interpretation (FEI), a machine learning algorithm that is able to reconstruct several thousand possible Btag decays. However, this algorithm is sensitive to inaccuracies in the modeling of the Monte Carlo simulation used throughout the analysis, leading to different efficiencies on simulated and recorded data. To correct these efficiency differences, I reconstruct the decay $\mathrm{e}^-\mathrm{e}^+ \to \Upsilon(4\mathrm{S}) \to \mathrm{B}_{\mathrm{tag}}(\to \text{hadrons})\,\mathrm{B}_{\mathrm{sig}}(\to X\ell\nu)$. Assuming that the reconstruction efficiencies of the $\mathrm{B} \to X\ell\nu_\ell$ decay are well-understood, any efficiency difference between data and Monte Carlo simulation can be attributed to the FEI and hence used for its calibration. Along with its effect on the overall reconstruction efficiency, the calibration also affects important observables used for background subtraction in $|V_{\mathrm{cb}}|$ and $\mathcal{R}(\mathrm{D}^{(*)})$ analyses.

I validate the calibration with a sample of $\mathrm{B} \to \mathrm{D}\ell\nu_\ell$ decays and perform additional studies to confirm the validity of core assumptions of the calibration procedure. The calibration of Btag mesons with correctly and incorrectly reconstructed flavors are found to differ, which was not accounted for in previous analyses. I successfully explore several correction strategies on a preliminary dataset. As the FEI is used heavily at both Belle and Belle II, this result has significant implications for many past and upcoming analyses.

The success of the FEI also highlights the importance of using state-of-the-art software technologies in modern measurements. To deliver the best possible science, large experimental collaborations are increasingly focusing on software education. I present recent developments in the software training activities that I have coordinated at Belle II and at the High Energy Physics Software Foundation (HSF).

Conventions

Unless otherwise declared, ℓ denotes either electrons or muons. Natural units are used throughout this thesis. A list of frequently used acronyms is given on page 231.

Contents

Chapter 1

Introduction

This chapter provides context for the work presented in this thesis. First, the Standard Model of particle physics is introduced, focusing on the electroweak sector and $|V_{\mathrm{cb}}|$. Signs for physics beyond the Standard Model, such as the tension in the $\mathcal{R}(\mathrm{D}^{(*)})$ observable, are discussed. Then, the experimental technique of B-tagging and the Full Event Interpretation (FEI) algorithm are described. Finally, measurement strategies of $|V_{\mathrm{cb}}|$ are compared, and the kinematics of the $\mathrm{B} \to \mathrm{D}^{*} \ell \nu_{\ell}$ decay are introduced.

1.1 The Standard Model of Particle Physics

This section describes the Standard Model of Particle Physics (SM), focusing on the electroweak symmetry breaking and the CKM matrix V_{CKM}. An extensive (and famous) description of the SM is given by Weinberg [1, 2]. Rather than following the historical development, a top-level view emphasizing the mathematical structure of the final construct is chosen here.[1] A very succinct explanation of the SM with a focus on the electroweak sector is [4].

1.1.1 The symmetries of the SM

The SM is a renormalizable gauge quantum field theory built around the symmetry group

$$\mathsf{SU}(3)_C \times \mathsf{SU}(2)_W \times \mathsf{U}(1)_Y, \tag{1.1.1}$$

where $\mathsf{SU}(\mathsf{n})$ is the special unitary Lie group of degree n, that is,

$$\mathsf{SU}(\mathsf{n}) := \{U \in \mathbb{C}^n \mid U^\dagger U = U U^\dagger = \mathbb{1},\, \det U = 1\}, \tag{1.1.2}$$

and $\mathsf{U}(1)$ is the unit circle, $\mathsf{U}(1) := \{z \in \mathbb{C} \mid |z| = 1\}$. The three symmetry groups correspond to particles of spin one (*gauge bosons*[2]), which mediate interactions between the matter particles of the SM. The dimensionality of the Lie algebras $\mathfrak{su}(n)$ generating the Lie groups corresponds to the number of degrees of freedom of the gauge bosons:

- Eight degrees of freedom of a gauge field G ($\dim \mathfrak{su}(3) = 8$). These particles are called *gluons* and mediate the strong interaction.
- Three degrees of freedom of a gauge field W ($\dim \mathfrak{su}(2) = 3$).
- One degree of freedom of a gauge field B.

[1] A summary of the historical development with a similar focus on the "big picture" is found in [3].
[2] Generally, *boson* refers to particles with integer spin.

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 $\mathsf{SU}(3)_C \times \mathsf{SU}(2)_W \times \mathsf{U}(1)_Y$. S denotes the spin of the particles. Antiparticles and the three generations of fermions are not listed separately.

The remaining particles of the SM can now be characterized by their interaction with the gauge fields, that is, by their transformation properties under their respective symmetry groups. For this, we use the shorthand notation $(\mathbf{m},\mathbf{n})_Y$ to designate a particle transforming as a m-tuple under $\mathsf{SU}(3)$, an n-tuple under $\mathsf{SU}(2)$ and carrying $\mathsf{U}(1)$ hypercharge[3] Y. The particles introduced in the following paragraphs are shown in Table 1.1.

The matter particles of the SM have spin 1/2 (*fermions*[4]) and are separated into *leptons* and *quarks*. Leptons are $\mathsf{SU}(3)$ singlets (that is, not subject to the strong interaction), whereas quarks are $\mathsf{SU}(3)$ triplets (subject to the strong interaction). Similarly, leptons can be divided based on their transformations under $\mathsf{SU}(2)$. This distinction coincides with a property called *chirality*: *Left-handed* fermions are $\mathsf{SU}(2)$ doublets (interact with W), *right-handed* fermions are $\mathsf{SU}(2)$ singlets (do not interact with W).

The two $\mathsf{SU}(2)$ degrees of freedom of the left-handed quarks are called *up* and *down*-type quarks (denoted q'_u and q'_d in the following). For leptons, we have *charged leptons* and *neutrinos* (denoted l and ν). Up and down-type quarks and charged leptons have right-handed partners, but no right-handed neutrino is included in the SM.[5]

To summarize, the fermions of the SM are

$$L'_\mathrm{L} = \begin{pmatrix} l'_\mathrm{L} \\ \nu_\mathrm{L} \end{pmatrix},\ Q'_\mathrm{L} = \begin{pmatrix} q'_\mathrm{uL} \\ q'_\mathrm{dL} \end{pmatrix},\ l'_\mathrm{R},\ q'_\mathrm{uR},\ q'_\mathrm{dR}. \tag{1.1.3}$$

For each of these particles, an *antiparticle* with opposite quantum numbers and opposite chirality exists. Furthermore, the fermions come in three *generations*:

$$\begin{bmatrix} l' & q'_\mathrm{u} \\ \nu' & q'_\mathrm{d} \end{bmatrix} = \begin{bmatrix} \nu'_e & \mathrm{u}' \\ e' & \mathrm{d}' \end{bmatrix}, \begin{bmatrix} \nu'_\mu & \mathrm{c}' \\ \mu' & \mathrm{s}' \end{bmatrix}, \begin{bmatrix} \nu'_\tau & \mathrm{t}' \\ \tau' & \mathrm{b}' \end{bmatrix}. \tag{1.1.4}$$

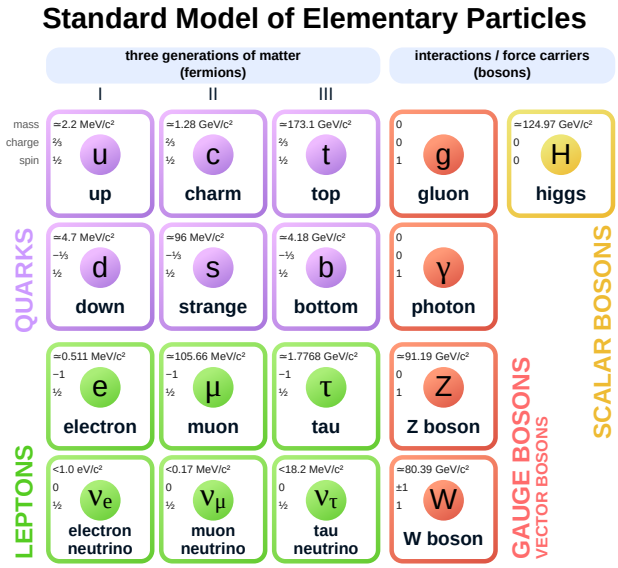
The corresponding particle names are shown in Figure 1.1a. The six different species of quarks and leptons are called the *flavors* of quarks and leptons.

Two degrees of freedom of W mix into an electrically charged gauge boson and its antiparticle. The corresponding interaction (*charged current*) connects up with down-type quarks of the same

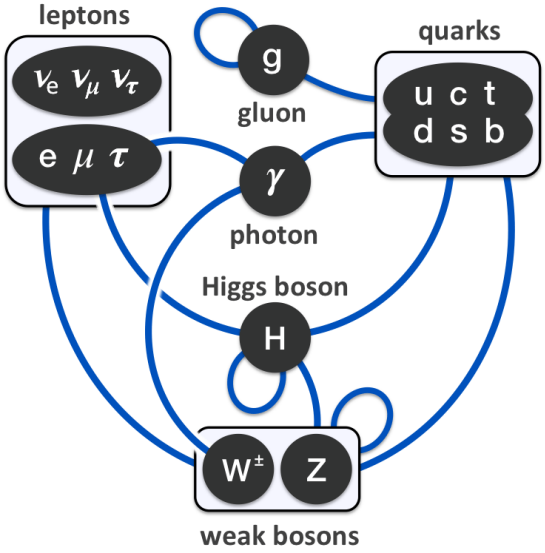
[3] Different conventions for Y exist. Here, $Y = Q - T_3$ is used, where Q is the electrical charge and T_3 is the third component of the weak isospin. Corresponding values can be found in [5][6, pp. 704, 713][4, p. 16].

[4] Generally, *fermion* refers to any particle of half-integer spin. However, we will use the word more narrowly to only refer to the SM fermions.

[5] However, many extensions of the SM include right-handed neutrinos.



(a) Particles with masses and quantum numbers. Figure from [7].



(b) Interactions of the SM particles. Blue lines connect particles with bosons with which they can interact. "Closed loops" indicate self-interactions. Figure from [8].

Figure 1.1: The Standard Model of Particle physics after electroweak symmetry breaking

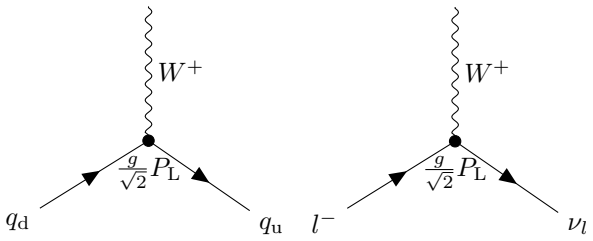


Figure 1.2: Charged current. The operator $P_\mathrm{L} := (1-\gamma^5)/2$ with $\gamma^5 := \mathrm{i}\gamma^0\gamma^1\gamma^2\gamma^3$ projects the fermion on its left-handed part.

generation and charged leptons with neutrinos of the same generation:

$$\mathcal{L}_\mathrm{CC} = \frac{g}{\sqrt{2}}\left[W_\mu^\dagger\left(\sum_{\substack{(q_\mathrm{u},q_\mathrm{d})=\\(\mathrm{u,d}),(\mathrm{c,s}),(\mathrm{t,b})}} \overline{q'}_\mathrm{uL}\gamma^\mu q'_\mathrm{dL} + \sum_{l=\mathrm{e},\mu,\tau} \overline{\nu'}_{l\mathrm{L}}\gamma^\mu l'_\mathrm{L}\right) + \mathrm{h.c.}\right]. \tag{1.1.5}$$

Here, the fermions are represented by spinors (4 vectors), and γ^μ are Dirac matrices. $\gamma^\mu \in \mathbb{C}^{4\times4}$ and $W_\mu \in \mathbb{C}^4$ contract via the Minkowski metric (Einstein notation). The chirality subscript is meant to precede the conjugation: $\overline{\psi}_\mathrm{L} := \overline{(\psi_\mathrm{L})}$. The Feynman diagram corresponding to this interaction is shown in Figure 1.2.

1.1.2 Electroweak symmetry breaking

The $\mathsf{SU}(2)_W \times \mathsf{U}(1)_Y$ symmetry does not allow mass terms for the fermions and the gauge bosons. A solution to this problem was proposed almost simultaneously by three independent groups in 1964 [9–11]: Instead of directly introducing mass terms, another particle is added to the theory. This *Higgs particle* H is a scalar[6] $(\mathbf{1},\mathbf{2})_{1/2}$ boson with gauge invariant couplings to both fermions and bosons. However, the potential of H is chosen in such a way that, while the potential itself and the set of degenerate ground states conserve the $\mathsf{SU}(2)_W \times \mathsf{U}(1)_Y$ symmetry, any particular ground state ϕ^0 of H breaks it (*spontaneous electroweak symmetry breaking*).

[6] i.e., has zero spin, even parity.

Three real degrees of freedom (*Nambu-Goldstone Bosons*[7]) of H can be absorbed using a $\mathsf{SU}(2)_W$ gauge transformation, leaving only the electrically neutral ground state ϕ^0. Expanding $\phi^0 = (v+H)/\sqrt{2}$ around the vacuum expectation value $v/\sqrt{2} := |\langle 0|\phi^{(0)}|0\rangle|$, the following term appears in the Lagrangian [5, pp. 46–48]:

$$v^2 \begin{pmatrix} W^1_\mu & W^2_\mu & W^3_\mu & B_\mu \end{pmatrix} \begin{pmatrix} g^2 & 0 & 0 & 0 \\ 0 & g^2 & 0 & 0 \\ 0 & 0 & g^2 & gg' \\ 0 & 0 & gg' & g'^2 \end{pmatrix} \begin{pmatrix} W^{1\mu} \\ W^{2\mu} \\ W^{3\mu} \\ B^\mu \end{pmatrix} \tag{1.1.6}$$

with constants $g, g' \in \mathbb{R}$. We can immediately identify mass terms for the first two degrees of freedom of W (the charged W boson and its antiparticle). Diagonalizing the matrix and thereby mixing W^3 and B, we can define two more mass eigenstates Z^0 and A that we can identify with the Z^0 boson and the photon.

The Lagrangian describing the coupling of the fermions to the Higgs field (*Yukawa coupling*[8]) can be written as[9]

$$\mathcal{L}_Y = -\left(1 + \frac{H}{v}\right)\left(\overline{\mathbf{d}}'_\mathrm{L}\mathbf{M}'_\mathrm{d}\mathbf{d}'_\mathrm{R} + \overline{\mathbf{u}}'_\mathrm{L}\mathbf{M}'_\mathrm{u}\mathbf{u}'_\mathrm{R} + \overline{\mathbf{l}}'_\mathrm{L}\mathbf{M}'_\mathrm{l}\mathbf{l}'_\mathrm{R} + \mathrm{h.c.}\right). \tag{1.1.7}$$

Here, $\mathbf{d}'$, $\mathbf{u}'$, and $\mathbf{l}$ denote the down-type, up-type, and charged lepton fermionic fields as 3-vectors of spinors for the three generations. For example, $\mathbf{u}'_\mathrm{L}$ denotes $(\,u'_\mathrm{L}\ c'_\mathrm{L}\ t'_\mathrm{L}\,)^T$ (u'_L, c'_L, $t'_\mathrm{L} \in \mathbb{C}^4$). $\overline{\mathbf{d}}'_\mathrm{L}$, $\overline{\mathbf{u}}'_\mathrm{L}$, $\overline{\mathbf{l}}'_\mathrm{L}$ denote the conjugate spinors, e.g.,

$$\overline{\mathbf{u}}'_\mathrm{L} := (\mathbf{u}'_\mathrm{L})^\dagger\gamma^0 = \begin{pmatrix} (u'_\mathrm{L})^\dagger\gamma^0 \\ (c'_\mathrm{L})^\dagger\gamma^0 \\ (t'_\mathrm{L})^\dagger\gamma^0 \end{pmatrix}, \tag{1.1.8}$$

where $\gamma^0 \in \mathbb{R}^{4\times4}$ is the zeroth Dirac gamma matrix. Finally, $\mathbf{M}'_\mathrm{d}$, $\mathbf{M}'_\mathrm{u}$ and $\mathbf{M}'_\mathrm{l}$ are $\mathbb{C}^{4\times4}$ matrices that parameterize the coupling. In the following, we will use the subscript f to denote one of d, u, l. Using the polar decomposition [15, pp. 42–45], we can write $\mathbf{M}'_f = \mathbf{H}_f\mathbf{U}_f$, where $\mathbf{U}_f$ is a unitary matrix and $\mathbf{H}_f$ is a positive semi-definite Hermitian matrix.[10] By the spectral theorem, $\mathbf{H}_f$ can be diagonalized as $\mathbf{H}_f = \mathbf{S}_f^\dagger\mathcal{M}_f\mathbf{S}_f$ with $\mathcal{M}_f$ real and diagonal and $\mathbf{S}_f$ unitary. Let $\mathcal{M}_\mathrm{u} =: \mathrm{diag}(m_\mathrm{u}, m_\mathrm{c}, m_\mathrm{t})$ and correspondingly for $\mathcal{M}_\mathrm{d}$ and $\mathcal{M}_\mathrm{l}$. Taken together, we have

$$\mathbf{M}'_f = \mathbf{S}_f^\dagger\mathcal{M}_f\mathbf{S}_f\mathbf{U}_f. \tag{1.1.9}$$

Defining

$$\mathbf{f}_\mathrm{L} := \mathbf{S}_f\mathbf{f}'_\mathrm{L} \quad \text{and} \quad \mathbf{f}_\mathrm{R} := \mathbf{S}_f\mathbf{U}_f\mathbf{f}'_\mathrm{R}, \tag{1.1.10}$$

and substituting Equation (1.1.9) into Equation (1.1.7), we get

$$\begin{aligned} \mathcal{L}_Y &= -\left(1 + \frac{H}{v}\right)\left(\overline{\mathbf{d}}_\mathrm{L}\mathcal{M}_\mathrm{d}\mathbf{d}_\mathrm{R} + \overline{\mathbf{u}}_\mathrm{L}\mathcal{M}_\mathrm{u}\mathbf{u}_\mathrm{R} + \overline{\mathbf{l}}_\mathrm{L}\mathcal{M}_\mathrm{l}\mathbf{l}_\mathrm{R} + \mathrm{h.c.}\right) = \\ &= -\left(1 + \frac{H}{v}\right)\left(\sum_{f=\mathrm{d,s,b}} m_f\overline{\mathbf{f}}\mathbf{f} + \sum_{f=\mathrm{u,c,t}} m_f\overline{\mathbf{f}}\mathbf{f} + \sum_{f=\mathrm{e},\mu,\tau} m_f\overline{\mathbf{f}}\mathbf{f}\right). \end{aligned} \tag{1.1.11}$$

[7] The mechanism of spontaneous symmetry breaking and its connection to the manifestation of bosons (Nambu-Goldstone Bosons) is a more general result predating the 1964 discovery. The original findings were in the context of condensed matter physics [12, 13] before being generalized in quantum field theory [14]. The 1964 achievement of the groups mentioned above was to formulate a *relativistic* field theory that correctly included *massive* Nambu-Goldstone Bosons and properties compatible with the experimental results.

[8] Generally, Yukawa couplings are couplings of the form $-g\overline{\psi}\phi\psi$ between a fermionic field ψ and a scalar field ϕ or $-\mathrm{i}g\overline{\psi}\gamma^5\phi\psi$ for a pseudoscalar (zero spin, odd parity) ϕ.

[9] This derivation of V_CKM follows [4, pp. 29–30].

[10] The polar decomposition is unique for invertible matrices. Jumping ahead, we see that $\det\mathbf{M}'_f \overset{(1.1.9)}{=} \det\mathcal{M}_f = \prod_i \mathcal{M}_{fii}$ because of unitarity. Because of the identification of $\mathcal{M}_{fii}$ with the fermionic masses in (1.1.11), indeed $\det\mathbf{M}'_f > 0$.

We can see that $\mathbf{f}$ are the mass eigenstates with masses m_f!

Can we simply redefine the fermions from Equation (1.1.4) to their generation-rotated mass-eigenstates given by Equation (1.1.10) while keeping the structure of the SM Lagrangian? Unfortunately not. The problem is the charged current introduced in Equation (1.1.5): Because it connects up- and down-type quarks, each having distinct rotations $\mathbf{S}_f$, Equation (1.1.5) takes the form

$$\mathcal{L}_{\mathrm{CC}} = \frac{g}{\sqrt{2}}\Big[W_\mu^\dagger(\overline{\mathbf{u}}_{\mathrm{L}}\mathbf{S}_u\gamma^\mu\mathbf{S}_d^\dagger\mathbf{d}_{\mathrm{L}} + \overline{\boldsymbol{\nu}}'_{l\mathrm{L}}\gamma^\mu\mathbf{S}_l^\dagger\mathbf{l}_{\mathrm{L}}) + \mathrm{h.c.}\Big]. \tag{1.1.12}$$

The charged current connects different mass eigenstate generations! This mixing of generation is parameterized by the matrix

$$V_{\mathrm{CKM}} \coloneqq \mathbf{S}_{\mathrm{u}}\mathbf{S}_{\mathrm{d}}^\dagger. \tag{1.1.13}$$

V_{CKM} is called the *Cabibbo-Kobayashi-Maskawa (CKM) matrix* [16, 17] and will be described in more detail in the following section.

The lepton sector of the SM is less complex: Because no right-handed neutrino is included, there are no neutrino mass terms in Equation (1.1.7). Therefore, we are not bound by Equation (1.1.10) and can instead redefine neutrinos as $\boldsymbol{\nu} \coloneqq \mathbf{S}_{\mathrm{l}}\boldsymbol{\nu}'$. This eliminates cross-generation couplings in the lepton sector.

However, the existence of neutrino oscillations, predicted by Pontecorvo [18] and observed by the Super-Kamiokande Observatory [19] and the Sudbury Neutrino Observatories [20] shows that neutrinos are, in fact, not massless particles.[11] While the mechanism by which mass terms are included in the Lagrangian remains unclear, this means that lepton generations do indeed mix. The effects are parameterized by the *Pontecorvo-Maki-Nakagawa-Sakata (PMNS)* matrix [22] in a similar way to the CKM matrix.

1.1.3 The CKM Matrix

Equation (1.1.13) shows that V_{CKM} is a unitary 3×3 matrix. Unitarity of a matrix $V \in \mathbb{C}^{3\times 3}$ is equivalent to an orthonormality requirement on its rows[12], that is,

$$\delta_{ij} \overset{!}{=} (VV^\dagger)_{ij} = \sum_k V_{ik}V_{jk}^* \tag{1.1.14}$$

for $i \geq j$ $(i, j = 1, 2, 3)$. The conditions $1 = (VV^\dagger)_{ii}$ constrain three real degrees of freedom (note that $(VV^\dagger)_{ii}$ is real by construction), the remaining conditions constrain three complex (six real) degrees of freedom. Therefore, only nine real degrees of freedom remain for a unitary matrix. In the case of V_{CKM}, we can furthermore absorb five complex phases into the quark fields [6, p. 724], leaving only four real degrees of freedom.

The orthogonality constraints $(i \neq j)$ on V_{CKM} are often visualized as a triangle in the complex plane (*unitarity triangle*): Most commonly, the combination $i = 1, j = 3$ is considered:[13]

$$V_{\mathrm{ud}}V_{\mathrm{ub}}^* + V_{\mathrm{cd}}V_{\mathrm{cb}}^* + V_{\mathrm{td}}V_{\mathrm{tb}}^* = 0. \tag{1.1.15}$$

Normalizing by the term with the lowest experimental uncertainties, this is equivalent to

$$\frac{V_{\mathrm{ud}}V_{\mathrm{ub}}^*}{V_{\mathrm{cd}}V_{\mathrm{cb}}^*} + 1 + \frac{V_{\mathrm{td}}V_{\mathrm{tb}}^*}{V_{\mathrm{cd}}V_{\mathrm{cb}}^*} = 0. \tag{1.1.16}$$

[11] For a more recent review of neutrino oscillations, see [21].

[12] This is equivalent to requiring orthonormality of the columns with similar conditions as Equation (1.1.14) and arriving at alternative unitarity triangles [23].

[13] The $i = 1, j = 3$ triangle has the nice property that all sides are of the same orders in λ and A (see the following paragraphs) so that all sides of the normalized triangle (Equation (1.1.16)) have a length of order one. This implies a large $\bar{\eta}$, that is, large $\mathcal{CP}$ violation in this sector. [23, 24].

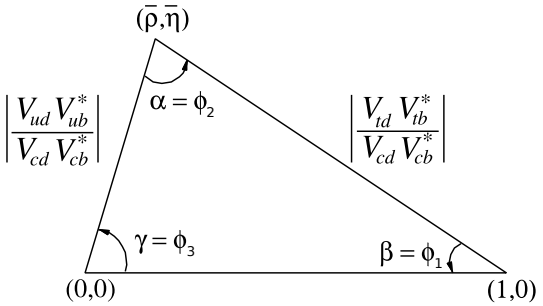


Figure 1.3: The unitarity triangle. Figure from [25].

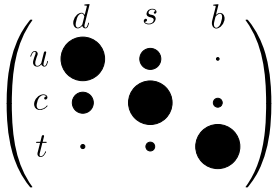


Figure 1.4: Hierarchy of the absolute values in V_{CKM} according to Equation (1.1.21).

This unitarity triangle is sketched in Figure 1.3; measurements constraining the apex are shown in Figure 1.5.

Different parameterizations of V_{CKM} exist to make the degrees of freedom apparent. The particle data group recommends [25] the parameterization by Chau and Keung [26]:

$$\begin{aligned} V_{\text{CKM}} &= \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} = \\ &= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \end{aligned} \tag{1.1.17}$$

where $s_{ij} := \sin\theta_{ij}$, $c_{ij} := \cos\theta_{ij}$ $(i < j = 1, 2, 3)$ and s_{ij}, c_{ij} can be assumed to be positive without loss of generality.

As a complex phase, δ introduces $\mathcal{CP}$ violation into the SM. In fact, it was the evidence of $\mathcal{CP}$ violation [27] that prompted Kobayashi and Maskawa to propose a third generation of fermions and to generalize the 2×2 Cabibbo matrix [16] to the CKM matrix [17]: In the Cabibbo matrix, no $\mathcal{CP}$ violating phase is possible [17].

The experimentally established hierarchical structure of $|V_{\text{CKM}}|$ (Figure 1.4) is made apparent in the parameterization by Wolfenstein [28]:

$$V_{\text{CKM}} = \begin{pmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}\left(\lambda^4\right). \tag{1.1.18}$$

This is an approximate parameterization based on an expansion in $\lambda = |V_{\text{us}}| \approx 0.22$. In particular, it should be noted that the matrix is not strictly unitary without including the $\mathcal{O}(\lambda^4)$ terms. However, an exact version of Equation (1.1.18) can be obtained by defining λ, A, $\bar{\rho}$ and $\bar{\eta}$ via $s_{12} \stackrel{!}{=} \lambda$, $s_{23} \stackrel{!}{=} A\lambda^2$ and $s_{13}e^{-i\delta} \stackrel{!}{=} A\lambda^3(\bar{\rho} - i\bar{\eta})$ and using Equation (1.1.18). This approach is

presented in [29] and adopted by the CKMFitter group [24]. Comparing this to the classical Wolfenstein parameterization $\bar{\rho}$ and $\bar{\eta}$ are now power series in λ: $\bar{\rho} = \rho(1 - \lambda^2/2 + \cdots)$ and $\bar{\eta} = \eta(1 - \lambda^2/2 + \cdots)$.

With these definitions, the apex of the unitarity triangle as presented in Figure 1.3 takes the convenient form of $\bar{\rho} + \mathrm{i}\bar{\eta}$ [24, p. 12]:

$$\bar{\rho} + \mathrm{i}\bar{\eta} = -\frac{V_{\mathrm{ud}}V_{\mathrm{ub}}^*}{V_{\mathrm{cd}}V_{\mathrm{cb}}^*}. \tag{1.1.19}$$

Global fits to all measurements constraining quark mixing give [25]

$$\begin{aligned} \lambda &= 0.22650 \pm 0.00048, \quad & A &= 0.790^{+0.017}_{-0.012}, \\ \bar{\rho} &= 0.141^{+0.016}_{-0.017}, & \bar{\eta} &= 0.357 \pm 0.011, \end{aligned} \tag{1.1.20}$$

resulting in [25]

$$|V_{\mathrm{CKM}}| = \begin{pmatrix} 0.97401 \pm 0.00011 & 0.22650 \pm 0.00048 & 0.00361\ ^{+0.00011}_{-0.00009} \\ 0.22636 \pm 0.00048 & 0.97320 \pm 0.00011 & 0.04053\ ^{+0.00083}_{-0.00061} \\ 0.00854\ ^{+0.00023}_{-0.00016} & 0.03978\ ^{+0.00082}_{-0.00060} & 0.999172^{+0.000024}_{-0.000035} \end{pmatrix}. \tag{1.1.21}$$

These values are also visualized in Figure 1.4. Another value that is oftentimes quoted is the *Jarlskog invariant* $\mathcal{J}$ [30], defined via

$$\mathrm{Im}[V_{ij}V_{ik}^*V_{lk}V_{lj}^*] = \mathcal{J} \sum_{m,n=1}^{3} \epsilon_{ilm}\epsilon_{jkn}, \tag{1.1.22}$$

where ϵ is the totally asymmetric tensor and $1 \leq i, j \leq 3$. Its significance stems from the fact that all $\mathcal{CP}$-violating amplitudes are proportional to $\mathcal{J}$. For this reason, it is also taken as a measure of the total amount of $\mathcal{CP}$-violation occurring in the SM. Furthermore, it is twice the area of the unitarity triangles given analog to Equation (1.1.15) (without normalization). The current value for $\mathcal{J}$ is [25]

$$\mathcal{J} = (3.00^{+0.15}_{-0.09}) \times 10^{-5}. \tag{1.1.23}$$

1.1.4 Measurements of the CKM Matrix and overconstraining the triangle

While the number of degrees of freedom of the CKM matrix V_{CKM} is only four, this four-dimensional parameter space can be constrained by measurements sensitive to various combinations of V_{CKM} elements. In addition to quadratic combinations of CKM elements, the angles of the unitarity triangle (Figure 1.3) are prominent examples [23]:

$$\begin{aligned} \alpha := \phi_2 &:= \arg\left(-\frac{V_{\mathrm{td}}V_{\mathrm{tb}}^*}{V_{\mathrm{ud}}V_{\mathrm{ub}}^*}\right) = \arg\left(-\frac{1 - \bar{\rho} - \mathrm{i}\bar{\eta}}{\bar{\rho} + \mathrm{i}\bar{\eta}}\right), \\ \beta := \phi_1 &:= \arg\left(-\frac{V_{\mathrm{cd}}V_{\mathrm{cb}}^*}{V_{\mathrm{td}}V_{\mathrm{tb}}^*}\right) = \arg\left(\frac{1}{1 - \bar{\rho} - \mathrm{i}\bar{\eta}}\right), \\ \gamma := \phi_3 &:= \arg\left(-\frac{V_{\mathrm{ud}}V_{\mathrm{ub}}^*}{V_{\mathrm{cd}}V_{\mathrm{cb}}^*}\right) = \arg(\bar{\rho} + \mathrm{i}\bar{\eta}). \end{aligned} \tag{1.1.24}$$

In addition to these, a fourth angle related to one of the less prominent unitarity angles is defined as [23, p. 4]:

$$\beta_s := \arg\left(-\frac{V_{\mathrm{ts}}V_{\mathrm{tb}}^*}{V_{\mathrm{cs}}V_{\mathrm{cb}}^*}\right) = \lambda^2\bar{\eta} + \mathcal{O}\left(\lambda^4\right). \tag{1.1.25}$$

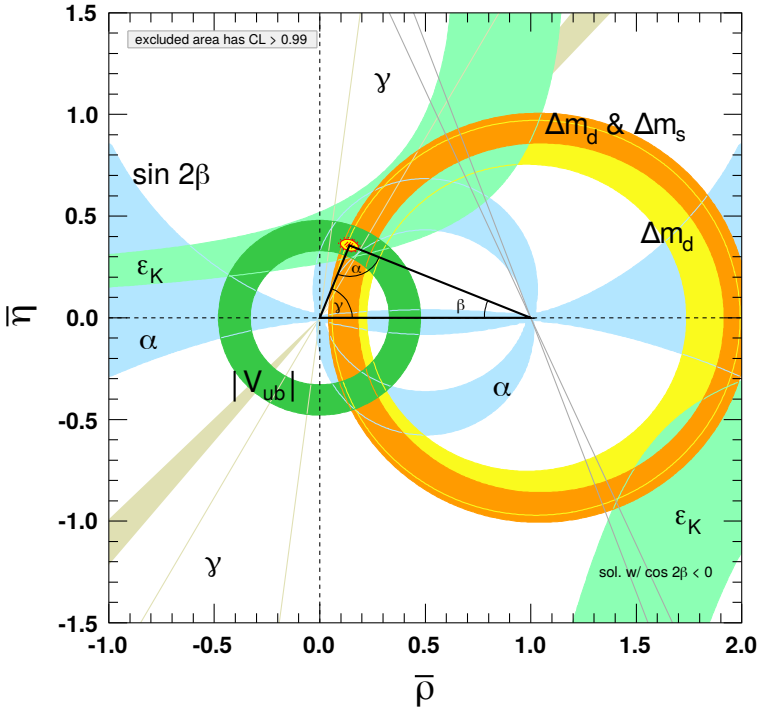


Figure 1.5: Constraining the apex of the unitarity triangle. Figure from [25].

A selection of the most sensitive measurement avenues is outlined in Table 1.2.

It is clear from the interpretation of α, β, γ in the unitarity triangle that $\alpha + \beta + \gamma = \pi$ (and this can also be seen from the expressions in $\bar{\rho}$ and $\bar{\eta}$). However, when expressing these angles in terms of V_{CKM} matrix elements (as in Equation (1.1.24)), this property only follows using the unitarity condition. This means that separate measurements of the three angles probe the unitarity of the SM!

Some of the redundancy in the measurements can be visualized as *overconstraining* the apex of the unitarity triangle. This is shown in Figure 1.5: Bands of different colors illustrate the constraints on the apex given by different measurements. For example, the dark green (orange) circle is a constraint on the length of the left (right) side of the triangle proportional to $|V_{\mathrm{ub}}|$ (Δm_d).[14] Similarly, the dark khaki (dark blue) bands represent the constraints of the measurements of γ (β).[15]

1.2 Shortcomings of the SM

The SM outlined in Section 1.1 has proven to be a remarkably successful theory. After being finalized in the mid-1970s, innumerable measurements have been carried out to constrain its free parameters, confirm its predictions, and challenge its validity. With the discovery of the Higgs particle in 2012 [31, 32], all predicted particles have been confirmed.

As mentioned in Section 1.1.2, an evident deficiency of the SM is the omission of neutrino masses. However, simple extensions of the SM to include neutrino mass terms can readily amend this. Depending on whether neutrinos are Majorana particles[16], a Majorana mass term or Dirac mass term can be added to the SM Lagrangian. Introducing these terms requires the inclusion of right-handed neutrinos as a new particle, but the overall structure of the SM remains essentially unchanged.

[14] The left side is also proportional to $|V_{\mathrm{ud}}|$. However, $|V_{\mathrm{ud}}|$ is measured to much higher precision, and thus it is $|V_{\mathrm{ub}}|$ that drives the uncertainty on this length.

[15] In the case of β, a total of four jets (two dark blue, two light gray) are shown in Figure 1.5. This is because the measurements of $\mathrm{B} \to (\mathrm{c\bar{c}})\mathrm{K}^{(*)}$ constrain $\sin(2\beta)$, resulting in a four-fold ambiguity.

[16] That is, they are their own antiparticles.

Constraint	Measurement	Theo. input	Notes
$\lvert V_{\mathrm{cb}}\rvert$	$\mathrm{B} \to \mathrm{D}^{(*)}\ell\nu_\ell$	$\mathrm{B} \to \mathrm{D}^{(*)}$ FFs	See Section 1.4.1
	$\mathrm{B} \to X\ell\nu_\ell$	OPE	See Section 1.4
$\lvert V_{\mathrm{ub}}\rvert$	$\mathrm{B} \to \pi\ell\nu_\ell$	$\mathrm{B} \to \pi$ FFs	
	$\mathrm{B} \to X_{\mathrm{u}}\ell\nu_\ell$	OPE	
$\lvert V_{\mathrm{ud}}\rvert$	$M \to \ell\nu_\ell$	Decay constant f_M	
	$M \to N\ell\nu_\ell$	$M \to N$ FF or $M \to N$ amplitude	
γ	$\mathrm{B} \to \mathrm{D}^{(*)}\mathrm{K}^{(*)}$		
$V_{\mathrm{ts}}V_{\mathrm{td}}^*$, $V_{\mathrm{cs}}V_{\mathrm{cd}}^*$	ϵ_{K} ($\mathrm{K}\overline{\mathrm{K}}^0$ mixing)	bag parameter B_{K}	
$\lvert V_{\mathrm{tb}}V_{\mathrm{td}}^*\rvert$	Δm_{d} ($\mathrm{B}^0\overline{\mathrm{B}}^0$ mixing)	bag parameter B_{B^0}	
$\lvert V_{\mathrm{tb}}V_{\mathrm{ts}}^*\rvert$	Δm_{s} ($\mathrm{B}_{\mathrm{s}}^0\overline{\mathrm{B}}_{\mathrm{s}}^0$ mixing)	bag parameter $B_{\mathrm{B}_{\mathrm{s}}^0}$	
β	$\mathrm{B} \to (\mathrm{c\bar{c}})\mathrm{K}^{(*)}$		
α	$\mathrm{B} \to \pi\pi, \rho\pi, \rho\rho$		
β_s	$\mathrm{B}_{\mathrm{s}}^0 \to \mathrm{J}/\psi\phi$		

Table 1.2: A selection of measurements constraining V_{CKM} [23, p. 5]. The following abbreviations are used: FF (hadronic form factor, introduced in Section 1.4.1), OPE (operator product expansion, see Section 1.4). An expanded version of this table with recent measurement values is shown in [23, p. 17].

Leaving this issue aside, we can still identify a series of theoretical shortcomings and experimental tensions that point to an incompleteness of the SM.

1.2.1 Theoretical problems

- The SM does not include a description of gravity. At large scales, the classical theory of general relativity has been very successful experimentally.[17] However, developing a quantized theory of gravitation is not straightforward: For example, general relativity is not renormalizable and features a more complex description of time and causality than quantum theories [33, p. 484]. The two most developed attempts at quantum gravity are string theory and loop quantum gravity [34].

- The theory of quantum chromodynamics (QCD) that describes the strong interaction in the SM permits $\mathcal{CP}$ violation. However, no such $\mathcal{CP}$ violation has been observed to this date. Explaining this absence in SM requires a particular choice, a *fine-tuning* in the QCD parameters, which can be considered *unnatural*[18] [36]. A popular solution to this issue is the *Peccei-Quinn mechanism* [37] in which a (pseudo-)Nambu-Goldstone boson from a spontaneously broken symmetry suppresses the $\mathcal{CP}$ violation in the QCD sector. The new boson is called an *Axion*. Besides their theoretical motivation, Axions (and axion-like particles, ALPs) are promising candidates for cold dark matter [25].

- Another issue of naturality arises when comparing the Higgs mass of $\approx 125\,\mathrm{GeV}$ to the gravitational scale of $\approx 1.22 \times 10^{19}\,\mathrm{GeV}$: Both scales are separated by seventeen orders of magnitude. This immense separation of scales can be perceived as surprising. Furthermore,

[17] Although we need a dark energy component to explain the accelerated expansion of the Universe.
[18] However, arguments of naturality can be criticized to be subjective [35].

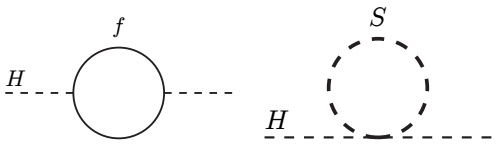


Figure 1.6: Corrections to the mass of the Higgs boson. Figures from [38].

(a) Fermion loop (b) Scalar loop

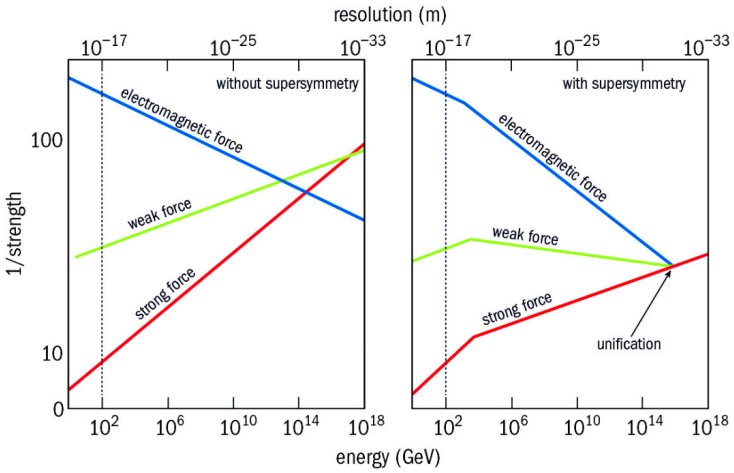


Figure 1.7: Running of coupling constants in the SM and in a potential grand unified theory (here on the example of SUSY). Figure adapted from [39].

the mass of the Higgs receives radiative corrections to its mass. Coupling terms of the form $\lambda_f H \bar{f} f$ (coupling to fermions f, Figure 1.6a) and $-\lambda_S |H|^2 |S|^2$ (coupling to scalars S, Figure 1.6b) lead to the following first-order corrections of the squared Higgs mass m_{H}^2 [38]:

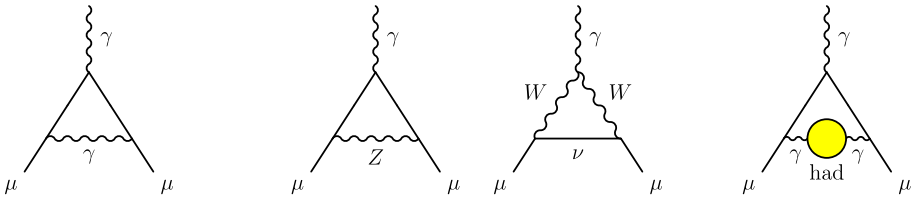
$$\Delta m_{\mathrm{H}}^2 = -\frac{|\lambda_f|^2}{8\pi^2}\Lambda_{\mathrm{UV}}^2 + \cdots \quad \text{resp.} \quad \Delta m_{\mathrm{H}}^2 = +\frac{\lambda_S}{16\pi^2}\Lambda_{\mathrm{UV}}^2 + \cdots, \tag{1.2.1}$$

where Λ_{UV} is the ultraviolet momentum cutoff used to regulate the loop integral. Identifying the cutoff with the presence of NP at this scale implies that Λ_{UV} must be relatively large. But then the correction terms of Equation (1.2.1) are many orders of magnitude larger than m_{H}^2! Again, fine-tuning between the different (and in SM unrelated) constants governing these corrections would be needed to make the corrections cancel. Observing that the terms in Equation (1.2.1) enter with opposite signs, a popular solution has been to introduce a new symmetry between fermions and bosons. This *Supersymmetry* (SUSY) introduces new fermions as partners to the SM bosons and vice-versa.

- The large number of free parameters of the SM is considered to be unsatisfying for a fundamental theory by many [33, p. 483]. This particularly applies to the Yukawa sector of the SM that contains 13 of these parameters (6 quark masses, 3 charged lepton masses, 4 real parameters for V_{CKM}). For example, the hierarchical structure of the masses and of the V_{CKM} elements (see Equation (1.1.18)) is not explained in the SM.
- The *Grand Unification Hypothesis* predicts that the interactions of the SM are aspects of one fundamental interaction described by one unified coupling constant and one gauge symmetry. This symmetry would then be broken at lower energies, giving rise to the three separate symmetry groups and couplings [33, pp. 508–509]. Extrapolating the running (i.e., energy dependence) of the three SM coupling constants to high energies, all three constants come close at one point but do not meet precisely in the SM. This can be taken as motivation for extensions like SUSY (Figure 1.7).

1.2.2 Dark matter

Cosmological observations have long concluded that the visible amount of matter is inconsistent with the gravitational effects in the Universe. The concept of a hypothetical, non-luminous type



(a) 1st order QED (Schwinger term [55])

(b) Lowest-order weak

(c) Lower-order hadronic

Figure 1.8: Loop corrections to g_μ. Figure from [25].

of matter to explain these observations can be traced back to the late 19th century [40]. One of the most famous pioneering works is that of F. Zwicky [41], who calculated the masses of galaxy clusters using the virial theorem.[19] Since then, the observational evidence for dark matter has become overwhelming (assuming that our description of gravity is valid): Dark matter has been found in cosmological structures of various scales [25], from dwarf galaxies [42], regular galaxies [43] to clusters of galaxies [44].

In the ΛCDM cosmological model that is strongly supported by the measurements of the cosmic microwave background [45], the overwhelming majority of the total matter density of the Universe is of non-baryonic, almost electrically neutral origin. Massive neutrinos can only explain a small fraction of this matter [46, pp. 15–16]. The remainder is called cold dark matter. Altogether, dark energy and dark matter dominate the total energy content of the Universe, while ordinary matter contributes less than 5 % of the energy budget [45].

Because dark matter has proven to be very elusive in terrestrial experiments, modified models of gravity have been proposed to explain the cosmological observations with only SM matter [47, 48]. However, these models typically remain in tension with the observed properties of galaxy clusters and gravitational waves [49–51].

1.2.3 Anomalous magnetic moment of the muon

In the SM, the magnetic moments of the electron and muon are given by

$$\vec{\mu}_\ell = g_\ell \left(\frac{q}{2m_\ell} \right) \vec{S}, \tag{1.2.2}$$

where g_ℓ is the gyromagnetic ratio (a dimensionless constant), q the (elementary) charge, and $\vec{S}$ the spin of the lepton. The tree-level[20] value of g_ℓ can be calculated from the Dirac equation [52] and is equal to two. However, loops in this interaction between the photon and the lepton introduce corrections $a_\ell := (g_\ell - 2)/2$ to this value (Figure 1.8). While the predicted and measured value of a_e agree to very high precision, the value of a_μ currently shows a 4.2σ tension [53]. The latest measurement [53] alone finds a tension of 3.3σ. A recent review of theoretical predictions and experimental measurements of $g_\mu - 2$ is [54].

1.2.4 Baryon asymmetry

There is a striking asymmetry between the amount of (baryonic) matter and antimatter in the observable Universe. In fact, antimatter is almost entirely absent from the cosmos, while all observable astronomical structures consist of matter. This was not always the case: The early Universe was hot, and pair creation and annihilation were in thermal equilibrium. Large quantities of both matter and antimatter were present. With the cooling of the cosmos, the pair

[19] F. Zwicky is sometimes credited with the invention of the term “dark matter”, but this is not historically accurate [40].

[20] That is, calculated without loop corrections.

production stopped, and most of the matter and antimatter annihilated to photons. Counting the presently observable baryons and photons, the baryonic asymmetry of the hot Universe (BAU) can be estimated as [45, 56]:

$$\eta := \left.\frac{N_B}{N_\gamma}\right|_{T=3K} = \left.\frac{N_B - N_{\bar{B}}}{N_\gamma}\right|_{T=3K} \sim \left.\frac{N_B - N_{\bar{B}}}{N_B + N_{\bar{B}}}\right|_{T\gtrsim 1\,\mathrm{GeV}} \sim 10^{-10}, \tag{1.2.3}$$

where $N_{B/\bar{B}/\gamma}$ is the number of (anti)baryons/photons. $T = 3\,\mathrm{K}$ indicates the present Universe, $T \gtrsim 1\,\mathrm{GeV}$ the early hot Universe.

For BAU to occur, three phenomenons have to be allowed [57] (*Sakharov conditions*):

i. Baryon number violation,

ii. $\mathcal{P}$ and $\mathcal{CP}$ violation

iii. deviation from the thermal equilibrium.

Technically, the SM satisfies all of these conditions:

i. The SM allows *sphaleron processes* [58, 59] that violate the baryon number. While these processes are heavily suppressed at low temperatures, the sphalerons become efficient at temperatures above the scale of electroweak symmetry breaking [60–62].

ii. Both $\mathcal{P}$ and $\mathcal{CP}$ are violated in the SM [27, 63]. $\mathcal{P}$ symmetry is violated maximally because the weak force only acts on left-handed particles, and the complex phase in V_{CKM} (as discussed in Section 1.1.3) is responsible for $\mathcal{CP}$ violation.

iii. This can be fulfilled by cosmic inflation.

However, the $\mathcal{CP}$ violation of the SM and the deviation from the thermal equilibrium that might occur during the electroweak symmetry breaking are believed to be orders of magnitude too small to explain BAU.

1.2.5 Flavor physics

The decays $\mathrm{B} \to \mathrm{D}^{(*)}\ell\nu_\ell$ have been studied extensively to measure $|V_{\mathrm{cb}}|$ (see Section 1.4). In contrast, $\mathrm{B} \to \mathrm{D}^{(*)}\tau\nu_\tau$ decays are more challenging experimentally because of the short life time of the τ and the additional neutrinos in its decay. Rather than measuring both decays separately, it is advantageous to consider the ratio

$$\mathcal{R}(\mathrm{D}^{(*)}) = \frac{\mathrm{BR}(\mathrm{B} \to \mathrm{D}^{(*)}\tau\nu_\tau)}{\mathrm{BR}(\mathrm{B} \to \mathrm{D}^{(*)}\ell\nu_\ell)}, \tag{1.2.4}$$

because many theoretical uncertainties (e.g., $|V_{\mathrm{cb}}|$, form factor normalizations) and experimental uncertainties (e.g., detector efficiencies) approximately cancel in the ratio. As can be seen in Figure 1.9, there is a significant discrepancy between the experimental measurements of $\mathcal{R}(\mathrm{D}^{(*)})$ and its theoretical predictions. Currently, this discrepancy exceeds three standard deviations.

In addition to the charged current observable $\mathcal{R}(\mathrm{D}^{(*)})$, neutral-current anomalies involving $\mathrm{b} \to \mathrm{s}\ell^+\ell^-$ transitions are observed. These anomalies encompass both deviations for e/μ universality [64–67] and discrepancies observed in decays with $\mu^+\mu^-$ pairs only [68, 69].

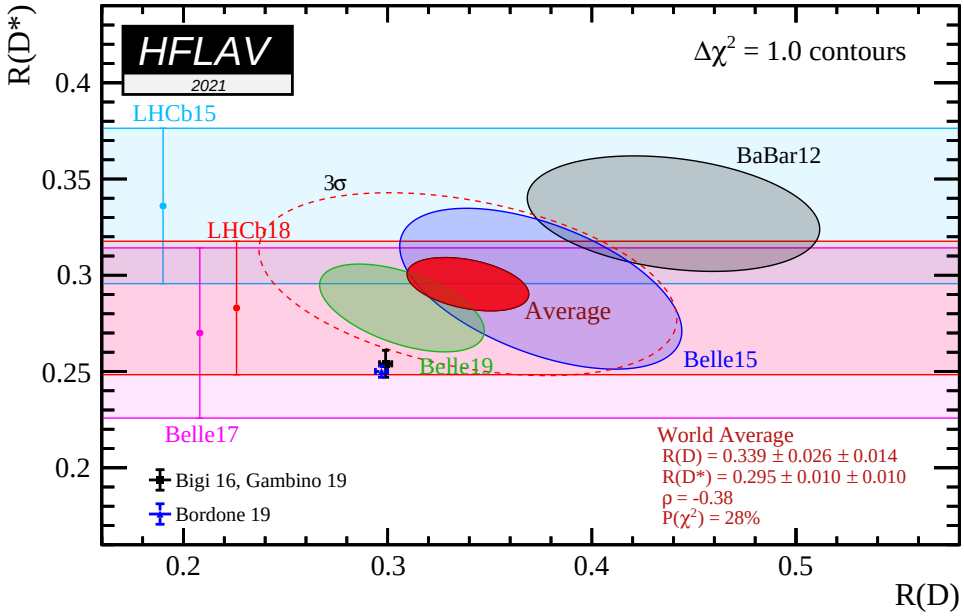


Figure 1.9: Measurements and predictions of $\mathcal{R}(\mathrm{D}^{(*)})$. Figure from [70]. Measurements and predictions are listed in Table 1.3.

	$\mathcal{R}(\mathrm{D})$	$\mathcal{R}(\mathrm{D}^*)$		τ-decay	Details
2013	$0.440 \pm 0.058 \pm 0.042$	$0.332 \pm 0.024 \pm 0.018$	BaBar	Leptons	Had. tag, $\rho = -0.27$ [71, 72]
2015	$0.375 \pm 0.064 \pm 0.026$	$0.293 \pm 0.038 \pm 0.015$	Belle	Leptons	Had. tag, $\rho = -0.49$ [73]
2015		$0.336 \pm 0.027 \pm 0.030$	LHCb	Leptons	[74]
2018		$0.270 \pm 0.035\ ^{+0.028}_{-0.025}$	Belle	Hadrons	Had. tag [75, 76]
2018		$0.280 \pm 0.018 \pm 0.029$	LHCb	Hadrons	[77, 78]
2020	$0.307 \pm 0.037 \pm 0.016$	$0.283 \pm 0.018 \pm 0.014$	Belle	Leptons	SL tag, $\rho = -0.51$ [79]
2021	$0.339 \pm 0.026 \pm 0.014$	$0.295 \pm 0.010 \pm 0.010$	HFLAV	$\chi^2/\mathrm{dof} = 8.8/7$ (C.L. = 0.28) [70]	
2016	0.299 ± 0.003		SM		[80]
2019		$0.254 \pm 0.007 \pm 0.006$	SM		[81]
2019	0.297 ± 0.003	0.250 ± 0.003	SM		[82]

Table 1.3: Measurements (top), their average (middle) and predictions (bottom) of $\mathcal{R}(\mathrm{D}^{(*)})$. For a visual comparison, see Figure 1.9. Abbreviations: *Had.* (hadronic), *SL* (semileptonic), C.L. (confidence limit), ρ (statistical and systematic correlation between the $\mathcal{R}(\mathrm{D})$ and $\mathcal{R}(\mathrm{D}^*)$ measurement).

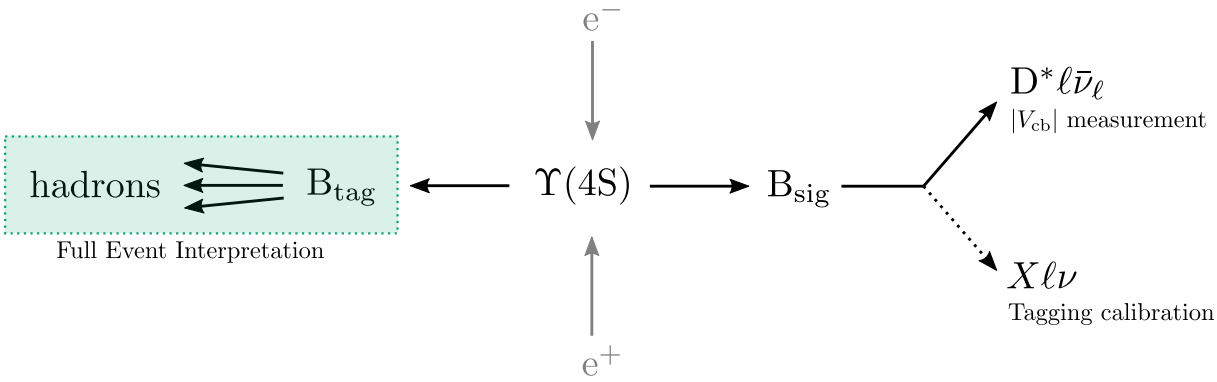


Figure 1.10: Illustration of the $\Upsilon(4S)$ reconstruction for the $B \to D^* \ell \nu_\ell$ measurement and the tagging calibration.

1.3 Hadronic tagging and the Full Event Interpretation

After having introduced the electroweak sector of the SM and possible signs of lepton flavor universality violation, we now turn to an experimental strategy that is central to this thesis: tagging.

The Belle experiment collides electrons and positrons to produce $\Upsilon(4S)$ particles that decay into pairs of B mesons (see Section 3.1 for details). *B-tagging* means reconstructing one of the two B mesons from the $\Upsilon(4S)$ decay (*tag B*, B_{tag}) in a well-understood decay channel that is not in the focus of the measurement. This allows constraining the kinematics of the other B meson (*signal B*, B_{sig}) that is reconstructed in the channel of interest (Figure 1.10).

This technique is particularly beneficial for semileptonic signal decays because they include a neutrino that makes direct B_{sig} momentum measurements impossible. However, if the B_{tag} is reconstructed in hadronic modes, the B_{tag} momentum can be used together with the knowledge of the beam energies to infer the momentum of the B_{sig}. Furthermore, if both B mesons are reconstructed correctly, all particles in the $e^- e^+$ collision are accounted for (because only a $\Upsilon(4S)$ has been produced). This can be used for very effective background suppression techniques, resulting in very pure samples. The disadvantage of tagging is the lower reconstruction efficiency because both B mesons need to be reconstructed successfully. Analyses that do not reconstruct the B_{tag} in exclusive decay modes are referred to as *untagged* analyses.

A new tagging algorithm, the *Full Event Interpretation* (FEI, described in detail in Section 4.1.1) surpasses the efficiency of the previously used hadronic tagging algorithm up to twofold, prompting reanalyses of several studies of the Belle dataset.

This includes a new measurement of $\mathcal{R}(D^{(*)})$ with leptonic τ decays [83, 84] that is to replace the earlier result [73].

Furthermore, this thesis shows preparations for a $|V_{cb}|$ measurement using hadronically tagged $B \to D^* \ell \nu_\ell$ decays (discussed in detail in Section 1.5 and Chapter 5). This measurement is closely related to [83, 84], because $B \to D^* \ell \nu_\ell$ is the normalization mode of $\mathcal{R}(D^*)$.

The FEI relies heavily on machine learning algorithms that are trained on Monte Carlo (MC) simulations of collision data. Differences between the simulated data and recorded data can lead to a substantial difference in reconstruction efficiencies of the FEI on the two datasets. This particularly affects the $|V_{cb}|$ measurement that requires measuring the (absolute) decay rate of $B \to D^* \ell \nu_\ell$ decays.

The performance of the FEI must thus be calibrated in a data-driven way. This is done by reconstructing hadronically tagged $B \to X \ell \nu_\ell$ decays (Figure 1.10). Assuming that the efficiencies of the $B \to X \ell \nu_\ell$ signal side reconstruction are well understood, the ratio of event yields on data

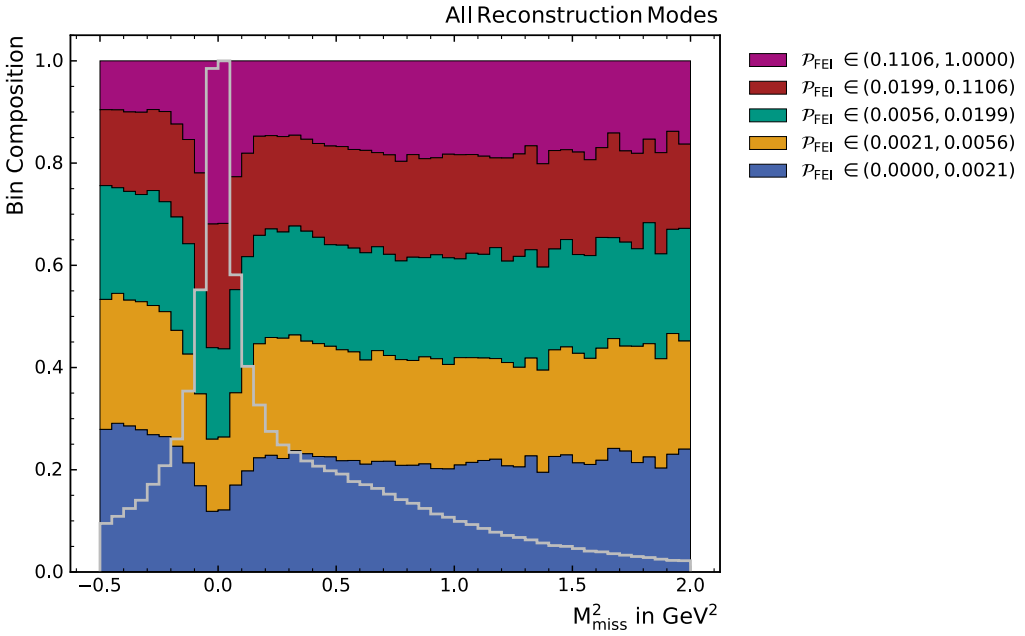


Figure 1.11: Correlation between $\mathcal{P}_{\text{FEI}}$ and the m^2_{miss} observable used as the fitting variable in the $|V_{cb}|$ analysis and as one of the variables in the $\mathcal{R}(D^{(*)})$ analysis. The gray line shows the m^2_{miss} distribution in the $\mathcal{R}(D^{(*)})$ analysis. For each bin in m^2_{miss}, the composition of $\mathcal{P}_{\text{FEI}}$ is shown in five bins. Correctly reconstructed $B \to D^{(*)}\ell\nu_\ell$ decays peak at $m^2_{\text{miss}} = 0\,\text{GeV}^2$, which is accompanied by a significant increase in the fraction of high values of $\mathcal{P}_{\text{FEI}}$. Figure from [83].

and MC can be used as a calibration factor.

The calibration factors are found to depend on several quantities, including $\mathcal{P}_{\text{FEI}}$, a classifier output of the FEI that is related to the probability of correctly reconstructing the B_{tag}. Because tagged analyses combine both reconstructed B mesons to a $\Upsilon(4S)$ and apply additional selection criteria, the probability of retaining an event with correctly reconstructed B_{tag} becomes correlated to the probability of retaining an event with correctly reconstructed B_{sig}. This correlates $\mathcal{P}_{\text{FEI}}$ with signal side observables that allow discriminating between signal and background. As an example, Figure 1.11 shows the correlation between $\mathcal{P}_{\text{FEI}}$ and the missing mass

$$m^2_{\text{miss}} := (p^\mu_B - p^\mu_{D^{(*)}} - p^\mu_\ell)^2. \tag{1.3.1}$$

Consequently, the $\mathcal{P}_{\text{FEI}}$ dependency of the calibration factors affects the shape of distributions used for background separation in hadronically tagged analyses. This includes the $\mathcal{R}(D^{(*)})$ and $|V_{cb}|$ analyses mentioned above.

1.4 Measuring $|V_{cb}|$

All experimentally competitive measurements of $|V_{cb}|$ are performed using semileptonic decays mediated by a $b \to c\ell\bar{\nu}_\ell$ transition [25]. The measurement strategies fall into two categories:

Exclusive measurements reconstruct a specific semileptonic decay with a $b \to c\ell\nu$ transition, e.g., $B \to D^*\ell\nu_\ell$. The decay rate is proportional to $|V_{cb}|^2$. However, to extract $|V_{cb}|$ from the decay rate measurement, hadronic matrix elements need to be evaluated. Because they cannot be calculated perturbatively, they are expressed in terms of *form factors* that are parameterized and fitted to the differential cross section. External input from lattice QCD can then be used to determine all remaining constants entering the amplitude, such that $|V_{cb}|^2$ can be extracted.

Exclusive measurements can further be categorized by whether or not they employ the B-tagging strategy introduced in Section 1.3. Tagged datasets have significantly higher purity, making them particularly useful for constraining hadronic form factor parameterizations. However, untagged measurements have significantly higher reconstruction efficiency and thus lower statistical uncertainties. Both strategies are therefore complementary.

Inclusive measurements measure the spectrum of all $b \to c\ell\nu$ decays, that is $B \to X_c\ell\bar{\nu}_\ell$. On the theoretical side, this total semileptonic rate is calculated using *Heavy Quark Expansion* (*HQE*) [85, 86]. At leading order, the free parameters are m_b and m_c. At $\mathcal{O}(1/m_b^2)$ and at $\mathcal{O}(1/m_b^3)$, two more parameters describing non-perturbative corrections are added (denoted μ_π, μ_G and ρ_D^3, ρ_{LS}^3 respectively). All of these parameters are *running*, that is depend on the renormalization scale μ. The experiments measure moments of the hadronic mass distribution $\langle M_X^n \rangle$ and the lepton energy distribution $\langle E_\ell^n \rangle$ ($n \in \mathbb{N}$) and then fit the theoretical parameterization (the parameters just described and $|V_{cb}|$) to these measurements. [87] provides a succinct introduction from an experimental perspective.

Recently, a novel measurement strategy based on a HQE symmetry called *reparameterization invariance* has been suggested [88]. This symmetry allows to significantly reduce the number of required non-perturbative parameters, such that only eight parameters are needed at $\mathcal{O}(1/m_b^4)$. The first measurement of the q^2 moments required for the calculation has been recently completed at Belle [89].

Currently, these values are [25]

$$\begin{aligned} |V_{cb}| &= (42.2 \pm 0.8) \times 10^{-3} \quad \text{(inclusive)}, \\ |V_{cb}| &= (39.5 \pm 0.9) \times 10^{-3} \quad \text{(exclusive)}. \end{aligned} \tag{1.4.1}$$

As is clearly visible, some tension exists between both measurements. In fact, the p value of a χ^2 test is only $p = 2\,\%$. Therefore the current average is quoted [25] with uncertainties scaled up by $\sqrt{\chi^2/1} \approx 2$ (see [90]) as

$$|V_{cb}| = (41.0 \pm 1.4) \times 10^{-3}. \tag{1.4.2}$$

One hypothesis for the origin of the tension between both strategies is discussed in Section 1.4.2.3.

1.4.1 $|V_{cb}|$ from $B \to D^* \ell\nu_\ell$

The main objective of this thesis is the preparation of a measurement of $B \to D^*\ell\nu_\ell$, including a determination of $|V_{cb}|$. This section introduces the basic principle of $|V_{cb}|$ measurements with exclusive semileptonic decays on the example of $B \to D^*\ell\nu_\ell$. The specific calculation of $|V_{cb}|$ in our analysis also includes additional helicity angle information and is discussed in more detail in Section 1.5 and Chapter 5.

We now consider the differential cross section $B \to D^*\ell\nu_\ell$ with respect to w (*recoil*), the product of the four-velocities of the B and D meson:

$$w := v_B^\mu v_{\mu D^*} = \frac{m_B^2 + m_{D^*}^2 - q^2}{2 m_B m_{D^*}} \quad \text{with} \quad q^2 := (p_B - p_{D^*})^2, \tag{1.4.3}$$

where v^μ, m_X, p_X denotes the four-velocities, masses and four-momenta of $X = B$ or $X = D^*$. The *zero recoil point* mentioned in Section 1.4 corresponds to $w = 1$, or $q^2 = (m_B - m_{D^*})^2$. This is in fact the minimum of w (maximum of q^2)[21], that is, the point of maximal amount of momentum transfer to the $\ell\nu$ system.

[21] As q^2 is Lorentz invariant, it can be evaluated in the B rest frame as $q^2 = (p_B - p_{D^*})^2 = m_B^2 + m_{D^*}^2 - 2(E_B E_{D^*} - 0 \cdot \vec{p}_{D^*}) \leq m_B^2 + m_{D^*}^2 - 2 m_B m_{D^*} = (m_B - m_{D^*})^2$.

	$\mathcal{F}(1) \times \lvert V_{cb}\rvert$	Exp.	Res.	Details
1997	$0.0319 \pm 0.0018 \pm 0.0019$	ALEPH	Z	[91]
2000	$0.0371 \pm 0.0010 \pm 0.0020$	OPAL	Z	[92]
2001	$0.0355 \pm 0.0014^{+0.0023}_{-0.0024}$	DELPHI	Z	[93]
2003	$0.0431 \pm 0.0013 \pm 0.0018$	CLEO2	$\Upsilon(4S)$	Untagged. Fitting $\cos\theta_{B-D^*\ell}$ [94]
2004	$0.0392 \pm 0.0018 \pm 0.0023$	DELPHI	Z	[95]
2009	$0.0359 \pm 0.0006 \pm 0.0014$	BaBar	$\Upsilon(4S)$	Untagged. Fitting $\lvert\vec{p}^*_\ell\rvert$, $\lvert\vec{p}^*_D\rvert$, $\cos\theta_{B-D^*\ell}$ to reconstructed $DX\ell\bar{\nu}$ states [96]
2019	$0.03483 \pm 0.00015 \pm 0.00056$	Belle	$\Upsilon(4S)$	Untagged. Fitting ΔM, $\lvert\vec{p}_\ell\rvert$, $\cos\theta_{B-D^*\ell}$, [97]
	0.0355 ± 0.0008	HFLAV	avg.	Uncert. inflated by $\sqrt{\chi^2/(\mathrm{ndf}-1)} = \sqrt{17.2/6} = 1.7$ [98]
	0.03527 ± 0.00038	HFLAV	eval.	[98]

Table 1.4: Exclusive $|V_{cb}|$ measurements via $B \to D^*\ell\nu_\ell$ that enter the current world average. Abbreviations: *Exp.* (Experiment), *Res.* (resonance: $e^-e^+ \to X$), *avg.* (average), *eval.* (evaluation).

The differential cross section of $B \to D^*\ell\nu_\ell$ can be expressed as [99, p. 98][100, p. 375]

$$\frac{d\Gamma}{dw}\left(B \to D^*\ell^-\bar{\nu}_\ell\right) = \frac{G_F^2 m_{D^*}^3}{48\pi^3}(m_B - m_{D^*})^2 \chi(w)\eta_{EW}^2 \mathcal{F}^2(w)|V_{cb}|^2, \tag{1.4.4}$$

where G_F is the Fermi constant, η_{EW}^2 is a electroweak correction [101], $\mathcal{F}(w)$ is the form factor (further discussed below), and $\chi(w)$ is a known phase space factor

$$\chi(w) := \sqrt{w^2-1}(w+1)^2\left(1 + \frac{4w}{w+1}\frac{1-2rw+r^2}{(1-r)^2}\right) \tag{1.4.5}$$

with $r := m_{D^*}/m_B$. The electroweak correction can be approximated as [102] $\eta_{EW} \approx 1 + \alpha/\pi \log(m_Z/m_B)$ (with α the electromagnetic coupling constant constant and m_Z the mass of the Z-boson) and is currently listed as [25]

$$\eta_{EW} = 1.0066 \pm 0.0050, \tag{1.4.6}$$

which includes an additional uncertainty for long-distance QED radiative corrections [103, pp. 28–29].

With this, the only unknown in Equation (1.4.4) is $\mathcal{F}(w)|V_{cb}|$. To extract $|V_{cb}|$ we thus need theoretical input for $\mathcal{F}(w)$. The most precise input comes from lattice QCD calculation but has until recently only been available at the zero recoil point as $\mathcal{F}(1)$. Recently, preliminary lattice calculations at additional non-zero recoil values up to $w \sim 1.1$ have become available [104, 105], but the precision is not yet comparable to the results at $w = 1$. Nonetheless, the results can be included as additional constraints [106].

Lattice results for $w = 1$ are reviewed by the FLAG collaboration in [107, pp. 182–183]. Averaging the results of [103] and [108], [25] lists

$$\mathcal{F}(1) = 0.904 \pm 0.012. \tag{1.4.7}$$

The fact that $\mathcal{F}(1) \approx 1$ is not surprising: In the limit of infinite mass, heavy quark symmetry predicts $\mathcal{F}(1) = 1$.

Knowing $\mathcal{F}(w)$ only at low w, we cannot simply measure $\Delta\Gamma/\Delta w$ at $w \gtrsim 1$: As evident from Equation (1.4.5), the differential cross section vanishes at zero recoil (and for any narrow selection in w, the statistical uncertainties become substantial). The alternative is to perform a differential measurement of $\Delta\Gamma/\Delta w$ and fit it to a parameterization of $\mathcal{F}(w)$ before using the lattice input to determine the normalization independently of $|V_{\mathrm{cb}}|$.

$\mathcal{F}(w)$ can be expressed in terms of three independent functions $h_{A_1}(w)$, $R_1(w)$ and $R_2(w)$ [99, p. 98]:

$$\begin{aligned}\chi(w)\mathcal{F}^2(w) = h_{A_1}^2(w)\sqrt{w^2-1}(w+1)^2\times \\ \times\left\{2\left[\frac{1-2wr+r^2}{(1-r)^2}\right]\left[1+R_1^2(w)\frac{w-1}{w+1}\right]+\left[1+(1-R_2(w))\frac{w-1}{1-r}\right]^2\right\},\end{aligned} \tag{1.4.8}$$

where the lepton mass is neglected. The ratios $R_1(w)$ and $R_2(w)$ are defined in terms of the functions $h_V(w)$ (vector form factor), $h_{A_1}(w)$, $h_{A_2}(w)$ and $h_{A_3}(w)$ (axial form factors):

$$R_1(w) = \frac{h_V(w)}{h_{A_1}(w)}, \quad R_2(w) = \frac{h_{A_3}(w)+rh_{A_2}(w)}{h_{A_1}(w)}. \tag{1.4.9}$$

Finally, the interpretation of these form factors becomes clear as a decomposition of the hadronic matrix elements:

$$\begin{aligned}\frac{\left\langle \mathrm{D}^*(v_{\mathrm{D}^*},\epsilon)|\bar{c}\gamma^\mu b|\mathrm{B}(v_{\mathrm{B}})\right\rangle}{\sqrt{m_{\mathrm{B}}m_{\mathrm{D}^*}}} &= h_V(w)\varepsilon^{\mu\nu\rho\sigma}v_{\mathrm{B},\nu}v_{\mathrm{D}^*,\rho}\epsilon^*_\sigma \\ \frac{\left\langle \mathrm{D}^*(v_{\mathrm{D}^*},\epsilon)\middle|\bar{c}\gamma^\mu\gamma^5 b\middle|\mathrm{B}(v_{\mathrm{B}})\right\rangle}{\sqrt{m_{\mathrm{B}}m_{\mathrm{D}^*}}} &= \mathrm{i}h_{A_1}(w)(1+w)\epsilon^{*\mu} - \mathrm{i}\left[h_{A_2}(w)v_{\mathrm{B}}^\mu + h_{A_3}(w)v_{\mathrm{D}^*}^\mu\right]\epsilon^*\cdot v_{\mathrm{B}},\end{aligned} \tag{1.4.10}$$

where v_{B}, v_{D^*} are the four-velocities, $\varepsilon^{\mu\nu\rho\sigma}$ is the antisymmetric tensor and ϵ the polarization of the D^*. This parameterization of the hadronic matrix elements in terms of h_V and $h_{A_{1,2,3}}$ is also called the heavy quark symmetry (HQS) basis.[22]

1.4.2 Form Factors

Form factors can be parameterized and calculated in different theoretical frameworks. Four different approaches can be distinguished [102]:

1. Constraining the structure of the form factors based on the functional properties of the hadronic matrix elements they describe. The notable example is the BGL parameterization (Section 1.4.2.1) which uses dispersion relations, analyticity and unitarity to describe form factors as a relatively unconstrained and model-independent series expansion.

2. Using heavy quark effective theory (HQET, see Section 1.4.2.2) to relate the different form factors to each other. The notable example is the CLN parameterization (Section 1.4.2.2). In the form most often used in experimental analyses, it only contains the normalizations $h_{A_1}(1)$, $R_1(1)$ and $R_2(1)$ and the slope parameter ρ as free parameters. However, leaving these parameters to independently float in fits to experimental data neglects correlations between them. Furthermore, the CLN model imposes stronger constraints than BGL and needs appropriate model uncertainties. At current experimental precision levels, these previously often neglected sources of uncertainty need to be included, as will be discussed in more detail in Section 1.4.2.2. The BLPR model [109] is a recently proposed improvement on the CLN description that addresses these concerns.

[22] Because of different conventions, expressions like Equation (1.4.10) might differ by a factor of ± 1 or $\pm\mathrm{i}$ [102].

3. Approximating the form factors using methods that take the quark picture into account, such as QCD sum rules or light cone sum rules. An example is the ISGW2 model [110, 111] which was used for the description of D^{**} decays in Belle MC (see Section 7.5.2).

4. Performing lattice QCD calculations. In contrast to the other strategies that only constrain the dependencies of the form factors and require additional experimental input, this approach allows for calculations of absolute values from first principles. However, lattice results are currently only available for a subset of kinematic regimes (as discussed in Section 1.4.1).

1.4.2.1 BGL parameterization

The Boyd Lebed Grinstein (BGL) parameterization [112, 113] uses the helicity basis to express the hadronic matrix elements in terms of form factors. A summary that focuses on the ideas from a theoretical perspective is presented in [102], a very concise list of formulas needed to calculate and use them is given in [114].

Coming from the HQS basis (Equation (1.4.10)), we define:

$$\begin{aligned} f(w) &:= h_{A_1}(w) m_B \sqrt{r}(w+1), \\ g(w) &:= \frac{h_V(w)}{r m_B}, \\ \mathcal{F}_1(w) &:= h_{A_1}(w) m_B^2 \sqrt{r}(w+1)[w - r - (w-1)R_2(w)]. \end{aligned} \tag{1.4.11}$$

The idea is to parameterize these three functions as a power series. To ensure the convergence of the power series, we first need to transform w to a variable z with $|z| < 1$. A simple conformal mapping is given by

$$z(w) = \frac{\sqrt{w+1} - \sqrt{2}}{\sqrt{w+1} + \sqrt{2}}. \tag{1.4.12}$$

Before we expand f, g and F_1 in z, we first remove poles at $z = z_P$ that correspond to known resonances below the B-D^*-threshold:

$$z_P := \frac{\sqrt{t_+ - m_P^2} - \sqrt{t_+ - t_-}}{\sqrt{t_+ - m_P^2} + \sqrt{t_+ - t_-}}, \tag{1.4.13}$$

where $t_\pm := (m_B \pm m_{D^*})^2$ and m_P are the masses of the resonances. They are removed using *Blaschke factors* [115, p. 117] $P_{1^\pm}$:

$$P_{1^\pm} := \prod_P \frac{z - z_P}{1 - z z_P} \tag{1.4.14}$$

where P_{1^+} is used for f and $\mathcal{F}_1$ and P_{1^-} for g. Furthermore, some additional q^2 dependencies and constants are contained in the weighting functions ϕ_i (with $i = f, g, F_1$).[23]

With that, we can write

$$\begin{aligned} f(z) &= \frac{1}{P_{1+}(z)\phi_f(z)} \sum_{n=0}^{\infty} a_n^f z^n, \\ \mathcal{F}_1(z) &= \frac{1}{P_{1+}(z)\phi_{\mathcal{F}_1}(z)} \sum_{n=0}^{\infty} a_n^{\mathcal{F}_1} z^n, \\ g(z) &= \frac{1}{P_{1-}(z)\phi_g(z)} \sum_{n=0}^{\infty} a_n^g z^n. \end{aligned} \tag{1.4.15}$$

[23] They are listed in equation (4.23) together with the parameters of table 1 of the original paper [113] or in verbatim in [114].

with coefficients $a_n^i \in \mathbb{R}$, which can currently only be determined by experimental measurements of $B \to D^* \ell \nu_\ell$. Of course, experiments will only ever be able to fit the parameters a_n^i up to finite $n = N$. In order to estimate the uncertainty on the extrapolation of $\mathcal{F}(w)$ to the zero recoil point, we thus need to be able to control the uncertainty that we incur by prematurely truncating this series expansion. In other words, we need an upper bound for the factors a_n^i. Using analyticity and an operator product expansion applied to correlators of the hadronic $\bar{c}b$ currents, the following *unitarity bounds* can be shown:

$$\begin{aligned} \sum_{i=0}^{\infty} |a_n^g|^2 &\leq 1, \\ \sum_{i=0}^{\infty} \left[|a_n^f|^2 + |a_n^{\mathcal{F}_1}|^2 \right] &\leq 1. \end{aligned} \tag{1.4.16}$$

1.4.2.2 CLN parameterization

The key ingredient for the parameterization by Caprini, Lellouch and Neubert (CLN) [116] is *Heavy Quark Effective Theory* (HQET) pioneered in [111, 117–119]. HQET is an effective field theory used to describe systems of a light quark and a heavy quark (of mass $m_Q \gg \Lambda_{\text{QCD}}$; in our case, $Q = b$ and $Q = c$), which is considered as a static source of the gluon field. Interactions with the heavy quark enter at higher orders of $1/m_Q$. In the limit of infinite mass, new symmetries appear in the HQET Lagrangian that are not present in the standard QCD Lagrangian. These symmetries also allow obtaining relations between form factors of heavy mesons.

These relations are used in the CLN model to relate all form factors to one single form factor, h_{A_1} (using the HQS basis from Equation (1.4.10) again), which is expanded up to the third order in z. Additionally, QCD sum rules are employed to constrain the $1/m_{c,b}$ corrections of HQET. In this way, h_{A_1} is famously approximated as [116, eq. (38)]

$$h_{A_1} \approx 1 - 8\rho_{A_1}^2 z + (53.\rho_{A_1}^2 - 15.)z^2 - (231.\rho_{A_1}^2 - 91.)z^3, \tag{1.4.17}$$

where ρ_{A_1} (oftentimes only referred to as ρ) is a free parameter. The precision of this approximation was estimated to be better than 2 %. While this uncertainty was negligible in comparison to the experimental data available at the time, this is no longer the case (see the precision in Equation (1.4.1)). It is thus problematic that this uncertainty is not included in most experimental results that use the CLN parameterization [109, 114].

The ratios $R_1(w)$ and $R_2(w)$ are expressed in powers of $w - 1$:

$$R_i(w) = R_i(1) + R_i'(1)(w-1) + R_i''(1)(w-1)^2. \tag{1.4.18}$$

The corresponding coefficients have been calculated by the original authors as [116, eq. (36)]

$$\begin{aligned} R_1(w) &\approx 1.27 - 0.12(w-1) + 0.05(w-1)^2, \\ R_2(w) &\approx 0.80 + 0.11(w-1) - 0.06(w-1)^2. \end{aligned} \tag{1.4.19}$$

It needs to be noted that these relations are subject to uncertainties from QCD sum rules as well as $\mathcal{O}(\Lambda_{\text{QCD}}^2/m_c^2)$ and $\mathcal{O}(\alpha_s \Lambda_{\text{QCD}}/m_c)$ corrections, where α_s is the coupling constant of the strong force. Again, these uncertainties were not considered in experimental analyses [109, 114].

Another issue common to many experimental analyses is that the normalizations $R_i(1)$ were fitted to data while $R_i'(1)$ and $R_i''(1)$ were fixed to the QCD sum rule predictions of Equation (1.4.19). This disregards the fact that $R_i(1)$ is correlated to $R_i'(1)$ and $R_i''(1)$ [102].

This and other issues are addressed in the BLPR model [109].

1.4.2.3 Form factors and the inclusive/exclusive puzzle

It has been suspected that the inadequate implementation of the CLN parameterization (or systematic uncertainties thereof) could be responsible for the tension between inclusive and exclusive measurements of $|V_{cb}|$. This debate was fueled by independent fits [114, 120] to the kinematic distributions of the unfinished tagged $B \to D^* \ell \nu_\ell$ analysis [121] that this thesis prepares to repeat and improve upon (see Chapter 5). Since then, additional studies have confirmed that $|V_{cb}|$ measurements using $B \to D^* \ell \nu_\ell$ can be sensitive to form factor model choices [81, 122].

In contrast, a very recent study [106] has used the data from the untagged Belle measurement of $B \to D^* \ell \nu_\ell$ [97] and compared the results for different form factor parameterizations, including also recent lattice QCD results at non-zero recoil. Using detailed studies of systematic uncertainties, it concludes that the results obtained with the CLN and BGL parameterizations are compatible on this dataset. In particular, the tension between the inclusive and exclusive measurements remains.

1.5 Hadronically tagged $B \to D^* \ell \nu_\ell$ decays

This thesis prepares for a measurement of $B \to D^* \ell \nu_\ell$ decays with hadronically reconstructed tag B. This decay is of interest for various reasons:

- It is the normalization mode of the $\mathcal{R}(D^*)$ measurement (Section 1.2.5). While the tension is most often attributed to an insufficient understanding or new physics in $B \to D^* \tau \nu_\tau$ (which is generally harder to measure experimentally because of the τ decays), the possibility of unaccounted effects in $B \to D^* \ell \nu_\ell$ must be excluded. This benefit particularly applies to the analysis presented here, because the dataset is directly produced by the framework of the upcoming $\mathcal{R}(D^{(*)})$ measurement [83, 84] (see Section 1.2.5), including all corrections to MC simulation (see Chapter 7).
- As already outlined in Section 1.4.1, $B \to D^* \ell \nu_\ell$ provides an excellent avenue to measure $|V_{cb}|$ and form factor parameters. Improved consideration of different form factor models could contribute to an improved understanding of the tension between the inclusive and exclusive measurements of $|V_{cb}|$. This particularly applies to this analysis, because inconsistencies between the $|V_{cb}|$ values obtained with different form factor models were demonstrated on previous preliminary results (see Section 1.4.2.3).
- $B \to D^* \ell \nu_\ell$ is a major source of background for reconstructing $B \to D^* \tau \nu_\tau$ and charmless semileptonic decays such as $B \to \pi \ell \nu$. Improved knowledge of the form factor parameters can help to reduce systematic uncertainties in reconstruction.
- The form factors are one of the ingredients for the theoretical predictions of $\mathcal{R}(D^*)$. Improving our understanding of the different parameterizations and increasing the precision of the measurements of the parameters can thus also weigh in on the debate about the flavor anomalies.
- The $B \to D^* \ell \nu_\ell$ decay can be used to calibrate the efficiency of hadronic tagging. Of course, this calibration mode is not possible for this analysis itself.
- $B \to D^* \ell \nu_\ell$ can be used as a tagging mode for measurements of other decays. The FEI tagging algorithm that is used for hadronic tagging in this thesis can also reconstruct semileptonic tags, which are dominated by $B \to D^{(*)} \ell \nu_\ell$ decays. Because the FEI is a machine learning algorithm trained on MC data, an accurate model of the decays of interest is crucial. Inaccuracies in the models used in MC generation can lead to bias, which then needs to be calibrated (as done for hadronic tagging in this thesis), leading to significant systematic uncertainties. A precise measurement of $B \to D^* \ell \nu_\ell$ and its form factors could help to reduce the bias and minimize calibration issues.

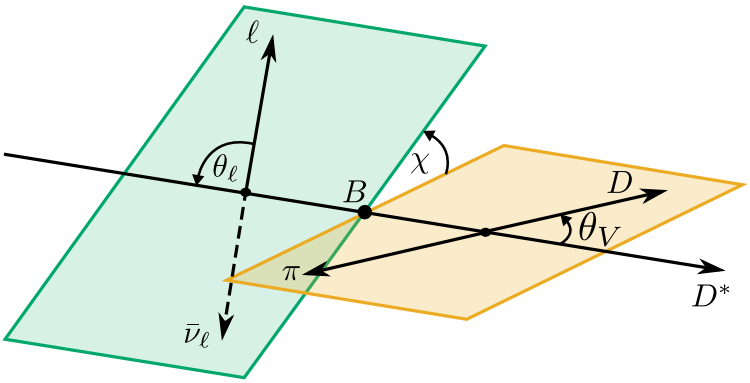


Figure 1.12: Definition of the helicity angles θ_ℓ, θ_V, and χ in the $\overline{\mathrm{B}} \to \mathrm{D}^*(\to \mathrm{D}\pi)\ell\bar{\nu}_\ell)$ decay. The green plane is spanned by the momenta of the leptons ℓ and $\bar{\nu}_\ell$. θ_ℓ is the angle between the momentum of the charged lepton and the opposite D^* momentum (seen from the $\ell\bar{\nu}_\ell$ rest frame). The yellow plane is spanned by the D and π momenta. θ_V is the angle between the D and D^* momenta (seen from the $\mathrm{D}\pi$ rest frame). Finally, χ is the angle between the green and yellow planes.

In our analysis, the D^* meson is reconstructed in the $\mathrm{D}^* \to \mathrm{D}\pi$ decay modes. The reconstruction modes and the general analysis strategy will be presented in more detail in Chapter 5. In addition to the distribution of w, we will measure the distributions of the three helicity angles defined in Figure 1.12.

In the limit of vanishing lepton mass, the differential cross section of this decay is given by

$$\begin{aligned}\frac{\mathrm{d}^4\Gamma(\overline{\mathrm{B}} \to \mathrm{D}^*(\to \mathrm{D}\pi)\ell\bar{\nu}_\ell))}{\mathrm{d}w\,\mathrm{d}\cos\theta_V\,\mathrm{d}\cos\theta_\ell\,\mathrm{d}\chi} = \frac{6m_{\mathrm{B}}m_{\mathrm{D}^*}^2}{8(4\pi)^4}\sqrt{w^2-1}\Big(1-2wr+r^2\Big)G_{\mathrm{F}}^2\eta_{\mathrm{EW}}^2|V_{\mathrm{cb}}|^2 \times \mathrm{BR}(\mathrm{D}^* \to \mathrm{D}\pi)\times \\ \times\Big[(1-\cos\theta_\ell)^2\sin^2\theta_V H_+^2 + (1+\cos\theta_\ell)^2\sin^2\theta_V H_-^2 + \\ + 4\sin^2\theta_\ell\cos^2\theta_V H_0^2 - 2\sin^2\theta_\ell\sin^2\theta_V\cos 2\chi H_+H_- + \\ - 4\sin\theta_\ell(1-\cos\theta_\ell)\sin\theta_V\cos\theta_V\cos\chi H_+H_0 + \\ + 4\sin\theta_\ell(1+\cos\theta_\ell)\sin\theta_V\cos\theta_V\cos\chi H_-H_0\Big],\end{aligned} \tag{1.5.1}$$

where the three helicity amplitudes are defined as

$$H_\pm = \left[(m_{\mathrm{B}}+m_{\mathrm{D}^*})r\frac{w+1}{2} \mp \frac{2}{r(1+r)}|\vec{p}_{\mathrm{D}^*}|R_1(w)\right]h_{A_1}(w), \tag{1.5.2}$$

$$H_0 = \frac{1}{2m_{\mathrm{D}^*}\sqrt{q^2}}\left[(m_{\mathrm{B}}^2-m_{\mathrm{D}^*}^2-q^2)(m_{\mathrm{B}}+m_{\mathrm{D}^*})\frac{w+1}{2}r - \frac{4m_{\mathrm{B}}^2}{r(m_{\mathrm{B}}+m_{\mathrm{D}^*})}|\vec{p}_{\mathrm{D}^*}|^2R_2(w)\right]h_{A_1}(w).$$

A full expression in terms of Wilson coefficients that parametrizes possible NP effects and accounts for non-vanishing lepton mass is given in [123]. It is also discussed in [124, App. A].

As we can see from Equations (1.5.1) and (1.5.2), the different form factors (here given in terms of h_{A_1}, R_1 and R_2) enter the differential cross section in linearly independent combinations of the angular terms. Therefore, measuring the angular distributions allows to constrain all form factors.

Chapter 2

Clustering of $B^+ \to \overline{D}^{(*)0}\tau^+\nu_\tau$ kinematic distributions with ClusterKinG

[1]New Physics can manifest itself in kinematic distributions of particle decays. The parameter space defining the shape of such distributions can be large, which is challenging for both theoretical and experimental studies. However, using clustering algorithms, the parameter space can be dissected into subsets (clusters) corresponding to similar kinematic distributions. Clusters can then be represented by benchmark points, which allow for less involved studies and a concise presentation of the results.

To make these techniques more accessible in a High Energy Physics context, I have developed the Python package `ClusterKinG`, an easy-to-use framework for the clustering of kinematic distributions.

As an example, $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ distributions are considered and possible implications for future experimental analyses discussed.

2.1 Introduction

New Physics (NP) contributions can influence the kinematic distributions of particle decays. While this opens up possibilities to find and determine the nature of NP, it can also be a nuisance for experimental studies because most measurements require assumptions on certain kinematic distributions, e.g., to distinguish signal from background or to determine efficiencies.

For example, assuming a two-Higgs-doublet model of type II changes the experimental measurement of $\mathcal{R}(D^{(*)})$ because discriminating between signal and background requires assumptions on the kinematic shapes of the signal, background, and normalization modes [71]. Such kinematic shapes are generally determined from Monte Carlo simulations.

Thus, many experimental measurements are model-dependent and are often only conducted under the assumption of the Standard Model (SM). Discrepancies between the SM prediction and the measured values are a good indication for NP. However, comparing different NP models based on their predicted results has to be taken with a grain of salt because the measurements themselves are model-dependent.

A further complication for both theoretical and experimental studies is the high dimensionality of the parameter space of typical NP models. If experimentalists wish to publish model-independent results, the studies must be repeated for a large sample of parameter points. This can be

[1]The results of this chapter have been published in JHEP [125], see also page 157. By and large, the content of this chapter is a verbatim copy of this paper. Exceptions are instances of redaction: restructuring the content, changing the wording, and extending some of the explanations. Because this project includes contributions of my co-authors, the individual work of all collaborators is listed separately in Appendix A.4.

computationally very expensive. Furthermore, numerical results and many visualizations can only be shown for specific (often arbitrary) parameter points, leaving their representative character unknown.

A possibility to reduce the complexity of this problem is to identify regions in the parameter space that lead to similar kinematic distributions. These regions can be found using clustering algorithms. From each cluster, a most representative point (*benchmark point*, BP) is chosen.

Experimental studies can focus on these BPs, thereby reducing the multi-dimensional problem to a small number of BPs to be considered. The results are then presented for each BP, allowing for a clear-cut numerical result and simpler visualizations.

Such a strategy has been employed for the first time in the context of Higgs boson pair production in [126–128]. An Effective Field Theory (EFT) approach has been adopted to parametrize the five-dimensional parameter space of anomalous Higgs couplings. It is shown that for current and future searches at the Large Hadron Collider (LHC), a total of 12 clusters gives a reasonable approximation of the considered parameter space [126]. In [127] several clusters are subjected to experimental limits from the CMS collaboration [129]. Finally, in [128] a method to extend the experimental sensitivity from the BPs to the other cluster members is discussed.

In recent years, substantial progress has been made in the EFT description of the SM in the form of the SM Effective Field Theory (SMEFT) [130] and the Weak Effective Theory (WET). The calculation of the complete one-loop SMEFT Renormalization Group Equations (RGEs) [131–134], the complete tree-level and one-loop matching from SMEFT onto WET [135–137] and the complete one-loop QCD and QED RGEs within WET [138, 139] allow for a general NP analysis of low-energy observables.

Various tools are dedicated to studying Wilson coefficients above and below the electroweak scale. They include the Match(runn)ers `DsixTools` [134, 140], `MatchingTools` [141], and `wilson` [142], the Wilson coefficient exchange format `WCxf` [143], the fitting tool `smelli` [144], basis codes like `BasisGen` [145], `DEFT` [146], and `abc-EFT` [147], the observable calculator `flavio` [148], `FlavorKit` [149], `SuperIso` [150], `SPheno` [151, 152], `FormFlavor` [153], as well as packages related to SMEFT, such as `SMEFT Feynman Rules` [154] and `SMEFTsim` [155].

However, public tools to cluster the phenomenology of NP operators systematically are still missing so far.

To fill this gap, we have written the Python package `ClusterKinG` (Clustering of Kinematic Graphs), which aims to make clustering techniques accessible in the context of EFTs and High Energy Physics in general. Despite this motivation, `ClusterKinG` is a general tool that can be applied to a diverse set of problems, even outside of physics.

We demonstrate the use of this package by clustering kinematic distributions of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ decays. These decays are of particular interest in view of the current B anomalies [156–164].

The rest of the chapter is organized as follows: In Section 2.2 we discuss the clustering method in general terms. In Section 2.3 we present the `ClusterKinG` package and describe its features. In Section 2.4 we apply the clustering method to kinematic distributions of the decays $\bar{B} \to D^{(*)}\,\tau^-(\to \ell\,\bar{\nu}_\ell\,\nu_\tau)\bar{\nu}_\tau$ and perform various consistency tests. Finally we conclude in Section 2.5.

2.2 Clustering

This section discusses the different steps involved in our clustering approach. As a first step, a suitable observable and the corresponding model of interest have to be chosen. After establishing the underlying parameter space, the following steps will be performed:

- Sampling the parameter space of the process.
- Computing the kinematic distributions for the points in the parameter space.
- Choosing a metric to measure differences between the kinematic distributions.
- Applying a suitable clustering algorithm on the parameter space.
- Selecting the BPs representing each of the clusters.

The above steps are explained in the following subsections.

2.2.1 Sampling of the parameter space

As discussed in the introduction, typical NP models depend on several parameters. Theoretical considerations (such as symmetry arguments) can often be used to limit the study to a subset of these parameters, thereby reducing the dimensionality of the problem. The considered range of these parameters can be motivated by existing exclusion limits or from theory.

From this subset of the original parameter space, *sample points* are chosen and used for the rest of the analysis. While large numbers of sample points will make the following steps more computationally expensive, it is important that the sampling is fine enough to represent the whole parameter space accurately.

`ClusterKinG` allows for an arbitrary selection of sampling points. The examples presented in this chapter use a uniform grid in the parameter space for simplicity. In order to limit the number of required points (and thereby computing time), it is also planned to implement adaptive sampling techniques in the future: After an initial run with a coarse grid, regions with significant variations in the kinematic distributions are identified and sampled in a finer grid. If needed, this procedure can then be applied several times.

2.2.2 Kinematic distributions

For every sample point, the corresponding kinematic distribution needs to be computed. If analytic expressions of the observable in terms of the parameters are available, this task can be achieved by evaluating the formulae. Otherwise, Monte Carlo (MC) simulations have to be used to generate the distributions. Since the generation of MC samples is generally resource-intensive, reweighting techniques can be used to generate samples corresponding to different parameter points from existing samples. For semileptonic B decays such methods are already implemented in the `HAMMER` tool [165, 166].

In this chapter, we only consider *binned* observables, and our kinematic distributions are thus histograms.

2.2.3 Metric

The objective of the clustering procedure is to partition the parameter space in such a way that parameter points generating similar kinematic distributions are grouped into the same cluster. For this, the "similarity" between kinematic distributions has to be quantified by a metric[2] in the space of kinematic distributions.

The choice of this metric follows from the interpretation of the distributions as potential measurements. As such, the metric of choice should give more weight to differences in bins associated with low expected experimental uncertainties while being less sensitive to differences

[2]Now and in the following, the term "metric" is used in a rather loose way, emphasizing the intuition of a distance measure while not necessarily fulfilling all requirements to be a metric in the mathematical sense.

in bins of less experimental significance. Estimating experimental uncertainties is also useful when deciding the number of clusters and benchmark points to choose: sufficiently many to cover the whole variety of distributions that lead to different experimental results, but not arbitrarily many. In this way, the number of clusters then also serves as an estimate for the sensitivity of a distribution to NP parameters.

A common choice for a metric measuring the similarity between binned distributions is a two-sample test statistic such as the Kolmogorov-Smirnov test [167], the Anderson-Darling test [168] or the χ^2 test.

In [126] a binned log-likelihood test statistic is used to distinguish between two distributions. This likelihood ratio is obtained by taking the ratio of the binomial distribution of the two individual samples and the binomial distribution, where both samples are assumed to be equal. The logarithm of this ratio can be shown to be χ^2-distributed up to a minus sign [169]. By basing the test statistic on binomial distributions, the metric incorporates the statistical significance of the different bins.

In this chapter, we use a χ^2 test operating on normalized distributions with uncertainties. As the distributions are not measured but generated (as described in Section 2.2.2), an uncertainty estimate is applied to them. It consists of configurable statistical uncertainties as well as relative and absolute systematic uncertainties that can be correlated between bins.

Let n_{ki} be the bin contents of two histograms H_k ($k = 1, 2$, $i = 1, \ldots, N$). Our null hypothesis is that the bin contents of the histograms are drawn from two distributions with identical means. We assume that the n_{ki} are distributed according to a multivariate normal distribution with covariance matrices $\Sigma_k = \left(\mathrm{Cov}(n_{ki}, n_{kj})\right)_{ij}$. We denote the corresponding normalizations as $N_k = \sum_{i=1}^{N} n_{ki}$, and define $\Delta_i = \frac{n_{1i}}{N_1} - \frac{n_{2i}}{N_2}$ and $\Sigma = \frac{\Sigma_1}{N_1^2} + \frac{\Sigma_2}{N_2^2}$. Our χ^2 measure is then given by

$$\chi^2(H_1, H_2) = \sum_{i,j=1}^{N} \Delta_i (\Sigma^{-1})_{ij} \Delta_j. \tag{2.2.1}$$

Under the null hypothesis, $\chi^2(H_1, H_2)$ approximates a χ^2 distribution with $N - 1$ degrees of freedom, henceforth denoted as χ^2_{N-1}.

It should be highlighted that this approximation can break down if the uncertainties are very imbalanced, though this does not usually happen if Poisson uncertainties are the dominant uncertainties. Appendix A describes toy studies that were used to validate the statistical treatment for all results shown in Section 2.4

In the following, we call two distributions distinguishable if their χ^2-distance $\chi^2(H_1, H_2)/(N-1)$ is larger than 1.125, corresponding to a p-value of 34% for $N = 9$.[3] Figure 2.1 shows the relationship between this *cutoff value*, the p-value, and the number of bins. This loose definition of distinguishability is conservative in the sense that it will lead to more clusters than a stricter criterion. The metric between the kinematic distributions gives rise to a metric d acting directly on the parameter space. In our case, we define

$$d(c_1, c_2) \equiv \chi^2(H_1, H_2)/(N-1) \tag{2.2.2}$$

for two sample points $c_{1,2}$ and their respective histograms $H_{1,2}$.

2.2.4 Clustering algorithm

In general a dataset can be either clustered *hierarchically* or *partitionally*. Partitional clustering methods [170] such as, for example, K-means algorithms [171] only perform one single partition

[3] The number of 1.125 was chosen for consistency with a previous version of the paper which incorrectly assumed N degrees of freedom.

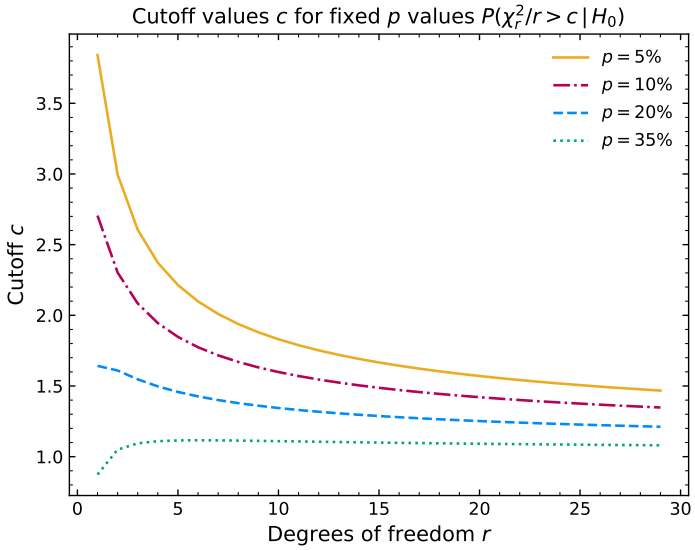


Figure 2.1: p-values and corresponding cutoff values.

of the data. Furthermore, the number of resulting clusters must be chosen beforehand as an input parameter. In the following, we focus on hierarchical clustering methods [172].

Hierarchical clustering algorithms group a given dataset in a nested sequence of clusters. Two approaches are common: Bottom-up (or agglomerative) algorithms successively merge clusters to form larger clusters, whereas top-down algorithms successively split clusters into smaller ones. In both cases, a stopping criterion is needed to avoid a trivial result. In our analysis, we will employ the agglomerative method with the following steps:

1. Associating each sample point to one cluster containing only this element.
2. Merging the nearest two clusters according to a metric D.
3. Repeating step 2 until the stopping criterion is fulfilled.

Note that the metric D in step 2 is not between points in the parameter space but between subsets (existing clusters) of this space. It makes sense to base D on the metric d introduced in Equation (2.2.2). Two canonical choices for the inter-cluster metric D are

$$D_\infty(C_1, C_2) \equiv \max_{c_1 \in C_1,\, c_2 \in C_2} d(c_1, c_2)\,, \quad \text{and} \quad D_1(C_1, C_2) \equiv \frac{1}{|C_1||C_2|} \sum_{c_1 \in C_1,\, c_2 \in C_2} d(c_1, c_2)\,, \tag{2.2.3}$$

where C_1 and C_2 are clusters, and $|C_{1,2}|$ denote their number of elements.

While [126] uses the 'average' metric D_1, we will employ the D_∞ metric in our analysis. This choice is more conservative in the sense that it usually leads to a larger number of clusters than in the case of D_1, simply because $D_\infty \geq D_1$.

The stopping criterion is chosen to be $D_\infty(C_1, C_2) > 1.125$ for all pairs of clusters. This means that if we merge two clusters, the resulting larger cluster does not contain any two distinguishable points. Consequently, all clusters of our final result contain only indistinguishable sample points. This is not the case for the 'average' metric in general.

2.2.5 Benchmark points

After the application of the clustering algorithm, BPs have to be determined for all of the resulting clusters. A BP is a cluster element c of a cluster C, which is chosen as a representative

of that particular cluster. Usually, it is taken to be the parameter point that minimizes a certain figure of merit, which is commonly based on the metric d of Equation (2.2.2). Examples are:

$$f_1(c, C) \equiv \frac{1}{|C|} \sum_{c_i \in C} d(c, c_i), \quad f_2(c, C) \equiv \sqrt{\sum_{c_i \in C} d(c, c_i)^2}, \quad f_\infty(c, C) \equiv \max_{c_i \in C} d(c, c_i), \tag{2.2.4}$$

which differ in their responsiveness to outliers in the data. The BPs are the key elements of the cluster analysis. They are determined to simplify experimental analyses, which can then lay their focus only on a finite set of BPs instead of the entire parameter space.

2.3 The ClusterKinG package

`ClusterKinG` is publicly developed as open-source software under the MIT license [173, 174]. The package aims for ease of use while also allowing users to manipulate and extend functionality.

The basic building blocks of `ClusterKinG` are worker classes: After initialization, a set of methods can be called for further configuration, before calling a `run` method that performs the desired action. By subclassing these worker classes, the functionality can be extended. Furthermore, these worker classes can be passed on to other worker classes to perform, e.g., stability checks that require repetitive calling with slight parameter variations.

To demonstrate the ease of use, a fully working example of code to generate clusters, benchmark points, and plots similar to the ones presented in Section 2.4 is shown in Appendix A.1.

A typical `ClusterKinG` workflow consists of the following steps:

1. Initialize a `Data` object: This is what the following worker classes read from and write to.

   Internally, this class holds a `pandas DataFrame` [175] that contains the bin contents of the kinematic distributions for each sample point. Additional information (*metadata*) is saved as a nested dictionary. Each of the following steps will add information to the metadata such that any output file contains a description of how it was generated.

   `Data` objects can be saved to and loaded from a single output file in SQL format, allowing to save all data in a single output file of comparably small size. Exports of the data to other formats (CSV, XLS, JSON, ...) are readily available.

2. Adding uncertainties (optional): After the distributions have been generated, an experimental uncertainty estimate can be added. Typically this consists of

   - statistical uncertainties, modeled as Poisson errors (and thereby dependent on the actual bin content for the distribution for each sample point), and
   - systematic uncertainties which can be given in absolute form or relative to the bin content for all sample points at once.

   To save memory space and improve performance, the uncertainties on the distributions are only calculated when needed, and only the procedure of *how* they are calculated is saved to the `Data` object. This also means that it is straightforward and fast to check the dependency of the clustering results on the uncertainties.

3. Scanning: (Kinematic) distributions are generated for the chosen sample points. This is done by a `Scanner` object or for convenient clustering in the space of Wilson coefficients its `WilsonScanner` subclass.

   Generally, this requires three steps:

   a) Providing a function to generate kinematic distributions. Any Python function can be used, in particular any observable from the `flavio` [148] package.

b) Defining the sample points, usually in the form of an equidistant grid in the parameter dimensions (though more fine-tunable methods are available). The input of `WilsonScanner` is given in the common `WCxf` format defined in [143]. The Wilson coefficients are specified at a certain scale and in a particular EFT and basis.

c) Running the scanning process. `ClusterKinG` supports multiprocessing to speed up calculations.

The resulting distributions and the spread of the bin contents among the sample points are visualized by calling some of the plot methods of the `Data` object.

4. Clustering: Different clustering algorithms correspond to different subclasses of the `Cluster` class. Hierarchical clustering is implemented in the `HierarchyCluster` worker which internally uses algorithms of the `scipy` library [176] and can be used with a range of different metrics (p-metrics, χ^2 metric, user defined functions).

5. Selection of benchmark points: This step is performed by the `Benchmark` class, which can be configured (and subclassed) to allow for different strategies.

Besides visualization, the data class also provides simple methods, such as to find the closest benchmark or sampling point given a point in parameter space.

Extensive technical documentation of this package is available online [177], and multiple usage examples are provided in the form of Jupyter notebooks in the main repository. For demonstration purposes, the notebooks can be run directly in the browser without installing any packages.

2.4 Clustering of $B \to D^{(*)}\tau\bar{\nu}$ distributions

2.4.1 Setup

In this subsection, we describe the setup used for our numerical analysis. Motivated by current B anomalies, we perform a clustering analysis on various $\bar{B} \to D^{(*)}\,\tau^-\bar{\nu}_\tau$ kinematic distributions. Such $b \to c$ transitions are described by the following effective Lagrangian:

$$\mathcal{L}_{\text{eff}} = -\frac{4G_F}{\sqrt{2}} V_{cb}[C_{VL}O_{VL} + C_{VR}O_{VR} + C_{SL}O_{SL} + C_{SR}O_{SR} + C_T O_T] + \text{h.c.}\,, \tag{2.4.1}$$

with the CKM matrix element V_{cb}, the Fermi coupling constant G_F, Wilson coefficients $C_{VL,VR,SL,SR,T}$, and the effective operators

$$\begin{aligned} O_{VL} &= (\bar{c}\gamma_\mu \mathcal{P}_L b)(\bar{\tau}\gamma^\mu \mathcal{P}_L \nu_\tau)\,, & O_{VR} &= (\bar{c}\gamma_\mu \mathcal{P}_R\, b)(\bar{\tau}\gamma^\mu \mathcal{P}_L \nu_\tau)\,, \\ O_{SL} &= (\bar{c}\mathcal{P}_L b)(\bar{\tau}\mathcal{P}_L \nu_\tau)\,, & O_{SR} &= (\bar{c}P_R b)(\bar{\tau}\mathcal{P}_L \nu_\tau)\,, \\ O_T &= (\bar{c}\sigma_{\mu\nu}\mathcal{P}_L b)(\bar{\tau}\sigma^{\mu\nu}\mathcal{P}_L \nu_\tau)\,. \end{aligned} \tag{2.4.2}$$

We use the notation $\mathcal{P}_{R,L} = \frac{1}{2}(\mathbb{1} \pm \gamma_5)$, $\sigma^{\mu\nu} = \frac{i}{2}[\gamma^\mu, \gamma^\nu]$ and b, c, τ, ν_τ for the quark and lepton fields. The Wilson coefficients in Equation (2.4.1) are in general complex quantities. For our analysis, we will assume the $\mathcal{CP}$ conserving limit, i.e., real Wilson coefficients. This is a common assumption[4] (see for example [157, 158]) and is mainly chosen for simplicity. Furthermore, for presentational reasons, we study three out of the five Wilson coefficients and choose one for each Dirac structure, namely:

$$C_{VL},\, C_{SL},\, C_T\,. \tag{2.4.3}$$

We will assume values for the first two Wilson coefficients in the interval $[-0.5, 0.5]$. The tensor operator is constrained [144] from the longitudinal polarization fraction F_L in $B \to D^{*0}\tau\nu$ [183] and we choose its Wilson coefficient to be in the interval $[-0.1, 0.1]$. For our analysis, we have

[4] Studies with $\mathcal{CP}$ violating contributions are analyzed in [178–182].

BP	C_{VL}	C_{SL}	C_T
0	−0.50	−0.17	−0.08
1	0.39	−0.28	−0.06
2	−0.50	−0.06	0.10
3	−0.17	−0.06	0.10
4	0.39	0.17	0.10
5	0.06	−0.06	0.03

Table 2.1: List of benchmark points for the distribution $d\mathrm{BR}(B \to D^{0*}\tau\nu)/d(\cos\theta_\tau)$ obtained from the clustering procedure given in terms of the left-handed vector, left-handed scalar and tensor Wilson coefficients C_{VL}, C_{SL} and C_T.

chosen an equidistant grid of 1000 sample points in the three-dimensional parameter space, where each of the Wilson coefficients lies within the specified intervals.

The clustering is performed using the `ClusterKinG` package. As mentioned in Section 2.2, we use the hierarchical clustering algorithm together with the χ^2-metric defined in Section 2.2. The stopping criterion is chosen such that the χ^2-distance between all distributions within the same cluster is ≤ 1.125, meaning that they are indistinguishable experimentally. Finally, the BPs are obtained by adopting the figure of merit f_1 from Equation (2.2.4).

The complete code that has been used for the generation of the results and plots below is provided in the `example` directory of the `ClusterKinG` repository together with usage instructions.

2.4.2 Results

2.4.2.1 $\cos\theta_\tau$-distribution of $\mathrm{BR}(B \to D^{0*}\tau\nu)$

As a first example, we consider nine bins of the $\cos\theta_\tau$-distribution of the branching ratio $\mathrm{BR}(B \to D^{0*}\tau\nu)$, where θ_τ denotes the angle between the tauon and the B meson in the dilepton mass frame. The kinematic distributions are generated using the `flavio` [148] package. With an assumed systematic uncertainty of 1 % and statistical uncertainties corresponding to a yield of 700 events, our clustering procedure leads to a total of six clusters and their corresponding BPs.

The clustered parameter space is shown as two dimensional cuts in the C_T-direction in Figure 2.2 and numeric values for the BPs are reported in Table 2.1. As can be seen in Figure 2.2, the parameter space exhibits a strong cluster variation in the direction of C_T. This fact is not surprising, considering the explicit dependence of the kinematic distribution, and it agrees with the findings of [184], where this "flat term" observable has been proposed in the context of charged $b \to c$ transitions involving light leptons.

The distributions corresponding to the sample points are visualized as a box plot in Figure 2.3. As expected, different clusters correspond to significantly different distributions. Furthermore, the distributions of the benchmark points are similar to the distributions given by the mean values of the bin contents of all distributions of the corresponding cluster.

2.4.2.2 $\cos\theta_V$-distribution of $\mathrm{BR}(B \to D^{0*}\tau\nu)$

As a second example we consider the $\cos\theta_V$-distribution of the process $B \to D^{0*}\tau\nu$. Here θ_V denotes the angle between D^{0*} and the B meson. The kinematic distributions for this process are again generated using `flavio` [148]. Assuming a signal yield of 700 events and a relative systematic uncertainty of 1 %, the clustering procedure leads to three clusters. The clustered

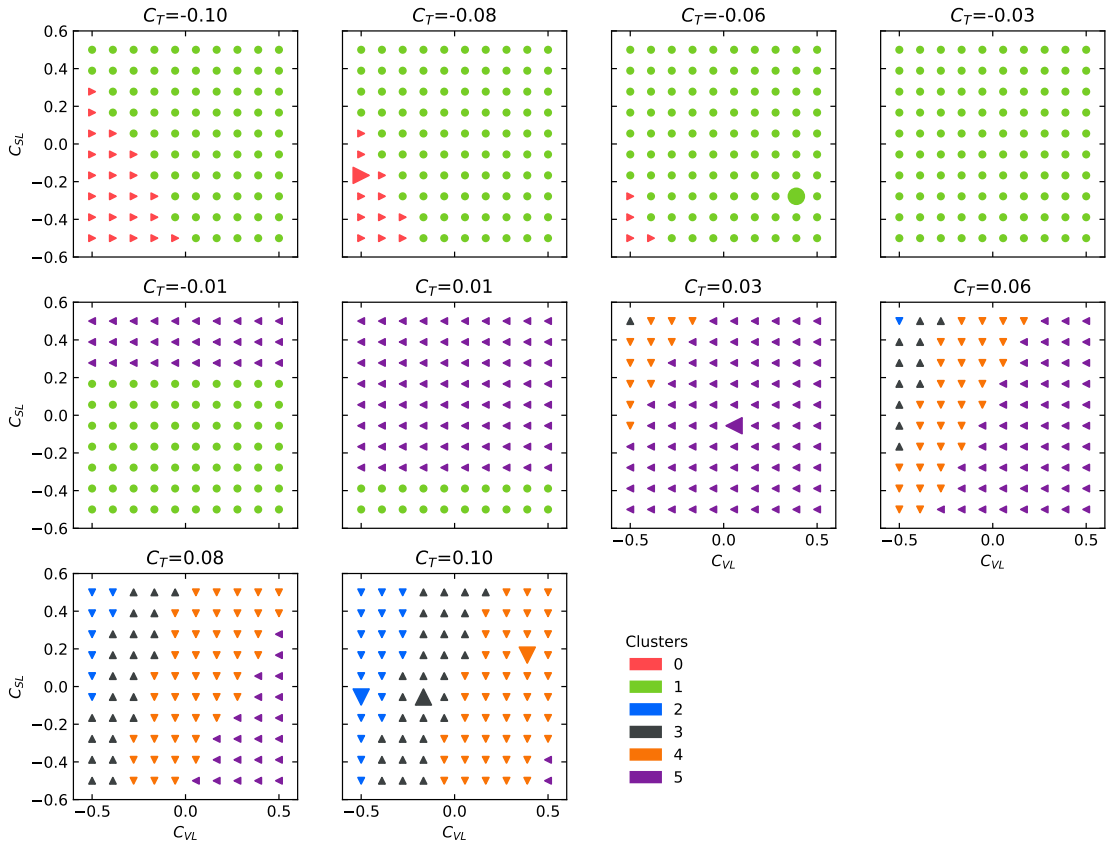


Figure 2.2: Two-dimensional cuts of the clustered parameter space resulting from dBR(B → D$^{0*}\tau\nu$)/d(cos θ_τ) distributions. The parameter space is spanned by the three Wilson coefficients C_{VL}, C_{SL} and C_{T}. Six different clusters are found, which are indicated with different markers and colors. BPs are given in boldface.

three-dimensional sample space is shown in Figure 2.4 and C_{T} can again be identified to be the most influential Wilson coefficient. A large subset of the parameter space belongs to cluster 2 (blue), whereas only a few sample points are contained in the first cluster (red), which is found at the edges of the sample space. The three BPs are reported in Table 2.2.

Finally, we show several example distributions for each cluster together with the BP distributions in Figure 2.5. While the distinction between the red cluster and the other two clusters is very clear, the kinematic distributions of the blue and green clusters are more similar.

Compared to θ_τ, fewer clusters are found, but the shapes of the clusters are different from the previous ones. The two observables can thus be considered complementary in their respective sensitivity to NP models.

2.4.2.3 q^2-distribution of BR(B → D$^0\tau\nu$)

The q^2-distribution of BR(B → D$^0\tau\nu$) has already been studied extensively in the literature. For our purpose, we consider this observable to study the influence of the systematic and statistical uncertainties on the resulting number of clusters. This is relevant for future experiments such as Belle II that reach for new luminosity records. The q^2-distributions were computed using `flavio` [148]. In Figure 2.6 we show the number of clusters as a function of the signal yield for various systematic uncertainties. As expected, the number of resulting clusters is dominated by the signal yield for small yields, while systematic uncertainties become the limiting factor for larger yields.

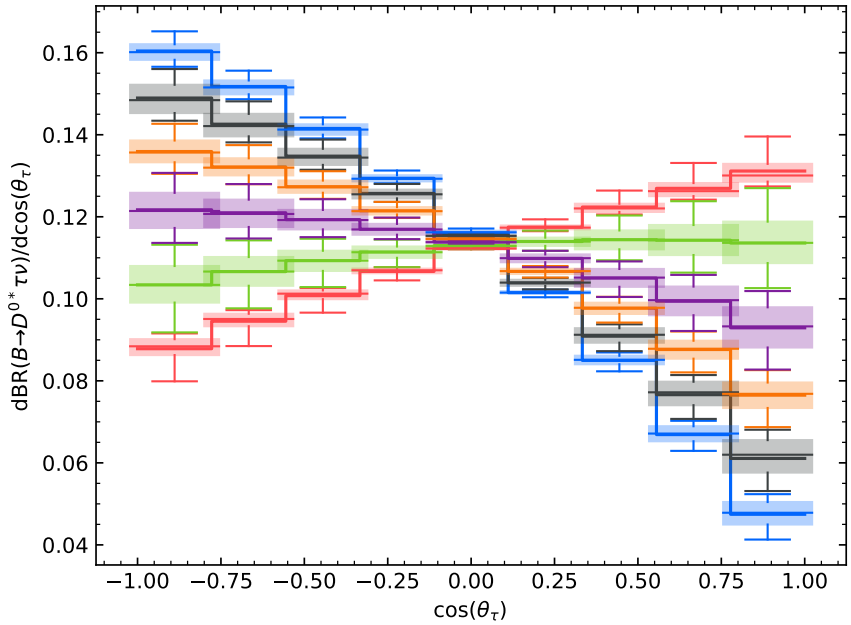


Figure 2.3: The distributions of the observable $\mathrm{dBR(B \to D^{0*}\tau\nu)/d(\cos\theta_\tau)}$ for the six different clusters (with colors matching these of Figure 2.2). The histograms corresponding to the BPs are shown as solid lines. The boxes extend from the upper to the lower quartile of the distribution of the bin contents within a cluster and a horizontal line indicates the median. Whiskers are used to further indicate the span of the data, covering six times the interquartile range. Points beyond this range are plotted as individual points (outliers).

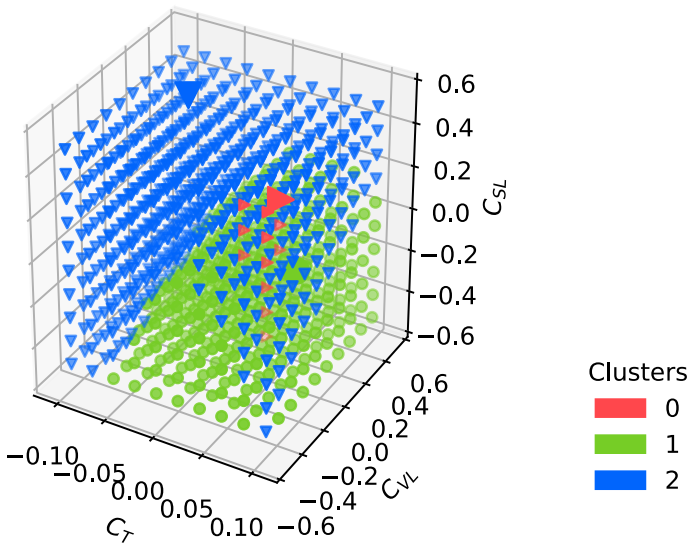


Figure 2.4: The clustered three-dimensional parameter space resulting from $\mathrm{dBR(B \to D^{0*}\tau\nu)/d(\cos\theta_V)}$ distributions. The parameter space is spanned by the three Wilson coefficients C_{VL}, C_{SL}, and C_{T}, varied with their respective ranges. Three different clusters are found in our approach, which are indicated with different markers and colors. BPs are given in boldface.

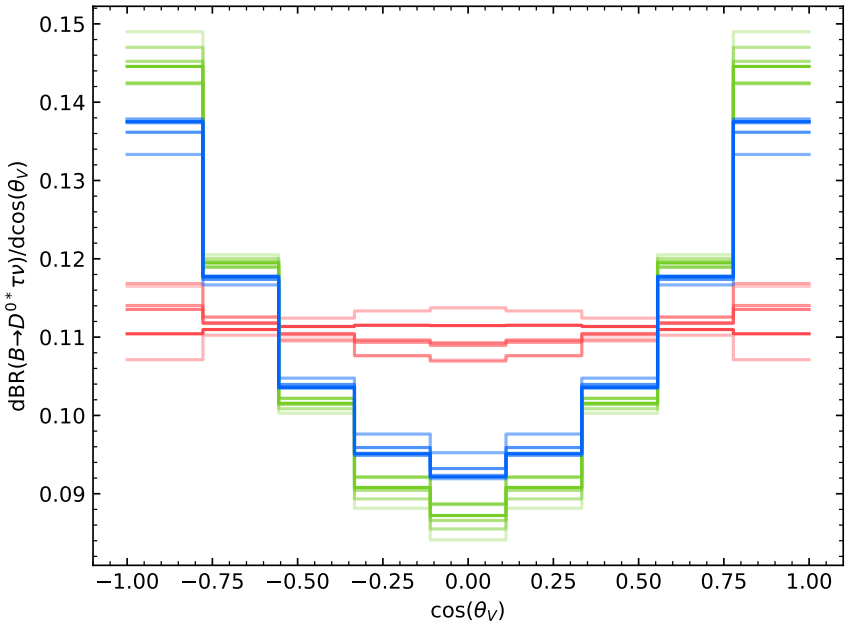


Figure 2.5: $\cos\theta_V$-distributions of the decay $\bar{\mathrm{B}} \to \mathrm{D}^{0*}\tau^-\bar{\nu}_\tau$ corresponding to three sample points and the BPs. Distributions corresponding to the same cluster are shown in the same color (matching the color scheme of Figure 2.4), with the sample distributions faded out slightly.

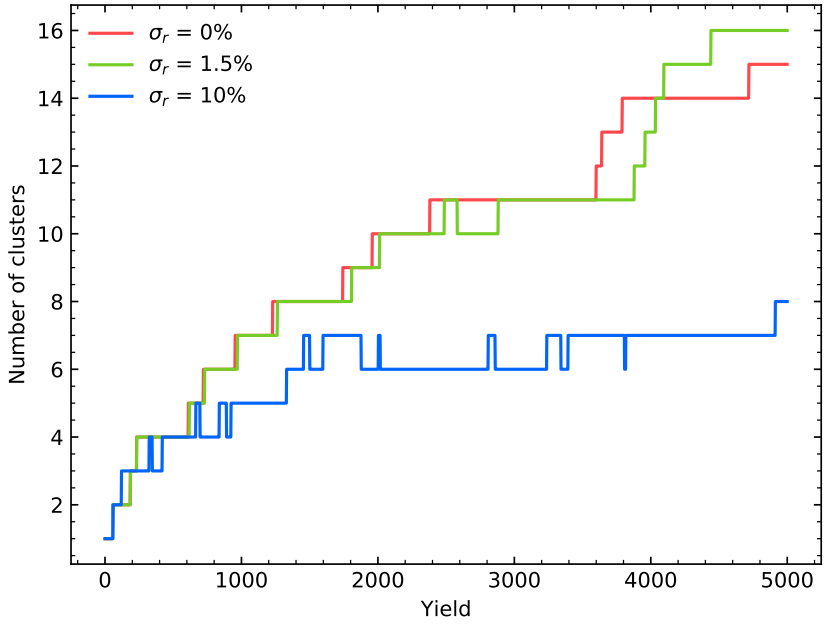


Figure 2.6: The number of clusters as a function of the signal yield for the observable $\mathrm{dBR}(\mathrm{B} \to \mathrm{D}^0\tau\nu)/\mathrm{d}q^2$ for three different relative systematic uncertainties σ_r.

BP	C_{VL}	C_{SL}	C_T
0	−0.39	0.50	0.10
1	0.17	−0.17	0.06
2	0.17	0.50	−0.06

Table 2.2: List of benchmark points for the distribution $dBR(B \to D^{0*}\tau\nu)/d(\cos\theta_V)$ obtained from the clustering procedure given in terms of the left-handed vector, left-handed scalar, and tensor Wilson coefficients C_{VL}, C_{SL} and C_T.

2.4.2.4 E_ℓ-distribution of $\Gamma(\bar{B} \to D\tau^-(\to \ell\,\bar{\nu}_\ell\,\nu_\tau)\bar{\nu}_\tau)$

Finally, we consider E_ℓ, the energy of the light lepton ℓ from the 5-body decay $\bar{B} \to D\tau^-(\to \ell\,\bar{\nu}_\ell\,\nu_\tau)\bar{\nu}_\tau$. To generate the kinematic distributions in terms of the chosen set of Wilson coefficients, we use the explicit expressions given in [185] as further outlined in Appendix A.2.

The study of this observable is motivated by the BaBar analysis in [71], where the experimental value of $\mathcal{R}(D)$ was extracted under the assumption of a two-Higgs-doublet model. Signal and background yields were extracted with a fit to the two-dimensional $m^2_{\text{miss}}, |\vec{p}^*_\ell|$ distribution. Here, m_{miss} and $|\vec{p}^*_\ell|$ denote the mass of the undetected neutrinos and the three-momentum of the light lepton in the rest frame of the B meson. The shape of these distributions for signal and background were taken from MC simulations. Such MC simulations are usually assumed to be SM-like. In [71] however, the (SM) MC simulations were reweighted assuming non-zero values for the parameter $p \equiv \tan\beta/m_{H^\pm}$ of the two-Higgs-doublet model of type II[5]. The (model-dependent) experimental values of $\mathcal{R}(D)$ were then extracted for 20 different values of p. The resulting distribution of $\mathcal{R}(D)$ measurements shows a sharp transition at $p = 0.35\,\text{GeV}^{-1}$, but is otherwise relatively independent of p.

This result motivates a clustering analysis to investigate the model dependency of the input kinematics.

In this chapter, we consider the E_ℓ-distribution, which can be taken as an approximation of $|\vec{p}^*_\ell|$ as considered in [71]. The results of our clustering analysis with respect to the parameter p are shown in Figure 2.7. The one-dimensional parameter space is clustered three times, assuming signal yields of 1000, 1800, and 2000 events as well as a relative systematic uncertainty of 1 %. The first sub-figure shows the two resulting clusters for a yield of 1000 events, which coincide with the findings of [71], where two different values of $\mathcal{R}(D)$ are obtained. However, in [71] the first value for $\mathcal{R}(D)$ is obtained for $p \leq 0.3\,\text{GeV}^{-1}$ and the second one for $p \geq 0.45\,\text{GeV}^{-1}$, whereas the first sub-figure suggests to have the same $\mathcal{R}(D)$ for $0.3\,\text{GeV}^{-1} \leq p \leq 0.7\,\text{GeV}^{-1}$ and another value for the rest of the parameter space.

Increasing the yield (and thereby reducing the uncertainties on the distributions) results in more clusters in the middle region (see Figure 2.7), again indicating that the shape of the kinematic distribution significantly changes between $0.3\,\text{GeV}^{-1} < p < 0.7\,\text{GeV}^{-1}$. However, as can be seen from Figure 2.8, the kinematics for low and high p are still very similar, incompatible with the result of [71].

On the other hand, since [71] used the m^2_{miss} distribution together with the $|\vec{p}^*_\ell|$ distribution, it is not too surprising to arrive at a rather different result, as the shape of the distributions in m^2_{miss} could behave very differently than $|\vec{p}^*_\ell|$. Applying clustering techniques to the 2D $m^2_{\text{miss}}, |\vec{p}^*_\ell|$ distribution will allow for a more thorough comparison and is left for future work.

[5] $\tan\beta$ denotes the ratio of the vacuum expectation values of the two Higgs bosons and $m_{H^\pm}$ is the mass of the charged Higgs boson.

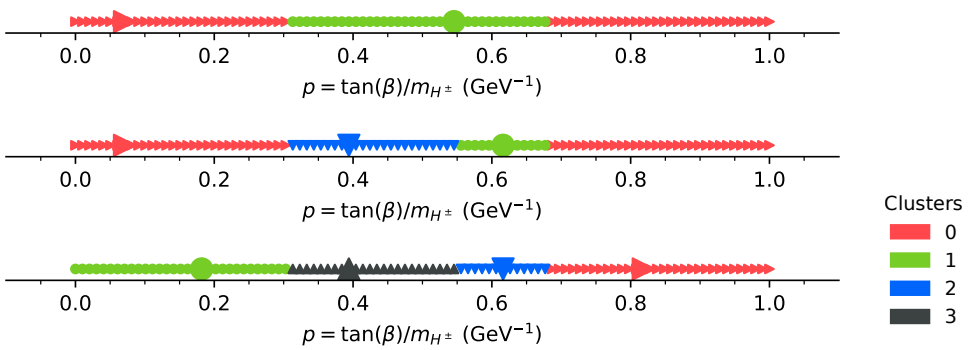


Figure 2.7: Clustering of the one-dimensional parameter space p of the observable $\mathrm{d}\Gamma(\bar{\mathrm{B}} \to \mathrm{D}\tau^-(\to \ell\,\bar{\nu}_\ell\,\nu_\tau)\bar{\nu}_\tau)/\mathrm{d}(E_\ell)$, with E_ℓ denoting the lepton energy. The NP parameter p stems from a two-Higgs-doublet model. The clustering is performed assuming signal yields of 1000, 1800, and 2000 events as well as a relative uncertainty of 1% and leads to a total of two, three, and four clusters, respectively.

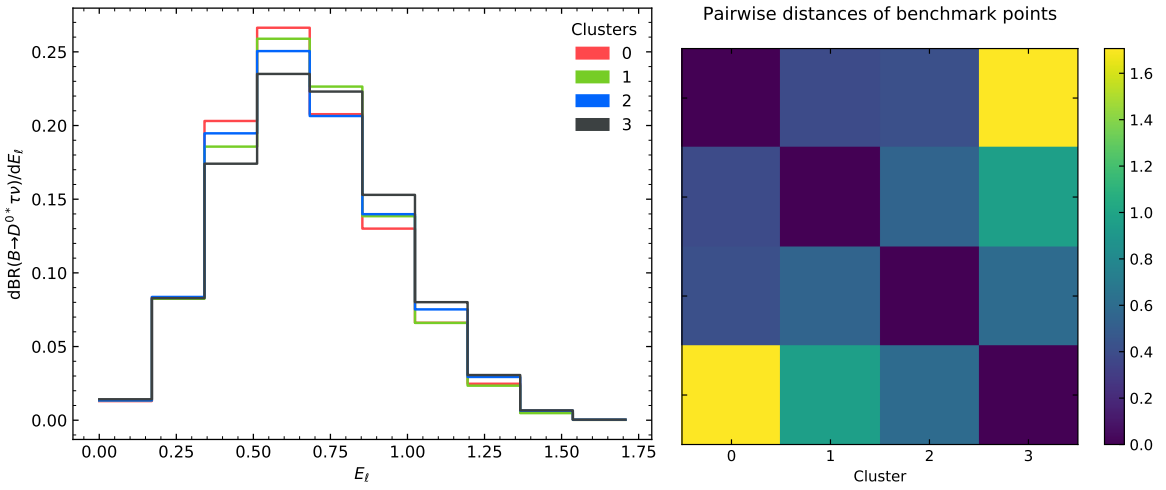


Figure 2.8: Benchmark E_ℓ-distributions of the branching ratio $\mathrm{BR}(\bar{\mathrm{B}} \to \mathrm{D}\tau^-(\to \ell\,\bar{\nu}_\ell\,\nu_\tau)\bar{\nu}_\tau)$ for a yield of 2000 events and a relative uncertainty of 1%. The matrix plot shows the pairwise distances between the kinematic distributions.

2.5 Conclusions

In this chapter, we discussed cluster analyses of kinematic distributions. These analyses divide the parameter space governing the distribution into subsets (*clusters*), in which all points correspond to similar kinematic distributions. Each cluster is then reduced to its most representative point (*benchmark point*, *BP*). Analyses relying on these kinematic distributions can then be carried out for the BPs only, rather than using the entire parameter space. This can drastically reduce the required computing power and make it easier to present numerical results and visualizations.

The results of the cluster analyses depend on the sampling of the parameter space, the clustering algorithm, and the metric measuring differences between kinematic distributions.

This chapter introduced the Python package `ClusterKinG` which implements the above steps and allows to perform clustering analyses without technical overhead. While it particularly aims to make clustering techniques more accessible for the High Energy Physics community, the software can also be applied to more general problems outside of particle physics. `ClusterKinG` is available as open-source software [173, 174] together with usage examples and technical

documentation [177].

We used the `ClusterKinG` package to study several kinematic distributions of the decays $\bar{B} \to D^{0(*)}\tau^-\bar{\nu}_\tau$. The θ_τ and θ_V-distribution of $\bar{B} \to D^{0*}\tau^-\bar{\nu}_\tau$ were studied, showing the clustered parameter space, the BPs as well as the corresponding distributions. A strong dependence of the θ_τ-distribution on the tensor Wilson coefficient C_T has been shown, which agrees with previous findings in the literature [184].

The influence of statistical and systematic uncertainties on the clustering result is shown on the example of the q^2-distribution of $\bar{B} \to D^0\tau^-\bar{\nu}_\tau$.

Finally, we analyzed the E_ℓ-distribution of the 5-body decay $\bar{\mathrm{B}} \to \mathrm{D}\tau^-(\to \ell\,\bar{\nu}_\ell\,\nu_\tau)\bar{\nu}_\tau$. The shape of this variable is an important input for some experimental measurements of $\mathcal{R}(\mathrm{D})$. The resulting model dependency that was observed in [71] on the example of type II two-Higgs-doublet models should also be seen from clustering the input kinematics. While not entirely consistent with the results of [71], our simplified approach correctly hints at the significant change of $\mathcal{R}(\mathrm{D})$ at $\tan\beta/m^2_{\mathrm{H}^\pm} \approx 0.3\,\mathrm{GeV}^{-1}$. However, a complete analysis, including also the shape of the m^2_miss distributions, remains to be done in order to thoroughly compare the results with those of [71].

Chapter 3

The Belle Experiment

The Belle experiment is an electron-positron collider experiment built to study the properties of B mesons. While data taking was stopped in 2010 to focus on the commissioning of the follow-up experiment Belle II, the analysis of the recorded data continues to this day. This chapter gives an overview of the instrumentation and experimental setup of the Belle experiment.

3.1 B factories, Belle and Belle II

The Belle experiment belongs to a group of collider experiments called *B factories* (sometimes *beauty factories*). B factories owe their name to their primary goal: To produce and detect as many B mesons as possible. While other experiments like LHCb or HERA-B were also specifically designed to detect large quantities of B mesons, the term is usually used specifically for electron-positron colliders that operate at the $\Upsilon(4S)$ resonance at a center of mass energy of $\sqrt{s} = 10.56\,\text{GeV}$ (Figure 3.1).

Because the $\Upsilon(4S)$ decays to pairs of charged or neutral B mesons with a branching fraction of more than 96% at 95% C.L. [186], this results in a very clean sample of B meson pairs. A Feynman diagram of the production of B meson pairs at a B factory is shown in Figure 3.2. Also visible is an important source of background at B-factories: note that the cross section contains a contribution that is relatively flat ("continuous") across the mass spectrum. This non-$B\overline{B}$ background is called the *continuum background*. It will be discussed in more detail in Section 8.3.

Concurrent to Belle, a second B factory was in operation: BaBar collected data till 2008. Belle and BaBar shared the same physics goals and had similar detector layouts and analysis strategies. This competition has been very beneficial because the physical results of one experiment can be checked and compared with the results of the other. Both detectors, collaborations, analysis methods, and results are reviewed and compared in [187].

A major scientific breakthrough by both experiments was the observation of $\mathcal{CP}$ violation in the B meson system [188] consistent with the theoretical model by Kobayashi and Maskawa [17], who were subsequently awarded the Nobel prize in 2008. Among the other achievements [189] of the Belle collaboration were the discovery of tetraquark states [190] and precise measurements of CKM matrix elements.

The Belle collaboration eventually included 470 collaborators from 72 institutions in 16 countries [187, p. 20]. In eleven years of data taking, Belle recorded a total integrated luminosity of $711\,\text{fb}^{-1}$ at the $\Upsilon(4S)$ resonance, significantly exceeding the $424.2\,\text{fb}^{-1}$ collected at BaBar, which was forced to stop data taking prematurely after budget cuts [187]. The large data sample size is also owed to the record-setting instantaneous luminosities achieved by the KEKB accelerator (Section 3.2).

In 2010, data taking at Belle was stopped in order to prepare for the upgrade to the next-

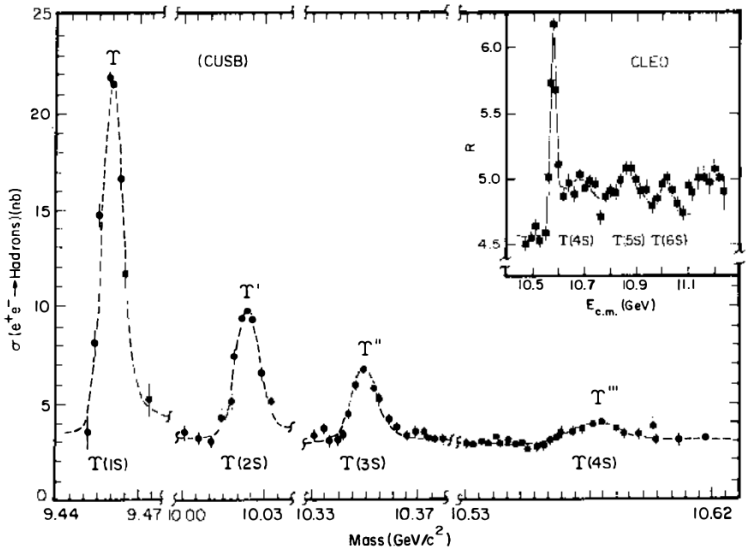


Figure 3.1: An early example of Υ spectroscopy by the CUSB and CLEO experiment. Figure from [191, p. 4].

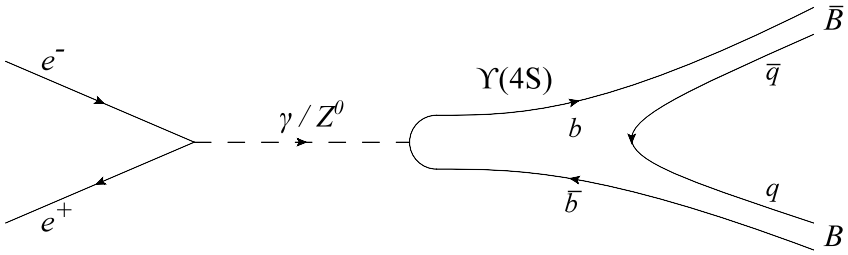


Figure 3.2: Producing pairs of B mesons via the $\Upsilon(4S)$ resonance. Figure from [192, p. 19].

generation experiment, Belle II [193], in conjunction with an improved accelerator complex, Super-KEKB [194, 195]. Together, they are designed to increase the total integrated luminosity by a factor of 50, allowing for an extensive physics program [196]. In addition to the hardware improvements, the analysis software framework has been rebuilt with various improvements to both performance and user-experience [197] (Section 4.1). In particular, it includes an improved algorithm to reconstruct hadronic B mesons in a multitude of channels (Section 4.1.1).

The first collisions were recorded by the Belle II detector in April 2018 [198]. As of 2022, Belle II has collected more than $250\,\text{fb}^{-1}$ of data [199]. While undoubtedly impressive, this dataset is still eclipsed by the Belle data, which is why Belle data continues to be analyzed. In this thesis, we combine the newer analysis strategies of the Belle II software with the larger dataset of the Belle experiment.

3.2 KEKB

The KEKB accelerator complex (Figure 3.3) delivers the electrons and positrons for the Belle experiment. Positrons are generated from collisions of electrons with a Tantalum target. Both electrons and positrons are then accelerated by a linear accelerator [200] to an energy of 8 GeV (electrons) and 3.5 GeV (positrons) and injected into two storage rings of 3 km diameter. Electrons circle clockwise in the *High Energy Ring* (HER), positrons circle counter-clockwise in the *Low Energy Ring* (LER). Radiofrequency cavities and wigglers are used to control beam properties in the storage rings and to counteract the energy loss from synchrotron radiation.

Both rings cross at the interaction point where the collisions occur. Because of the asymmetric

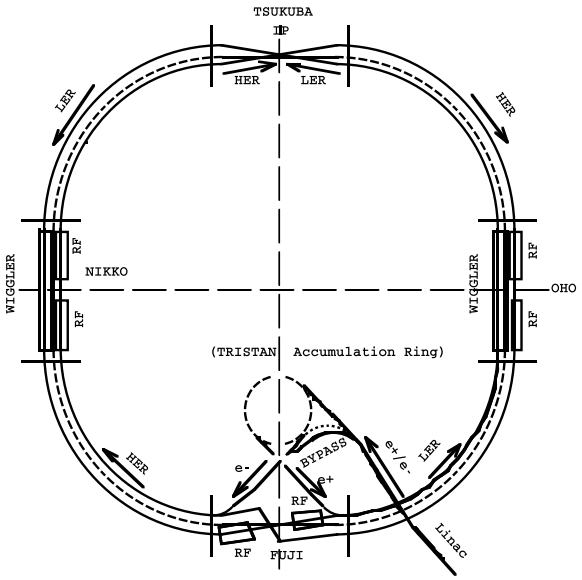


Figure 3.3: The KEKB accelerator complex. The arrows indicate the circling direction of electrons and positrons. The following abbreviations are used: *IP* (interaction point), *RF* (radiofrequency cavity), *linac* (linear accelerator). *Tsukuba*, *Oho*, *Fuji*, *Nikko* are cities and landmarks used to denote the different directions. Figure from [201, p. 3].

beam energy, the center of mass system of the collision moves with

$$\beta\gamma = \frac{E_{e^-} - E_{e^+}}{\sqrt{s}} = 0.425, \tag{3.2.1}$$

as seen from the laboratory frame. This *boosting* of the center of mass system facilitates measurements of time-dependent effects by measuring the displacement of decay vertices.

Luminosity records

On June 17, 2009, KEKB achieved a new world record when it reached a peak luminosity of $2.210\,83 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ [202, p. 1]. After a short period where the LHC held a new record with proton-proton collisions, Super KEKB reclaimed its record with a luminosity $2.22 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ in 2020 [203]. On its way to a 40 times higher target luminosity than that of KEKB [193], Super KEKB has since been ramping up luminosity until reaching $3.1 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ in late 2021 [204].

3.3 The Belle detector

The Belle detector was located at the interaction point of the KEKB storage rings. At this point, both beams are crossing at an angle of 11 mrad and collisions occur. Cross sections of the Belle detector are shown in Figures 3.4 and 3.5. We will use the following coordinates:

- The z-axis is parallel to the e^+ beam. Positive z points in the clockwise direction of the storage rings (that is, opposite to the e^+ beam).
- The y-axis lies in the horizontal plane and points towards the middle of the storage ring.
- The x-axis points upward so that x, y, z form a right-handed orthogonal coordinate system.
- The polar angle θ is the angle in the (x, z)-plane with respect to the z-axis.
- The azimuthal angle ϕ is the angle in the (x, y) plane.

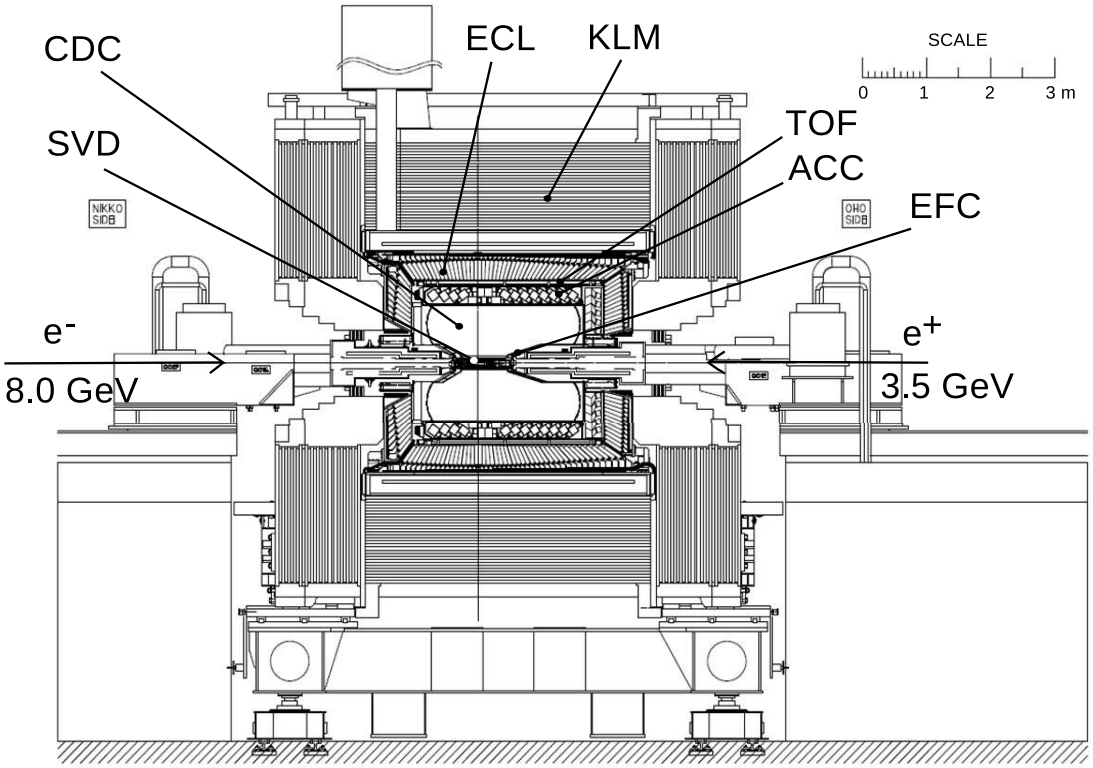


Figure 3.4: The Belle detector: Cross section along the beam axis. The abbreviations for the detector sub-systems are introduced in the following sections. Figure from [187, p. 24].

- $r := x^2 + y^2$ is the distance from the origin in the (x, y) plane.

The individual components of the Belle detector are outlined below.

3.3.1 Beam pipe

Directly surrounding the interaction point is a double-wall Beryllium beam pipe. A cooling liquid flows through the gap between both walls to counter beam-induced heating. The z-vertex resolution of the SVD (the next layer) is nearly proportional to the radius of the innermost SVD layer, providing a strong incentive to minimize the diameter of the beam pipe. At the same time, a decrease in beam pipe diameter increases the amount of beam background generated, which in turn also degrades vertexing performance. To minimize scattering on the beam-pipe wall, the thickness of the wall should thus be as small as possible.

The original design featured a Beryllium beam pipe with an inner diameter of 40 mm and a wall thickness of 0.5 mm. The gap between both walls was 2.5 mm and Helium was used for cooling. Together with the SVD upgrade, the beam pipe was replaced in 2003. The new design featured a reduced inner diameter of 30 mm, wall thicknesses of 0.6 mm (inner) and 0.35 mm (outer) and a 0.5 mm gap through which C_7H_{16} was routed. 40kg of tantalum and tungsten were used for additional shielding to deal with the increased beam backgrounds resulting from the reduced inner diameter.

The original beam pipe setup is described in [206], the update in [207].

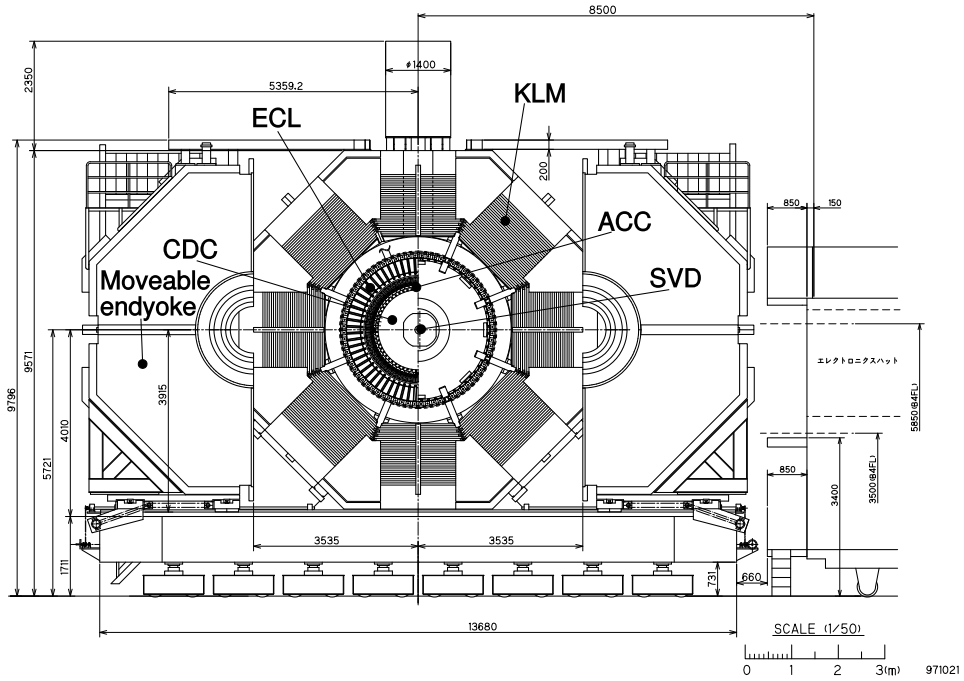


Figure 3.5: Front view of the Belle detector. The abbreviations for the detector sub-systems are introduced in the following sections. Figure from [187, p. 24] (which is an adapted version from [205, p. 5]).

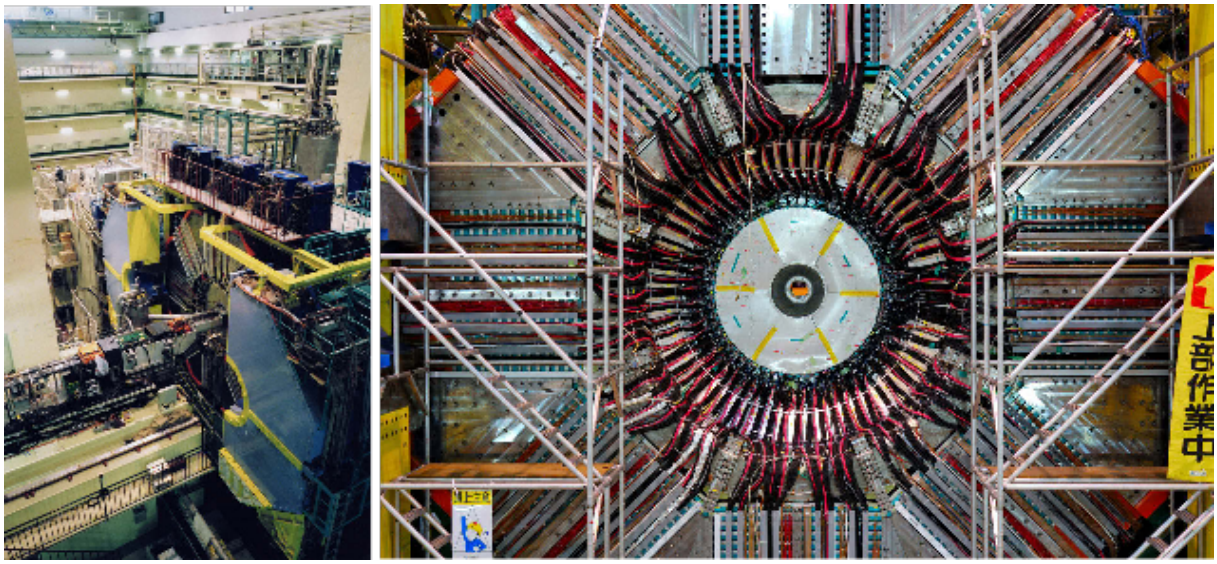


Figure 3.6: View of the beam pipe leading to the Belle detector and front view of the detector. Figure from [205, p. 25].

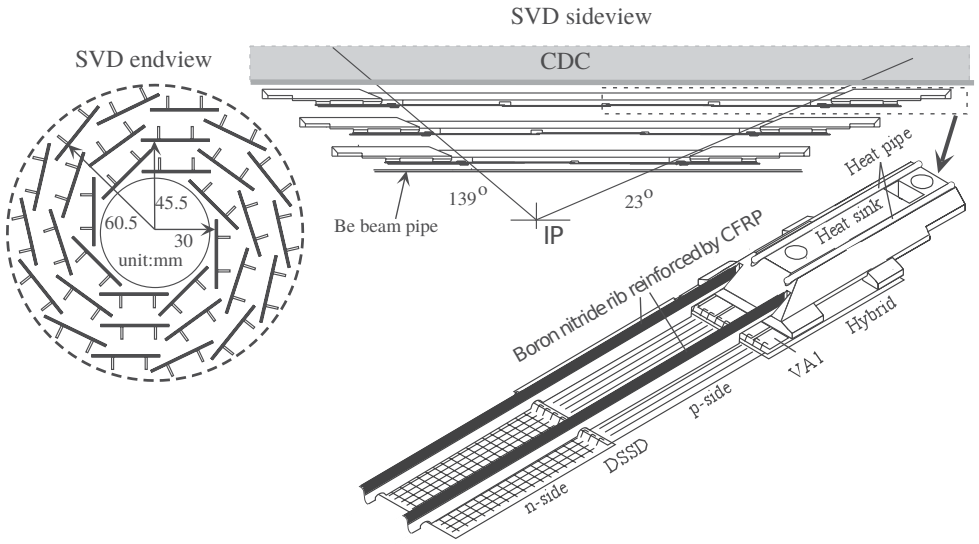


Figure 3.7: Silicon Vertex Detector before 2003 (SVD1). Figure from [209, p. 2].

3.3.2 Silicon Vertex Detector

One of the constitutive goals of the Belle experiment was the observation of $\mathcal{CP}$ violation in the B meson system, which manifests itself in the time-dependent decay rates of $B\overline{B}$ pairs. Due to the boost of the CMS system, the lifetimes of the B mesons can be directly inferred from the distance of the interaction point and the B decay vertex if the vertexing is precise enough.

The SVD consists of a large number of double-sided strip detectors (DSSDs). Charged particles traversing DSSDs create electron-hole pairs in *pn*-junctions. An applied bias voltage draws the electron-hole pairs to sense strips, resulting in an electrical signal. Because the sense strips of the two sides are perpendicular, z-ϕ information can be inferred.

Several DSSDs are arranged in z-direction to make up one *ladder*. Multiple ladders are arranged in a near-circular fashion in the $r-\phi$ plane and form a *layer*. An overlap in the $r-\phi$ plane helps with alignment studies.

The initial SVD configuration (SVD1) depicted in Figure 3.7 consisted of three layers of 8, 10 and 14 ladders with a total of 102 DSSDs covering $23° \leq \theta \leq 139°$.

In 2003 the SVD1 was replaced with an upgraded configuration (SVD2) depicted in Figure 3.7. This new setup had improved radiation hardness, lower dead time, and better overall performance. It consisted of four layers of 6, 12, 18, and 18 ladders with a total of 246 DSSDs covering $17° \leq \theta \leq 150°$. Because of the thinner beam pipe, the innermost layer could also be moved closer to the interaction point. The new setup provided a significantly improved impact parameter resolution and allowed the reconstruction of charged-particle tracks using only SVD information.

The SVD1 is described in detail in [206], the SVD2 in [208, 209].

3.3.3 Calorimeters

Calorimeters measure the energy of particles. By segmenting the calorimeter, additional spatial resolution can be provided. In an electromagnetic calorimeter, particles like electrons or photons deposit their energy by initiating electromagnetic showers (cascading processes of bremsstrahlung or pair production depending on the particle's energy) in the material. In homogeneous scintillating crystal calorimeters as used by Belle, the particle is absorbed in a single crystal, and the deposited energy is measured by counting the number of scintillation photons using photomultipliers. [210]

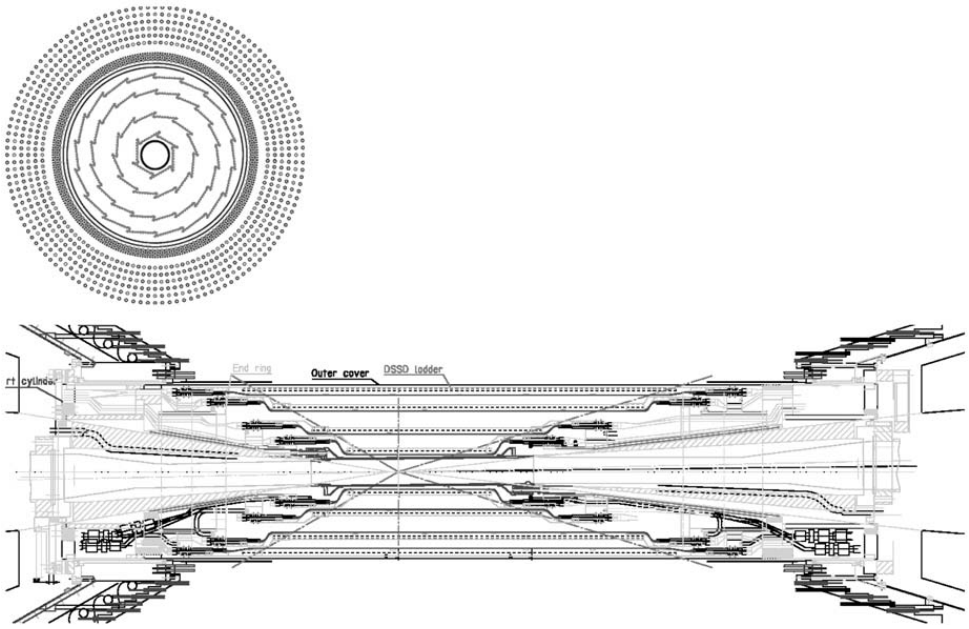


Figure 3.8: Silicon Vertex Detector after 2003 (SVD2). Figure from [209, p. 3].

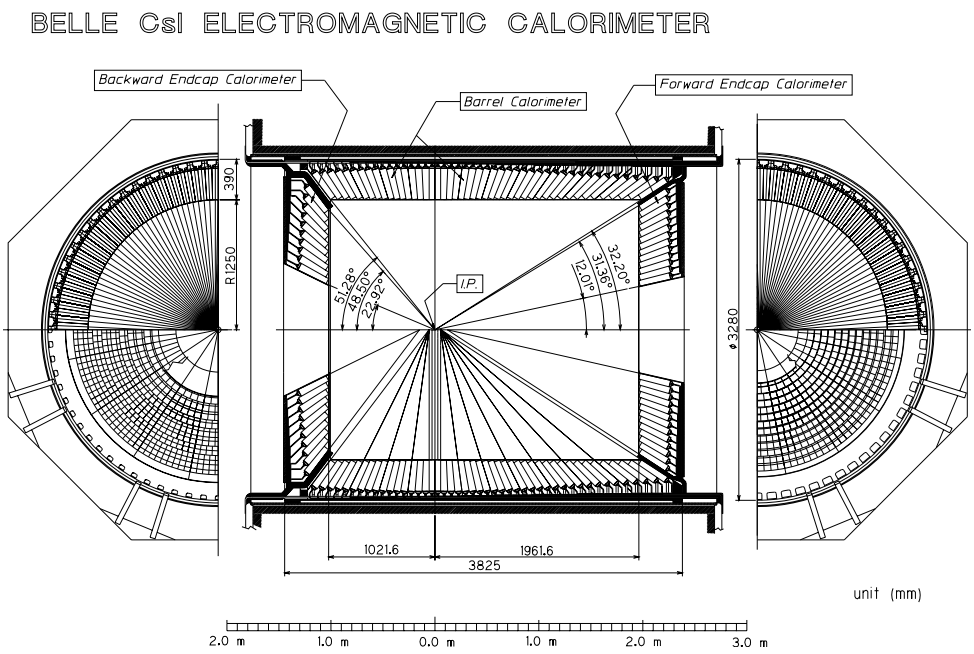


Figure 3.9: Electromagnetic Calorimeter (ECL). Figure from [206].

Electromagnetic Calorimeter (ECL)

The main objective of the ECL is the detection of photons with good energy and location resolution. Adequate performance for photons of a wide range of energy spanning from few 10 MeV (photons from cascade decays) up to 4 GeV (two-body decays like $B \to K^*\gamma$) is required. By comparing the energy deposit to the momentum of charged particles as measured by the other detectors, the ECL also helps with electron identification.

The ECL consists of 8763 CsI(Tl) crystals arranged in a barrel section and two endcaps, weighing 43 tons in total. The setup is shown in Figure 3.9. The crystals almost point to the IP but for a slight tilt to avoid missing photons in the gap between neighboring crystals. Except for a small gap between the barrel and endcap sections, the ECL provides a coverage of $12.4° \leq \theta \leq 155.1°$.

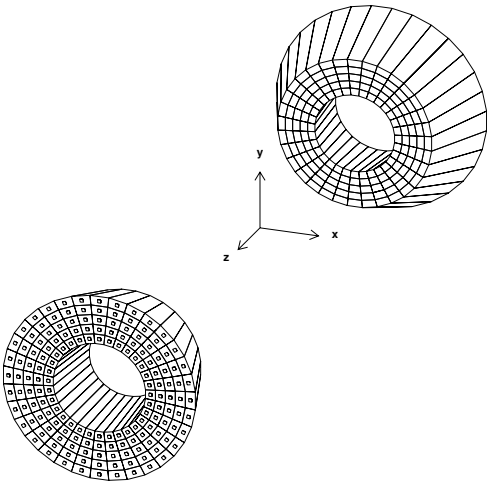
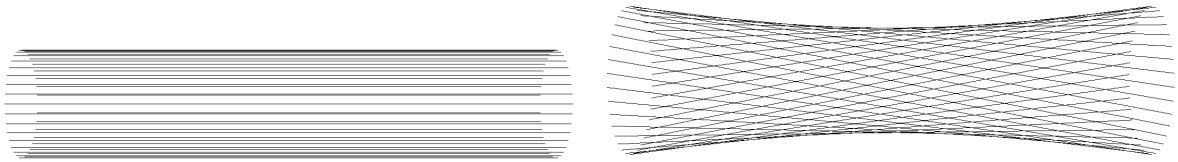


Figure 3.10: Extreme Forward Calorimeter (EFC). Figure from [206].



(a) Axial layer (b) Stereo layer

Figure 3.11: Illustration of the wire configurations of the axial and stereo superlayers of the CDC. The skewedness of the stereo layers has been exaggerated. Figure from [212, p. 16].

Extreme Forward Calorimeter (EFC)

The EFC extends the polar range provided by the ECL by covering $6.4° \leq \theta \leq 11.5°$ and $163.3° \leq \theta \leq 171.2°$. It is shown in Figure 3.10. The necessary proximity to the beam pipe and the short distance to the interaction point exposes it to very high levels of beam-induced radiation. To achieve the necessary radiation hardness BGO (Bismuth Germanate, $Bi_4Ge_3O_{12}$) crystals are used as scintillating crystals.

The EFC monitors beam parameters, measures the luminosity, and supports the ECL in measurements like $B \to \tau\nu$ or two-photon physics. It also shields the CDC from beam background.

3.3.4 Central Drift Chamber

In a drift chamber [211], charged particles ionize a gaseous medium on their path. The freed electrons and ions drift to an anode or cathode wire, resulting in an electric signal. The position of the particles can then be calculated from the time that the electrons drift to the anode and the drift velocity of the electrons in the medium.

The Belle CDC consists of 50 radial (cylindrical) layers of sense wires and three cathode strip layers. Due to the boost of the B mesons, the CDC is asymmetric in the z direction and covers $17° \leq \theta \leq 150°$.

The cathode strip layers are close to the interaction point. As they are segmented in z-direction, they can provide fast and precise z-coordinate measurements that are more difficult to achieve for the remaining layers [213].

The remaining layers are divided into axial and (small angle) stereo superlayers (each containing up to six of the 50 radial layers). In the axial superlayers, the wires are parallel to the z-axis, and no information about the z-position of a traversing particle can be inferred. However, the

wires of the stereo superlayers are slightly skewed to the beamline and thus provide input to reconstruct the z-coordinate. The two wire configurations are shown in Figure 3.11. In radial layers of both types, each sense wire is surrounded by eight field wires, forming one drift cell. Each radial layer has the same number of drift cells (in ϕ direction) for a total number of 8400 drift cells. [206, 214]

For the best possible momentum resolution of the low momentum decay products of the B mesons, multiple Coulomb scattering in the gas must be minimized. For this reason, a low-Z gas is preferred[1] and a 50% helium 50% ethane mixture was chosen, where the ethane component ensures a good $\mathrm{d}E/\,\mathrm{d}x$ measurement (see below). The small photo-electric cross section of low-Z gas also helps to reduce background from synchrotron radiation [215].

The CDC performs multiple tasks in the Belle detector:

- it records precise three-dimensional trajectories (*tracks*) of charged particles (*tracking*). The data from the CDC can be further combined with hits from the SVD to improve the accuracy of the tracking.
- it provides high resolution momentum measurements via the curvature of the tracks in the 1.5 T magnetic field (see Section 3.3.7)
- it allows measuring the energy loss $\mathrm{d}E/\,\mathrm{d}x$ of charged particles by measuring the amplitude of the electric pulses. Because the energy loss depends on the speed of the particle (Bethe-Bloch formula [216]), this information can be combined with the momentum measurement to infer the particle's mass. In particular, the $\mathrm{d}E/\,\mathrm{d}x$ measurement provides a good separation between kaons and pions up to a momentum of more than 1 GeV. [206, 217]
- it provides input to the trigger system. Especially the fast z-direction measurements from the strip cathodes help to exclude background from beam-gas scattered electrons and positrons interacting with the beam pipe (Z-trigger) [218].

3.3.5 Aerogel Cherenkov Counters

If the speed v of charged particles traversing a medium exceeds the phase velocity of light c_{medium} in this medium, electromagnetic radiation is emitted. Using the definition of the refractive index n of the medium, the requirement for this *Cherenkov radiation* reads:

$$\beta = \frac{v}{c} > \frac{c_{\text{medium}}}{c} = \frac{1}{n}. \tag{3.3.1}$$

The *Aerogel Cherenkov Counters* (ACC) of the Belle detector used this condition to discriminate between charged particles, in particular kaons and pions: The refractive indices of the used media were chosen between 1.01 and 1.03 (depending on the polar angle[2], such that for a momentum of $1.2 < |\vec{p}|/\mathrm{GeV} < 3.5$, Cherenkov radiation is only produced by pions, not kaons. The ACC thus complements the particle identification provided by the CDC.

The ACC consists of 960 modules in the barrel region and 228 modules in the forward end-cap region (Figure 3.12). Each module points to the interaction point and together cover the polar angles $17° \leq \theta \leq 127°$. Note that the ACC only returns a boolean result (*threshold counter*) and does not make use of the additional information that can be obtained from measuring the opening angle of the Cherenkov radiation cone [219].

The ACC is described in more detail in [206, 220]. A schematic of the individual modules is shown in Figure 3.13.

[1] Here, Z refers to the proton number.

[2] The asymmetric beam energies result in a correlation between the polar angle and the particle momenta. Therefore, the optimal refractive indices are varied accordingly [219, p. 596]. The different refractive indices are highlighted in Figure 3.12.

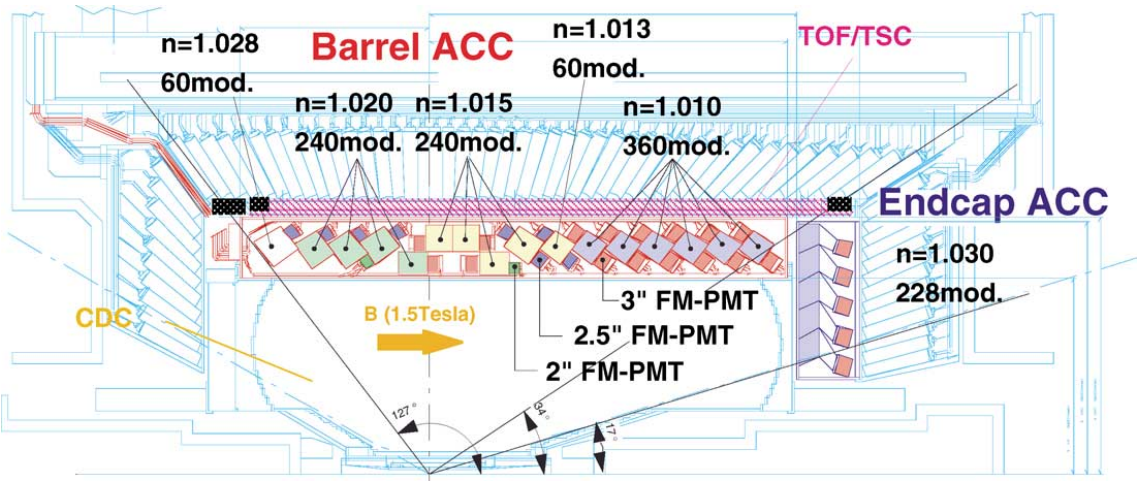
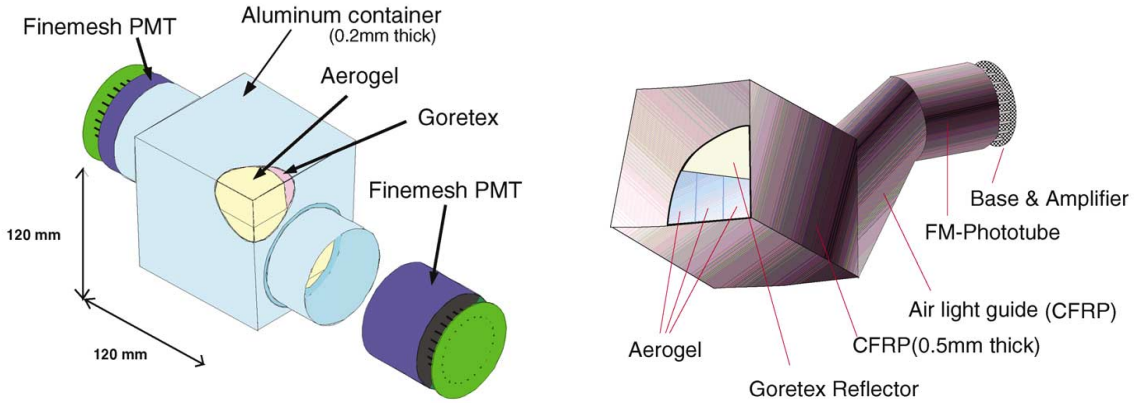


Figure 3.12: Position of the ACC modules in the Belle detector. Figure from [206, p. 156].



(a) Barrel module

(b) Endcap module

Figure 3.13: ACC modules. The Cherenkov radiation is produced in the aerogel, reflected by optical reflectors (*Goretex*) or guided by light guides, and detected by photomultiplier tubes (marked *PMT*/*FM-Phototube*). Figures from [206].

3.3.6 Time Of Flight system

The Time Of Flight (TOF) detector system measures the time of a particle to travel from the initial collision (provided by a reference clock synchronized with the beam collisions) to the TOF system. The TOF system consists of 64 modules at $r = 1.2\,\mathrm{m}$ covering a polar angle range of $34° \leq \theta \leq 120°$. Each TOF module consists of two 4 cm thick TOF counters with directly attached photomultiplier tubes and one 0.5 cm thick trigger scintillation counter (TSC) with the photomultiplier tube connected via a small light guide.

A minimum momentum of 0.28 GeV is required to reach the TOF module. Together with momentum measurements, the 100 ps resolution allows distinguishing kaons and pions of momenta lower than 1.25 GeV [221, p. 5], complementing the ACC and CDC.

The TOF system also provides timing signals for the trigger system. In particular, this signal controls the readout of the ECL and the CDC. Because the trigger rate is capped at 70 kHz, a coincidence requirement on TOF and TSC signals is used to lower the trigger rate and reduce background. The trigger system and front end electronics are described in detail in [222], the general setup in [206, 221].

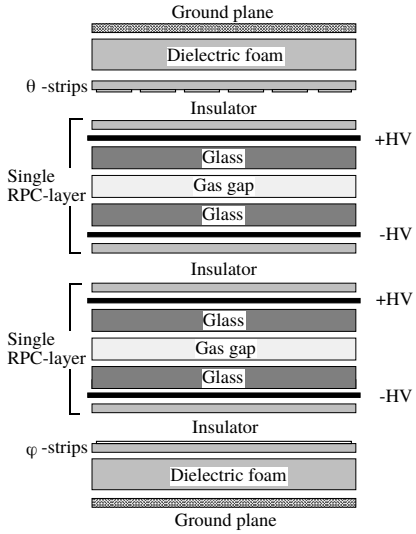


Figure 3.14: Structure of an RPC superlayer. Figure from [224].

3.3.7 Superconducting Solenoid

The superconducting solenoid (SCS) provides a homogeneous 1.5 T magnetic field. The coil has an effective radius of 1.8 m and is located in a barrel-shaped cryostat of 1.70 m (2.00 m) inner (outer) radius and 4.4 m length. The coil is made from a Niobium-Titanium-Copper alloy (NbTi/Cu) with aluminum as the supporting structure. Operated at a nominal current of 4400 A, 35 MJ of energy are stored in the magnetic field. The SCS is surrounded by the iron yoke that serves as flux return for the magnetic field and is described in the next section. The SCS is described in detail in [206].

3.3.8 Iron yoke, K_L^0 and Muon Detector

The iron support structure is used as a magnetic flux return and contains the last subdetector: The K_L^0 and muon detector (KLM), detecting neutral long-lived kaons and muons. Located outside of the SCS, the structure consists of alternating layers of 4.7 cm thick iron plates and resistive plate counter (RPC) superlayers detecting traversing charged particles.

Resistive plate counters consist of two highly resistive plate electrodes with a gas-filled gap. Traversing charged particles ionize the gas and create a local discharge between the plates that induces an electrical signal on pickup strips [223, 224]. The high resistance of the plates is important to stop the discharge.

For the Belle detector, the plates are made of coated glass with a resistance exceeding $10^{10}\,\Omega\,\mathrm{cm}$ and are operated at a gap voltage of 8 kV. The cross section of an RPC superlayer is shown in Figure 3.14. Each superlayer contains two RPCs with orthogonal pickup strips to provide a 3-dimensional location of the hit.[3]

K_L^0 mesons interact with the material of the ECL or the iron plates (totaling 4.7 interaction lengths) and cause showers of ionizing particles that are detected. In contrast to muons, they cannot be associated with extrapolated tracks from the CDC.

Muons are comparatively weakly interacting particles. Therefore, they are less deflected by detector material, cause no showers in the KLM and can penetrate more layers of the KLM than hadrons. By matching extrapolated tracks from the CDC to KLM hits, muons can be

[3] In fact, the two RPC layers also provide redundancy (that is, high efficiency), because a discharge on either of the RPCs will induce a signal on both readout strip planes [205, p. 6].

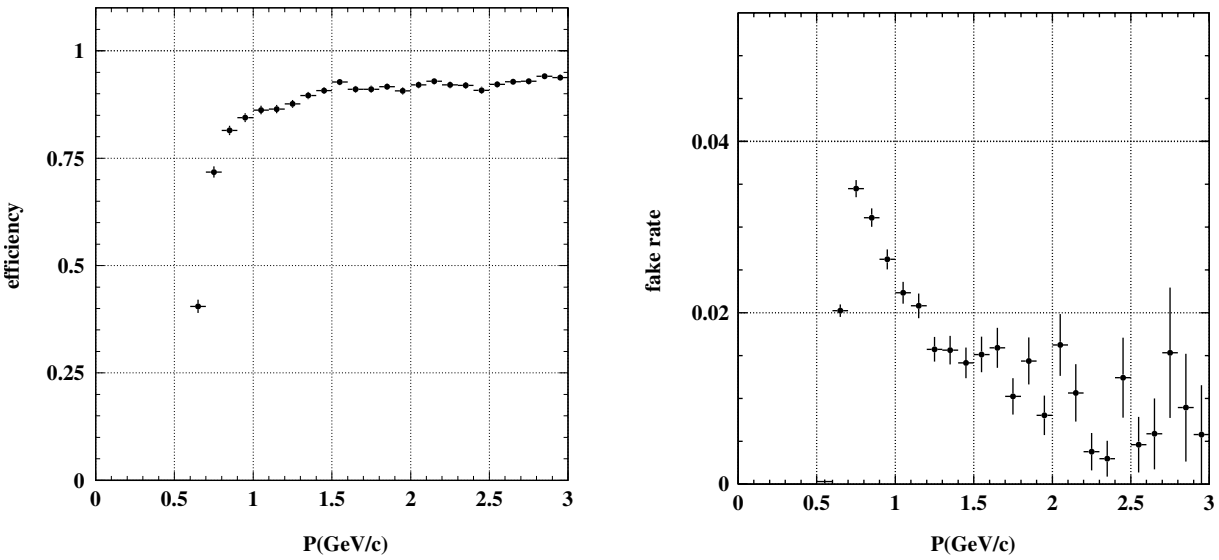


Figure 3.15: Efficiency and fake rate of muon identification versus the momentum for $\mathcal{L}_\mu > 0.66$, where $\mathcal{L}_\mu$ is the muon identification likelihood. Figures from [206, pp. 201–202].

distinguished from charged hadrons. The muon identification method is described in more detail in [205].

A minimum transverse momentum of 600 MeV is required to reach the KLM. As can be seen in Figure 3.15, the muon identification efficiency and purity increase with the particle momentum. For an optimal hadron rejection (less than 2 % of hadrons in the muon sample), muons need to traverse half of the KLM, requiring a momentum of more than 1 GeV [205, p. 3].

A detailed description of the KLM is found in [205, 206, 224].

3.4 Particle Identification

Particles with a long enough lifetime to traverse most of the detector are called *final state particles* (FSPs).

Charged FSPs (e^-, μ^-, π^-, K^-) are reconstructed from tracks with information from the CDC and SVD. By combining the tracks with additional information from the ECL and other sub-systems, the different particles can be distinguished. The likelihood of different particle hypotheses is expressed as likelihood ratios $\mathcal{L}$. Electron identification is described in detail [225], muon identification in [205].

Neutral FSPs (γ, K^0_L) are identified from clusters in the ECL or KLM. Photon candidates can also be formed from decays to two charged particles. In these detector signatures, the two tracks form a characteristic V shape, giving them the name *V0*s. They also allow the reconstruction of K^0_S mesons. Similarly, candidates for the short-lived π^0 mesons are formed from pairs of photons candidates.[4]

[4] Technically, of course, π^0 and K^0_S mesons, as well as photons reconstructed from V0s are not FSPs because they are reconstructed from other decay products.

Chapter 4

Software and Software Training at Belle II and in the HEP community

The amount of data collected at modern HEP experiments grows to ever new orders of magnitude. In 2020, a long way from the target luminosity, Belle II was already collecting up to 50 TB of raw data per day [226]. The target integrated luminosity of $50\,\mathrm{ab}^{-1}$ is expected to correspond to around 60 PB [226] of raw data – and this number is still dwarfed by the total data collected at the LHC experiments. At this point, the physics output of a collaboration hinges on its ability to keep up with the latest software technologies in all stages of data processing. Moreover, it is not only the amount of data that puts software in the spotlight: New algorithms can also significantly improve the precision of physical results. For example, using the improved tagging algorithm of the Belle II software to measure $|V_{\mathrm{cb}}|$ with the Belle dataset (as prepared in this thesis) is expected to increase the reconstruction efficiency up to twofold.

At the same time, most new members joining the HEP experiments lack formal training in software engineering. Because they have typically already spent years developing physical intuition and understanding, these technical skills are often the limiting factor in their early scientific progress. However, even for new members with a strong background in software engineering, the large stack of domain- or experiment-specific software can be challenging. For this reason, software documentation and training material have a significant impact on the output of the experiment: Adequate training activities quickly bring new members up to speed and instill best practices that can have significant long-term benefits. Furthermore, centralizing these efforts increases their efficiency and frees up resources of more senior members.

This chapter introduces the Belle II software framework used throughout the thesis. In particular, the Full Event Interpretation, whose calibration makes up most of this thesis, is described and compared with its predecessor. I then describe a complete rebuild of the Belle II training material that I have coordinated as convener of the Belle II software training and documentation group, which I have led since mid-2020. Finally, I highlight coordination work as one of the conveners of the Software Training and Careers working group [227] of the High Energy Software Foundation (HSF) [228].

4.1 The Belle II software framework

The Belle II Analysis Software Framework (*basf2*) consists of more than 40 different packages with a wide range of responsibilities, including the high-level trigger, data acquisition, tracking, event reconstruction, and post-reconstruction analysis tools. The basic user-facing building block of most packages is that of a *module* which represents a single unit of processing. Central to each module is the `event` method, which processes a single collision event and adds or modifies the corresponding event data in the `DataStore`. Additional data that are not related to specific events, such as detector configuration and calibration constants are accessed from the Conditions

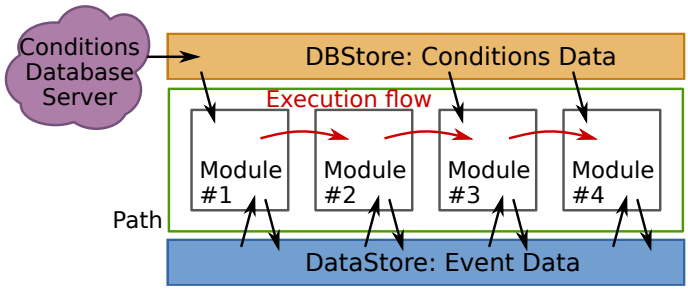


Figure 4.1: Execution flow in the Belle II Analysis Software Framework. Figure adapted from [231, 232].

Database Server and stored in the `DBStore` [229, 230]. As shown in Figure 4.1, multiple modules are arranged in a linear fashion to form a processing chain (called *path*). Files that configure such processing chains are called *steering files*. They are written in Python or C++.

The analysis framework is described in detail in [197, 233] and is available as open source [232]. Extensive documentation is available publicly [231] (more about this in Section 4.2).

Three packages are described in more detail in the next sections:

- The Full Event Interpretation (FEI, Section 4.1.1) is the hadronic tagging algorithm whose calibration is described in this thesis. Its increased efficiency is one of the main reasons for repeating the $|V_{cb}|$ analysis of Belle data with the Belle II software.
- The `b2bii` package (Section 4.1.2) converts Belle data into a format that the Belle II framework can process.
- The validation framework (Section 4.1.3) is responsible for performance and integration testing of the Belle II software. It is mentioned here because of my role in its continued development and maintenance.

4.1.1 Full Event Interpretation

As already mentioned in Section 1.4, *B-tagging* means reconstructing one of the two B mesons from the $\Upsilon(4S)$ decay (tag B, B_{tag}) in a well-understood decay channel that is not in the focus of the measurement. This allows constraining the kinematics of the other B meson (signal B) that is reconstructed in the channel of interest. We can distinguish between two tagging strategies [196, p. 109]:

Hadronic tagging The B_{tag} is reconstructed in a hadronic mode (i.e., without neutrinos). This means that we can infer the B_{tag} momentum and thereby the momentum of the B_{sig}, even if we miss neutrinos on the signal side. This advantage is essential for the $B \to D^* \ell \nu_\ell$ analysis because it allows using the m^2_{miss} observable for background discrimination. Generally, the complete kinematic information obtained with hadronic tags allows for very pure samples. The downside of this method is the low overall efficiency.

Semileptonic tagging The B_{tag} is reconstructed in semileptonic modes, profiting from the large branching fraction available. However, because neutrinos are involved, the available kinematic information is limited, leading to samples of lower purity than with hadronic tagging.

The following discussion is limited to the hadronic tagging because of its use in the $|V_{cb}|$ analysis. Since the weak spot of tagging is its low efficiency, tagging algorithms try to consider as many decay channels as possible.

4.1.1.1 Comparison with previous algorithms

The first version of hadronic tagging algorithms in the Belle collaboration used a series of rectangular cuts in the reconstruction and selection of hadronic tag candidates. This method, called (cut based) *Full Reconstruction* (FR) is for example used in [234, 235]. Later, the NeuroBayes neural network [236, 237] was developed and used to select the candidates. The new hierarchical multivariate approach led to up to twofold increased efficiencies [238].

The *Full Event Interpretation* (FEI) [239, 240] is the successor of the FR for the Belle II experiment. It uses a similar hierarchical approach but considers more decay channels and has an improved candidate selection. Instead of NeuroBayes, boosted decision trees in a speed-optimized implementation (*FastBDT*) [241, 242] are used to calculate the likelihood of correctly reconstructing candidates.

The performance of a tagging algorithm is commonly described in terms of

tag-side efficiency the fraction of $\Upsilon(4S)$ for which a correct tag has been reconstructed,

tag-side purity the fraction of events with correctly reconstructed tag among the events with any reconstructed tag.

For every $\mathrm{B}_{\mathrm{tag}}$ candidate that has been reconstructed, the FR and the FEI return a quantity (*classifier output* $\in [0, 1]$, also called *signal probability*) related to the likelihood of the candidate to be correctly reconstructed. Placing cuts on this classifier increases the tag-side purity but lowers the tag-side efficiency. Comparing the performance of a tagging algorithm thus means comparing both quantities as a function of each other.

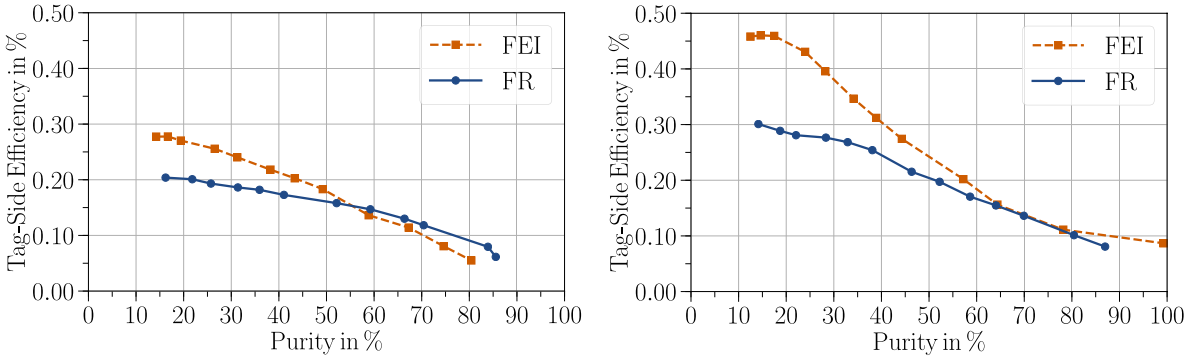
This is shown in Figure 4.2. For this study (described in more detail in [239]), the performance of the algorithms was tested on the full Belle dataset. The same selection criteria were used on the sample with FR and with FEI reconstruction, namely a lower bound on $m_{\mathrm{bc}}^{\mathrm{tag}}$ and a cut on ΔE. Here, the *beam constrained mass* $m_{\mathrm{bc}}^{\mathrm{tag}}$ and the *energy difference* ΔE are defined as

$$m_{\mathrm{bc}}^{\mathrm{tag}} = \sqrt{E_{\mathrm{beam}}^{*2} - \vec{p}_{\mathrm{B}_{\mathrm{tag}}}^{*2}} \quad \text{and} \quad \Delta E_{\mathrm{B}_{\mathrm{tag}}} = E_{\mathrm{B}_{\mathrm{tag}}}^{*} - E_{\mathrm{beam}}^{*}, \tag{4.1.1}$$

where E_{beam}^{*} is the center-of-mass (CMS) energy of the beam (half of the total energy $\sqrt{s}$ of the $\mathrm{e}^{-}\mathrm{e}^{+}$ system), $\vec{p}_{\mathrm{B}_{\mathrm{tag}}}^{*}$ is the $\mathrm{B}_{\mathrm{tag}}$ CMS momentum and $E_{\mathrm{B}_{\mathrm{tag}}}^{*}$ is the CMS momentum of the $\mathrm{B}_{\mathrm{tag}}$. Both variables and their complementary use in background suppression are described in [187, pp. 85–86].

The $\mathrm{B}_{\mathrm{tag}}$ candidates with the highest classifier output were chosen in the best candidate selection. In order to extract the yield of correctly tagged events, the $m_{\mathrm{bc}}^{\mathrm{tag}}$ spectrum was fitted. Note that the efficiencies for purities higher than 70 % were reported to be not reliable because of a strong dependence on the signal and background model used in the fit. It should also be pointed out that because this study was performed on data rather than on MC, the signal definition used in the definition of the purity is very coarse ("signal is what peaks in $m_{\mathrm{bc}}^{\mathrm{tag}}$"). In particular, the definition has likely not considered the correctness of the flavor of the reconstructed B meson. This is an important detail for the discussion in Chapter 11.

We can see that the FEI outperforms the FR consistently (if we ignore the possibly questionable efficiency values of the high-purity region in B^{0}), in particular at low purities, where the efficiency gain can surpass 50 %. An explanation for this is that the decay channels with particularly clean signatures are shared between the FEI and the FR and are reconstructed with similar efficiencies [239, p. 7], such that the additional decay channels of the FEI mainly contribute to the low purity region. If no classifier cut is applied, this also means that the FEI sample is expected to be less pure than that of the FR. This will play a role in Chapter 11.



(a) B^0 (b) B^+

Figure 4.2: Receiver operating characteristic of the FEI vs. the FR. Figure from [239].

Figure 4.2 should only be regarded as a very general comparison in performances because the impact on a particular analysis cannot be easily estimated. This is because analyses combine the tag and signal candidates and perform additional selection steps based on the combined information. In this combined selection, different B_{tag} candidates can be selected than when only considering the tag side separately. This means that for actual analyses, the tag-side efficiency and purity do not only depend on the B_{tag} candidate with the highest classifier output (as in Figure 4.2), but on having a variety of good quality B_{tag} candidates. This is another area of significant improvement of the FEI, which generally provides more B_{tag} candidates of adequate quality than the FR [239, p. 7].

4.1.1.2 The algorithm

The hierarchical structure of the FEI is illustrated in Figure 4.3. Starting with detector information, we can divide the algorithm into six stages:

1. In the first step, information from the detector is combined to form candidates for *final state particles* (particles that do not decay within the detector and are thus considered to be "stable" for our purposes). Charged leptons, kaons, and pions are reconstructed from tracks, K_L^0 mesons from clusters in the KLM. Photon reconstruction combines V0 objects with clusters in the ECL.

2. J/ψ mesons are reconstructed from e^-e^+ or $\mu^-\mu^+$, π^0 mesons from pairs of photons.

3. K_S^0 are reconstructed from pairs of muons, charged pions, neutral pions, or a V0 object

4. D and D_s mesons are reconstructed in a variety of channels. D_s mesons contribute relatively little to the overall efficiency of the hadronic FEI. The D channels are shown in Table D.2.

5. D^* (D_s^*) mesons are reconstructed from D (D_s) mesons and one additional (neutral or charged) pion or photon (also see Table D.2). The contribution of D_s^* mesons to the efficiency of the hadronic FEI is tiny.

6. Finally, the reconstructed particle candidates from the previous steps are combined to B_{tag} mesons. There are a total of 54 hadronic B_{tag} reconstruction modes, listed in Table 9.1 and shown with frequencies in Figure 6.2.

The combinations for all steps along with relevant selection criteria are also shown in [240, app. C.1].

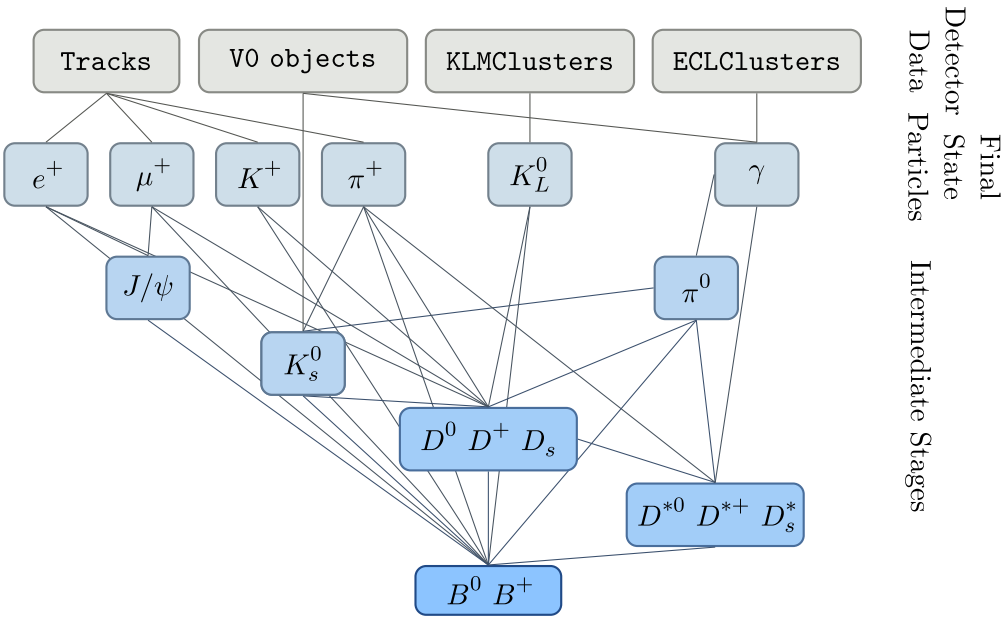


Figure 4.3: B_{tag} reconstruction with the FEI. Detector information is shown as gray nodes, reconstructed particle candidates as blue nodes. The lines connect particles to their decay products or the relevant detector information. Charge conjugated particles are implied (e.g., e^+ denotes e^- as well). Figure from [239].

4.1.2 Belle to Belle II Conversion

In order to apply the FEI to data recorded by the Belle experiment, it must be converted to the new data format of the Belle II software. This is achieved by the basf2 module `b2bii` [240, 243]. Rather than converting raw detector information (the immediately recorded output of the various detector systems), `b2bii` converts the so called *mDST* data. In this dataset, calorimeter clusters and tracks have already been reconstructed from the raw data. This has the advantage that the data objects are more detector-independent.

Specifically, the following information is retrieved from the mDST data files:

- ECL and KLM clusters,
- Tracks and V0 objects,
- PID information,
- Beam parameters,
- Generator level information (MC only, see Section 7.1).

Based on this input, the usual basf2 objects are created, which allows to apply the various basf2 modules for analysis.

4.1.3 The validation framework

The Belle II software is still evolving at a rapid rate. Algorithms for tracking, reconstruction, and other tasks are continued to be improved and tuned to the experimental conditions to deliver the best physical results. While *unit tests* ensure the basic functionality of individual software components, they cannot ascertain their performance. Usually, a sufficiently large sample of events needs to be processed to determine a new algorithm's efficiency, runtime, or accuracy. Such tests require significant computation time and manual checks and decisions by experts

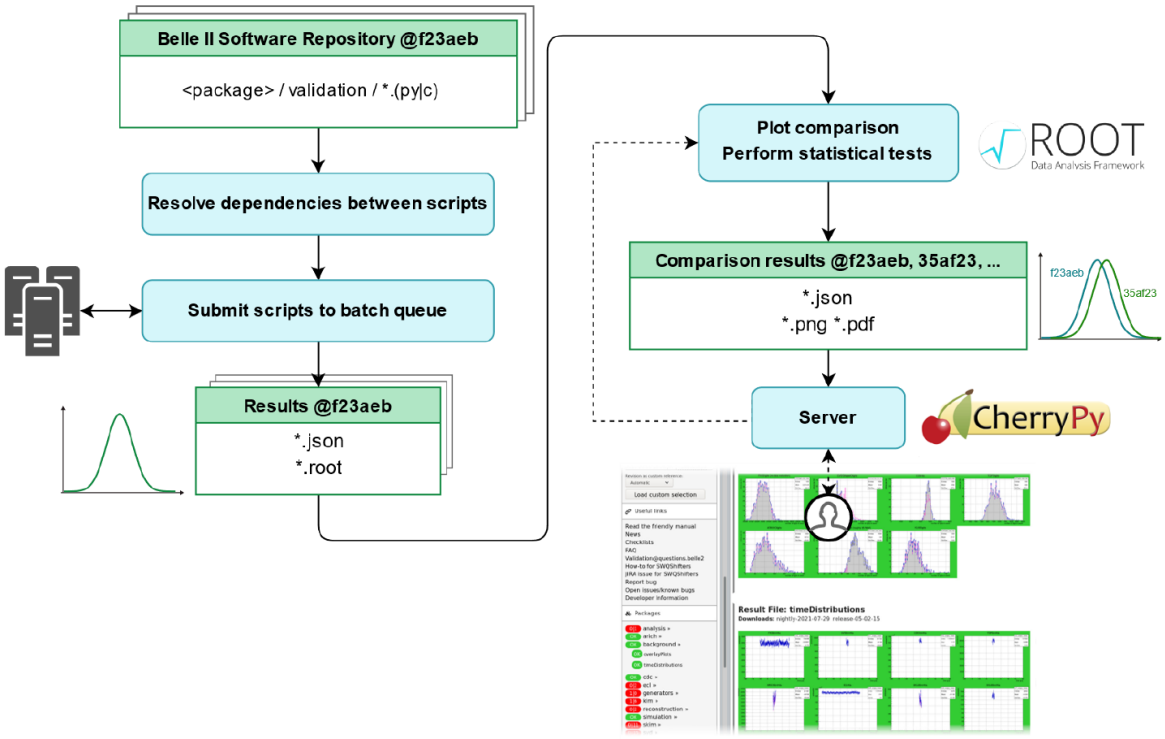


Figure 4.4: The validation framework. The stacked boxes on the left side show that the validation framework produces results for different software versions (two examples are denoted `f23aeb` and `35af23` by their git commit hashes).

(for example, situations featuring bias-variance-runtime tradeoffs cannot readily be evaluated automatically).

Checks and evaluations of this kind are facilitated by the validation framework. Every basf2 package contains a subfolder with a series of basf2 steering files. The validation framework collects these steering files, resolves dependencies between them, and submits them to a batch system in the correct order. Besides intermediate results, these steering files ultimately produce a series of histograms and other physical quantities of interest. The validation framework executes the steering files with different basf2 versions and records the outputs for every version. They can then be compared across versions in plots displayed on a dynamic web page run by a `cherrypy` [244] web server. The whole process is depicted in a flowchart in Figure 4.4.

While the validation package can also be run locally to compare arbitrary software versions, a central service that compares nightly and predefined versions is offered for convenience. A buildbot [245] instance schedules and manages the required runs of the validation framework. Besides visualizing differences between software revisions, automatic comparisons with references are performed. If a significant difference is spotted, the comparison is highlighted, and emails are automatically sent to contact persons. Furthermore, software quality shifters regularly scan the results, determine if an action is required, and follow up with the package experts to ensure any discrepancy is well understood and possibly resolved.

I have been serving as the librarian (principally responsible person) since end-2018. During this time, I have implemented various stability and usability improvements, extended the comparison functionalities to avoid false positives, provided user support, and improved the overall quality of the source code.

4.2 A new Software Training Model at Belle II

By providing a straightforward python interface, the Belle II Software Analysis Framework presents significant usability improvements over its predecessor at the Belle experiment (basf [246]). Nonetheless, the sheer amount of available functionality can be overwhelming for new members of the collaboration who do not possess an overview and general sense of orientation yet. For this reason, induction activities and materials are essential.

This section presents a paradigm shift in the training material and model that I have coordinated as convener of the Belle II software training and documentation group, which I have led since mid-2020.

Much of the content has been presented at the ACAT 2021 conference [247] and is documented in its proceedings[1]. As such, the text of this section has significant overlap with the proceedings. The proceedings were entirely written by myself (but for the implementation of minor suggestions by my co-authors), and verbatim copies are not marked explicitly throughout this section. This particularly applies to Sections 4.2.1 to 4.2.5, while the remaining content is presented for the first time. In particular, the results of several surveys are presented and discussed to inform the decisions of the next convener and their team.

4.2.1 Software training at Belle II

Each year, around one hundred new members join the Belle II experiment. Among these are students working on their Bachelor's, Master's, or Ph.D. projects and more senior physicists. While their previous experience and knowledge vary greatly, everyone shares one goal: to make progress with their projects as soon as possible. At the same time, the amount of information collected in the various spaces of the collaboration is so overwhelming that the proverb "to drink from a firehose" might best describe the learning experience of many newcomers.

To ensure a smooth start for everyone, newcomers need dedicated training material and support from more experienced members. Because of the central role of software in experimental high energy physics, this particularly concerns the experiment-specific software frameworks.

Events for beginners are also one of the best opportunities to raise awareness of best practices: After newcomers get started with their scientific projects, their focus tends to narrow, and they become more challenging to reach. General recommendations and well-written code examples in the tutorials can significantly improve long-term success.

4.2.2 Pivoting to a self-study friendly training model

The recent remodeling of the Belle II software training material was triggered by the Covid-19 pandemic, which made in-person training events impossible. However, even with unimpeded travel and public meetings, holding training primarily in the form of live workshops still has several drawbacks:

- **Divergent needs:** The different levels of seniority, previous experience, and knowledge result in very different learning speeds, which require different sessions/tracks and more personpower to run the event. While there are some ways to help individual participants struggling with particular problems at in-person events, this is significantly more difficult at online events.
- **Logistics, funding and the environment:** For in-person events, the required travel logistics and funding can limit both the number of attendees and the frequency of training

[1] Currently in review; to be published by IOPscience in the *Journal Of Physics: Conference Series*, see page 157; draft available at [248] (Belle II Internal).

events. The amount of air travel required also causes a significantly larger ecological footprint.

- **Scheduling:** Due to the different academic schedules of the more than 100 institutions at Belle II, newcomers join throughout the year. So far, logistical and personpower considerations have limited the number of training events to three events per year. As a result, most newcomers were forced to start their research long before attending. This causes an even more significant divergence of previous experience and diminishes the efficiency of the training.
- **Duration:** Logistics and personpower limit the training events to several days. As a result, the information presented at the events is necessarily very compressed and can quickly become overwhelming.

Taking the large-scale disruption caused by the pandemic as an opportunity, we decided to pivot the focus of the training group on material that allows efficient self-study.

4.2.3 Challenges for training material

In general, the training material is subject to the following challenges:

- **Versioning:** The analyst-facing interface of the Belle II software is still evolving. New software versions can introduce backward-incompatible changes or change the recommendations and best practices.
- **Testability:** Wherever possible, code snippets in the lessons should be tested automatically. This is particularly important when changing the recommended software versions. Requiring passing tests might also be an incentive to keep the material maintainable and on-topic, reducing the aggregation of outdated legacy material and other forms of software erosion.
- **Maintainability and Sustainability:** Lessons should be consistent, stable, and easy to update. This is an issue for more complicated code snippets incrementally built up within a lesson or building on top of code explained in a previous lesson, as changes need to be propagated carefully.
- **Interactivity:** Exercises can greatly improve the learning experience and keep readers engaged. However, it is crucial to strike the correct difficulty level for each individual. Because of the diverse audience, this can only be achieved by providing additional optional hints for each exercise, or exercises of different difficulty levels. Complete solutions should be available for all exercises.
- **Connecting resources:** Enabling newcomers to use the existing documentation resources is one of the objectives of the training itself. At the same time, referring to other resources allows the training material to be more concise and maintainable. However, links to external material can be brittle when resources move or change.

4.2.4 General principles

To face the challenges outlined in the previous section, we adopted the following principles:

- **Didactic style:** The style of the material is inspired by the work of the Software Carpentries [249–252], the work of HSF training [227, 253] and by the LHCb StarterKit [254]. The material is organized in lessons. The beginner lessons are to be studied in order and build on top of each other, while intermediate and advanced lessons are designed to be independent. Lessons consist of verbose text, code snippets, and exercises. A selection of callout boxes provides extra information, overviews, or summaries.

- **Versioning:** The training material is hosted in the git repository of the main Belle II software (basf2). This enforces a natural correspondence of software versions and documentation/training versions. It also makes failed unit tests block the merging of pull requests: It is impossible to merge a change if it breaks one of our lessons! Reviewers of pull requests can also require the opener to make necessary updates to the training material. In short: We avoid the training material becoming an afterthought of the software development by bundling them together. As a nice side effect, this also increases the visibility of contributions to the training material.

- **Avoiding redundancy:** Wherever possible, we refer newcomers to the API documentation and similar pages, either by making it an exercise to find a piece of information or by directly linking to it. Because the API documentation and most other technical documentation are generated with the same system and hosted in the same repository, links remain functional, even for older software and training material versions.

- **Software prerequisites:** For the basics of bash, git, and python, we refer to the training material from the software carpentries but provide additional exercises with which newcomers can test their knowledge. We further extend the python training by lessons on the pandas data analysis framework. In addition, we provide a lesson on SSH.

4.2.5 Technical details

The implementation of the training material is based on the following technical solutions:

- **Generation:** The training material is available as web pages that are rendered with the Sphinx documentation generator [255]. reStructuredText [256] is used as markup language. We have created custom callout boxes for exercise blocks, foldable hints and solutions.

- **Rollout:** The training material is developed via pull requests to the main branch of the basf2 git repository. The training material from the main branch is built on a nightly basis ("development version"). The training material corresponding to releases of the software is built from the corresponding *release branches*. Major versions branch off the main branch; new commits on these branches are published as minor and patch versions. The *recommended* training version is that of the latest release. This means that updates to the training material reach the recommended version with every major release. However, hotfixes or substantial improvements can also be cherry-picked to a release branch and then included in a minor or patch release. Alternatively, newcomers can also be pointed to the development version for specific updates.

- **Preview:** Opening a pull request (and pushing additional commits) triggers a build that runs checks and gives access to a website preview. However, these builds take between 30 and 45 minutes to complete (depending on the overall resource use). Local builds are faster (1-3 minutes per build) but require an initial compilation of (part of) the main software.

- **Code inclusion:** Generally, code snippets are kept in separate source files and are included in the lessons via sphinx directives. This has several advantages:

  - The files can usually be executed as unit tests. When unit testing is not feasible, static code-checking can still ensure some level of correctness.
  - Formatting tools can be used to enforce a consistent coding style.
  - The file can be included partially in multiple places to build a more extensive example.

  Partial code inclusion is achieved with the `start-after`/`end-before` directives of sphinx that include code after/before a particular search string is found in the file. To avoid ambiguity, we use short marker strings included in the comments as search strings rather

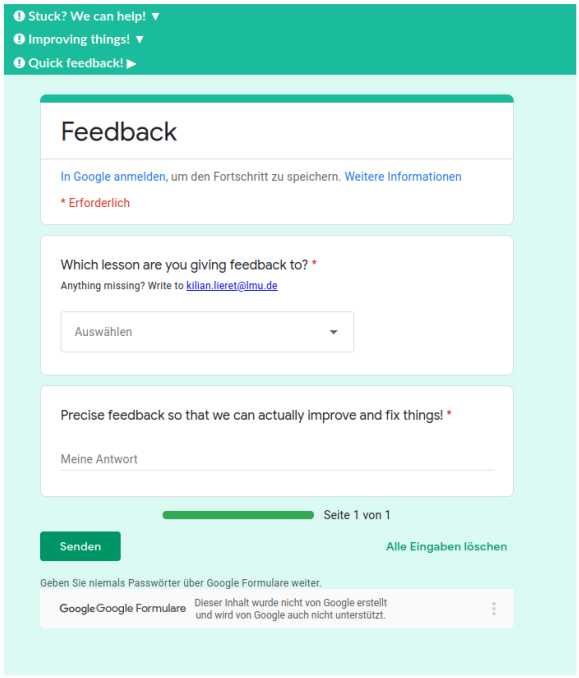


Figure 4.5: Footer including the form for quick feedback submission.

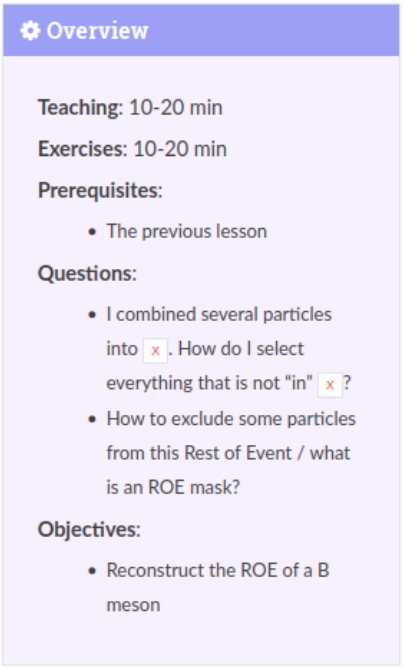


Figure 4.6: Overview box.

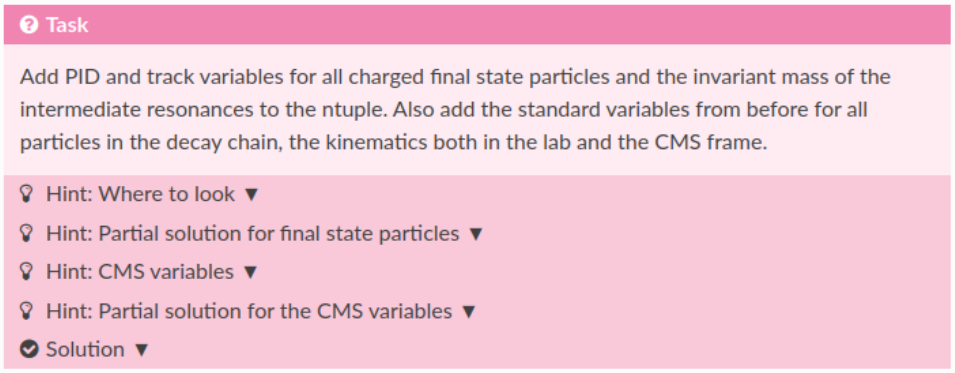
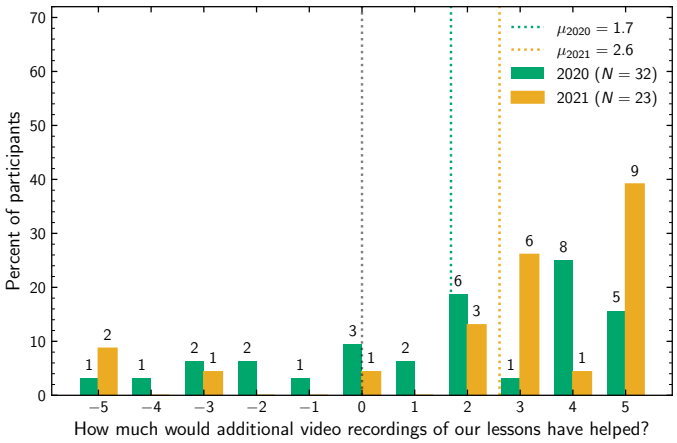


Figure 4.7: Stacked boxes for exercise, hints, and solution.

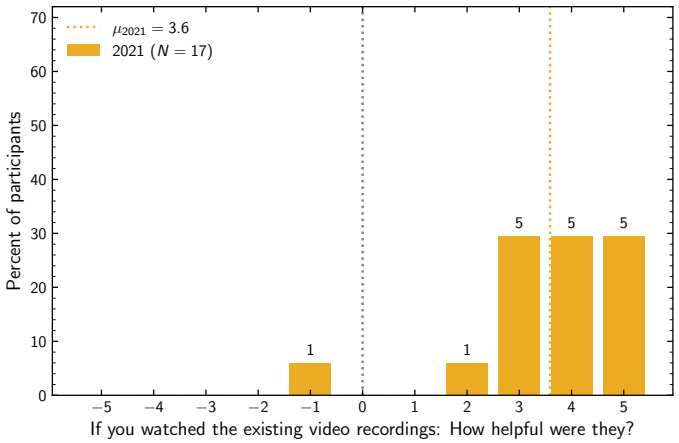
than parts of the code itself. These comments also make it evident which parts of the code are separately included while reading the file. This code inclusion scheme avoids the use of line numbers which are brittle and cumbersome to maintain.

- **Quick feedback:** Below every lesson, a set of three expandable containers provides guidance on where to get help with further questions, report problems, and how to contribute to the lesson. In particular, a tiny google form allows newcomers to report issues with the lesson anonymously with just four clicks. Entries in this google form are regularly checked and, if actionable, converted to JIRA issues that the training group uses for internal organization. However, despite the ease of use, we have received comparatively few reports through this channel.

Contributions to our training material require basic knowledge of git, pull request workflows, reStructuredText and sphinx. Acquiring these skills can be a steep learning curve, especially for newcomers – though it is very rewarding because it teaches hands-on experience in the entire software development workflow. We hope that additional step-by-step tutorials and dedicated hackathons will help grow the number of unique contributors in the long term.



(a) "How much would additional video recordings of our lessons have helped you? E.g., a full walk through of the first lessons."



(b) "If you watched the existing video recordings: How helpful were they?"

Figure 4.8: Survey results regarding the use of video recordings to supplement the lessons.

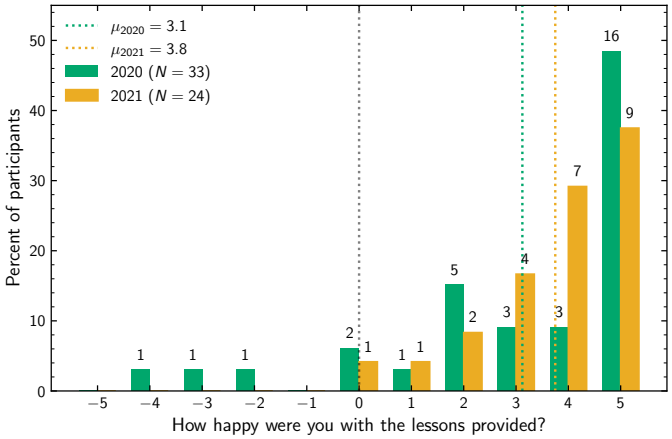


Figure 4.9: General satisfaction with the lessons. Positive numbers indicate satisfaction.

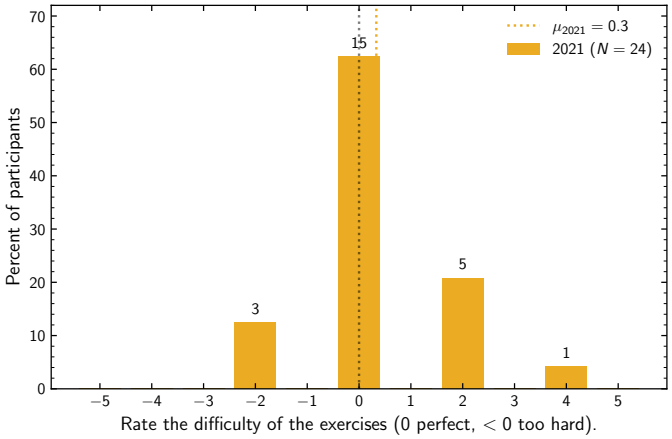


Figure 4.10: Perceived exercise difficulty.

4.2.6 Video recordings

Our surveys show that videos to complement the lessons would be welcomed by the majority of newcomers (see Figure 4.8a). In order to easily include the videos into the Sphinx setup, they need to be accessible publicly (that is, accessible without additional passwords or login requirements – they can, however, be hidden from search engines). After evaluating several options, we have chosen YouTube as host. Generally, however, videos have the disadvantage of being time-consuming to create and very difficult to update, restructure, and improve later on. For example, most video hosts (such as YouTube) do not allow for the overlay of text banners that could quickly alert students of a mistake in the video. Because it is also not possible to replace a YouTube video without changing its URL, we use the URL shortener *rebrandly* [257] to be able to quickly switch out destinations (without having to propagate URL changes into the different versions of the material).

In 2021, several shorter videos focusing on collaborative tools were uploaded and very positively received (Figure 4.8).

4.2.7 Feedback on the training material and future developments

The overall happiness has been very high: In 2020, almost half of the participants gave it the highest ranking available (Figure 4.9). The three negative ratings in 2020 did not leave concrete suggestions, but by 2021 not a single participant gave a negative rating, and the average score improved slightly.

The strategy of having exercises of various degrees of difficulty but providing ample hints seems to have paid off, with more than 60 % of newcomers rating their level as "perfect" (Figure 4.10). Of the remaining votes, most perceived the exercises as slightly too easy, which could be amended relatively quickly by adding additional optional exercises that are more open-ended.

Despite these successes, the material must keep evolving. The coupling of training material and software seems to have been very successful in ensuring the maintenance of existing material (taking care of breaking changes and changed recommendations). However, *expansion* of the material requires more extensive efforts. A promising avenue is to advertise such work as service tasks[2]. Another possibility is the organization of dedicated *hackathons*, such as done in 2021 [258]: this helps contributors block off time for documentation efforts and generally feels more motivating.

While the current material covers all fundamentals of basf2, topics of offline analysis (that is, analysis after data processing with basf2, such as fitting, plotting, and more) remain largely incomplete. These topics are challenging because no single common framework has been established for these tasks, and different software stacks are used. However, this is also an opportunity: If we settle on concrete software recommendations paired with sufficient documentation and training material, the recommended frameworks will gain increasing traction within the collaboration and slowly establish new standards.

Because these topics are less experiment-specific, there are also more options for collaboration, for example, with the HSF Training group (Section 4.3).

4.2.8 Training events

Week-long online training events ("StarterKits") were offered with the new material in 2020 and 2021 [259, 260].

An important finding is that most newcomers only have around half of their time available for the workshop on average (Figure 4.12a). However, this availability varies greatly between participants: In 2020, the histogram showed an almost uniform distribution between 0 % and 100 %! This might make it challenging to combine self-study lessons with live sessions. Generally, the workshop duration of one week is just right for the majority, with everyone else preferring a longer time span (Figure 4.12b).

The event registration is managed with indico [261]. Registration starts around a month in advance and is possible until the end of the event. There is no limitation on the number of registrants. The primary use of the registration is to collect time zone information (to schedule the live sessions) and email addresses (to send reminders to participants). For detailed information, a central page on the wiki system Atlassian confluence [262] is linked and advertised in the emails. In order to participate, a number of different accounts are necessary. Participants are asked to complete the necessary registrations as early as possible.

In 2021, participants were also asked to complete the first three chapters of the training material (collaborative tools, physics background, and software prerequisites) before the event. Completing these lessons can take up to three days of full-time work but varies significantly between participants because of the different levels of prior knowledge. Ensuring this minimal level of

[2] That is, tasks that Belle II collaborators complete in order to get added to the list of collaboration authors.

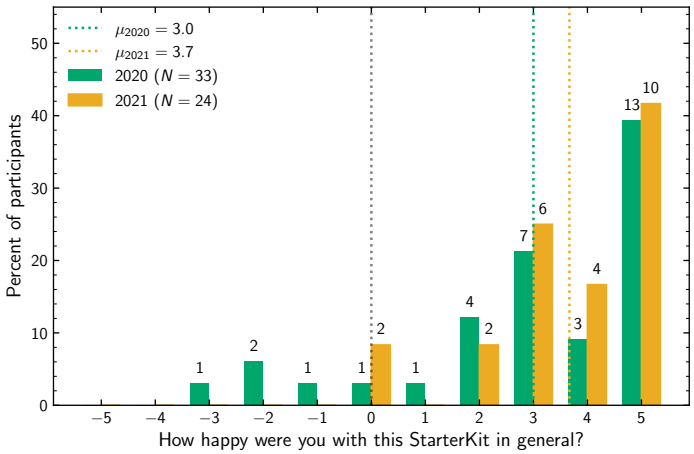


Figure 4.11: General satisfaction with the StarterKit. Positive numbers indicate satisfaction.

prior knowledge helps to synchronize the learning speed of participants during the events. It also frees up more time for the most important part of the workshop: the experiment-specific software.

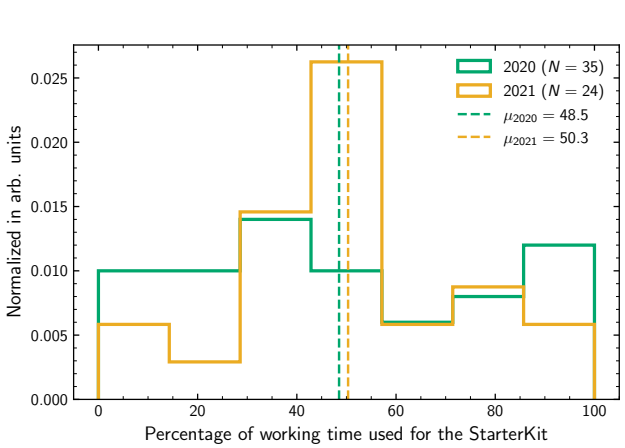
The first session of the workshop week is the *kickoff session* (offered two times for different time zones). The kickoff session emphasizes essential organizational information and tries to connect participants among each other. However, the most important purpose is to ensure that everyone has a setup that allows them to continue with the exercises. In 2021, we added a series of elementary exercises that helped to make the session more active and significantly improved the ratings (Figure 4.13a). The majority of participants will not have any issues, while some will require considerate technical help. Therefore it is best to keep a certain pace during the exercises and help with complicated issues after the official part of the session ends (possibly using breakout rooms). It is also advantageous to have experienced members who use different operating systems available (particularly for Windows).

The workshop primarily consists of self-study time, during which help is provided via chat channels. We also offer several Q & A sessions. Because of the completeness of the lessons provided, the need for additional guidance has been low (this was also confirmed in several surveys). For this reason, we have filled the Q & A sessions with additional content (code-along exercises and more) in the 2021 iteration, which improved ratings significantly (Figure 4.13b).

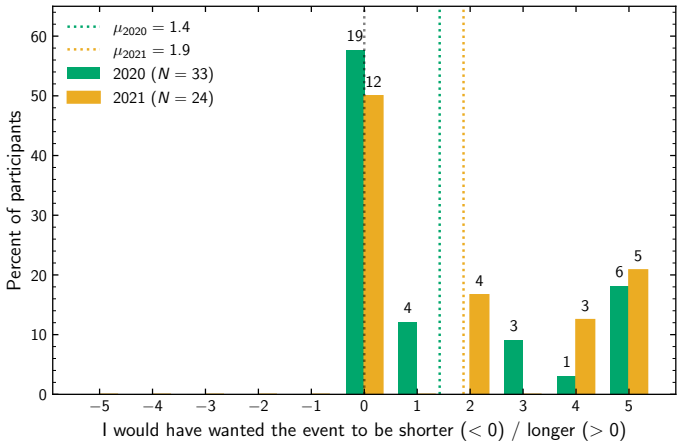
The last day of the StarterKit workshop contains *mentoring sessions* in which students are paired with experienced members in small groups (up to five students per mentor). Depending on the individual dynamics, these sessions can take various forms, ranging from intense debugging and detailed questions to ask-me-anything sessions. Due to the already complicated scheduling process involving different time zones and availabilities of mentors, it is challenging to split sessions by the experience of the participants (this might explain some of the negative ratings in 2020, see Figure 4.13c). Around half of the students gave the mentoring sessions the highest rating, and the overall ratings still increased in 2021 (Figure 4.13c). Clearer expectation management on both sides might further improve satisfaction: Students who have already started their projects should be aware that very detailed questions might be out of the scope of the session. Mentors should be aware that many students do not bring specific questions but still appreciate general (and perhaps subjective) guidance that is harder to convey in the lessons.

4.2.9 Future events

The new training material has drastically reduced the need for guidance. As such, one might ask provocatively whether training events in their current form are obsolete and should be replaced

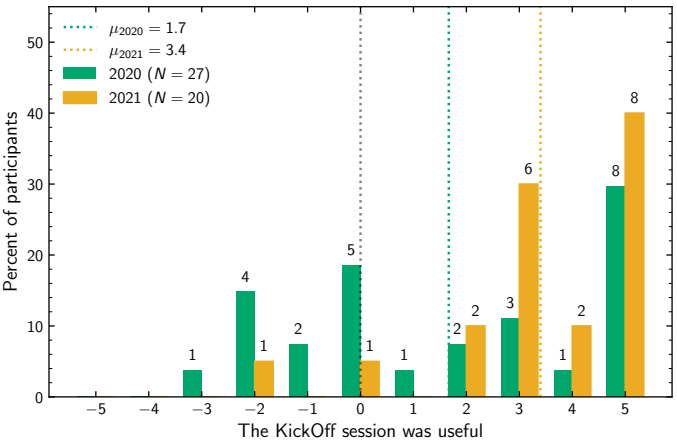


(a) "Which percentage of your working time did you use for the StarterKit during the workshop week?"

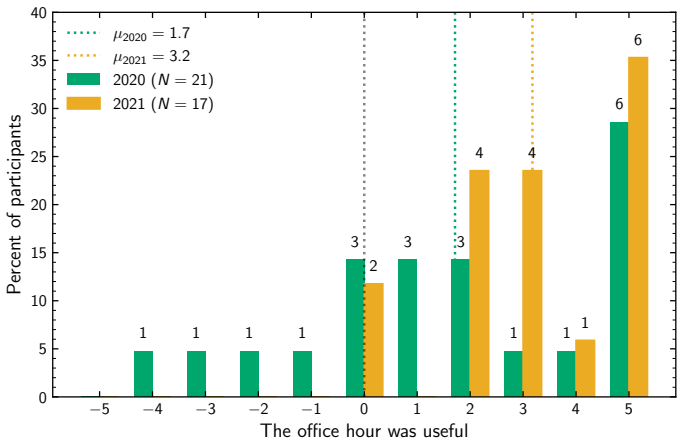


(b) "I would have wanted the event to be shorter / longer."

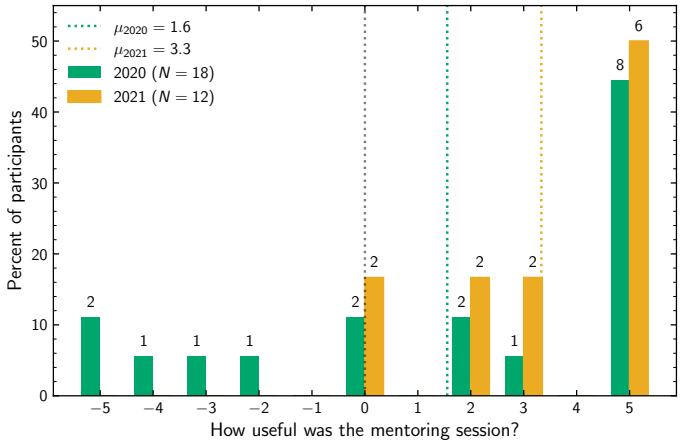
Figure 4.12: Survey results regarding the duration of the event and available working time.



(a) KickOff sessions

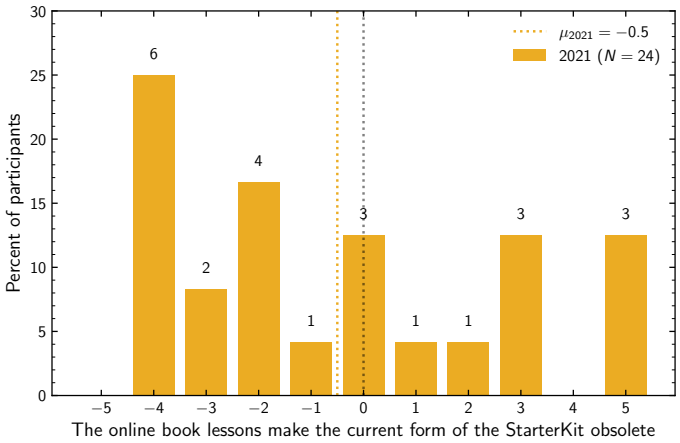


(b) Office hours

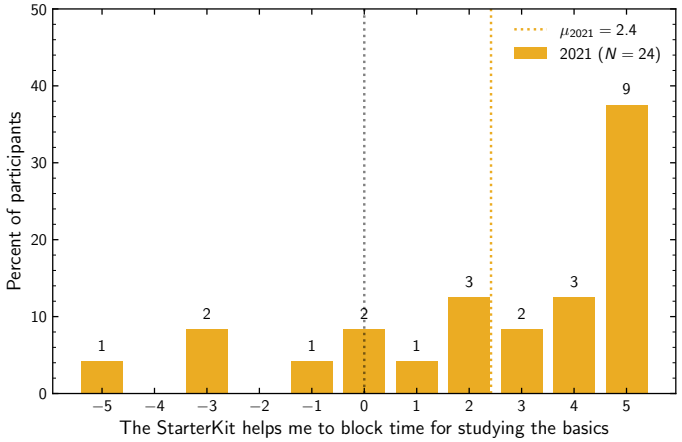


(c) Mentoring sessions

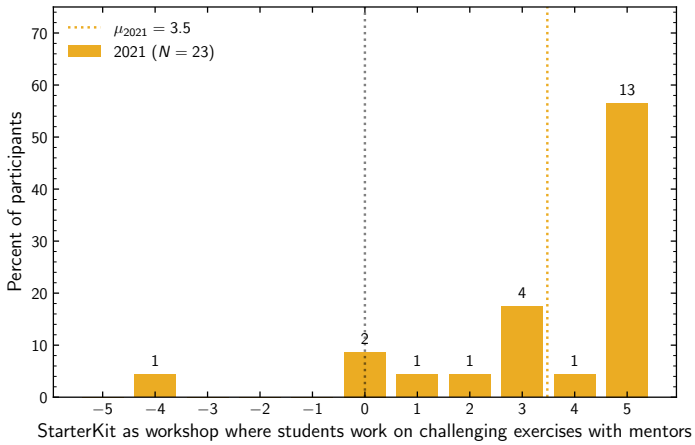
Figure 4.13: Ratings of the different live sessions of the StarterKit. Positive numbers indicate satisfaction.



(a) "The online book lessons make the current form of the StarterKit obsolete, because everything is explained there already. We should think of a different form of a workshop."

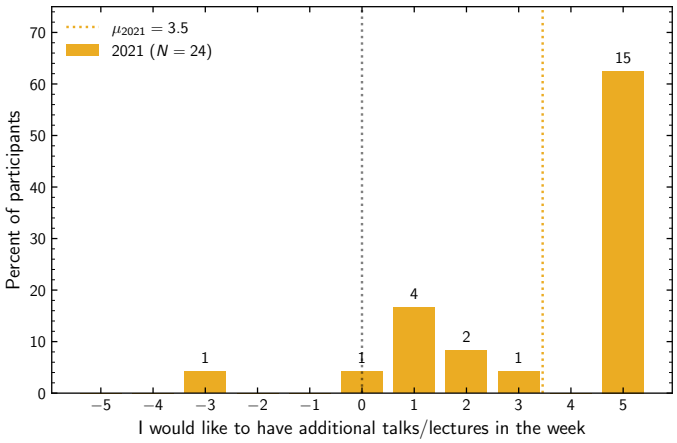


(b) "The StarterKit helps me to block time for studying the basics of the software (otherwise I'd be too caught up in working on my projects to learn about these things)."

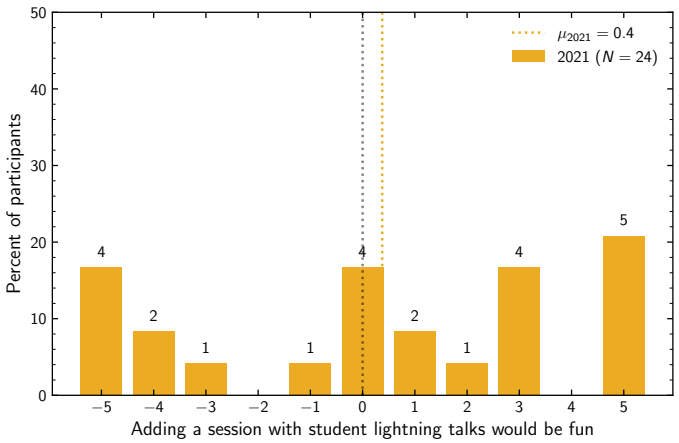


(c) "Especially when in person events are possible again: It makes sense to change the form of the StarterKit into a workshop where students work on more challenging exercises together with mentors. The online book should then be a prerequisite."

Figure 4.14: Survey results regarding the future of the StarterKit workshops. Positive numbers indicate agreement.



(a) "I would like to have additional talks/lectures in the week that e.g., introduce parts of the detector or talk about the physics. Note that these might not be in your timezone, so you might have to watch recordings."



(b) "Adding a session with student lightning talks would be fun. E.g. 5 minute talks for each student."

Figure 4.15: Survey results for additional activities during the StarterKit workshop. Positive numbers indicate agreement.

by a completely different kind of workshop. The audience of the 2021 StarterKit was relatively undecided on this question (Figure 4.14a). It should be noted that even if the workshop consists predominantly of self-study time, it still helps participants to block time for studying the basics (Figure 4.14b): Without a workshop, the immediate short-term progress of the project and other commitments is always be prioritized over the long-term investment of studying the basics.

Nonetheless, a possible future direction is to consider more and more lessons as requirements to be completed before our events and then focus on more complicated exercises to be solved in teams with mentors. This idea is received very positively (Figure 4.14c).

Focusing on more challenging exercises to be completed during the workshop would greatly profit from increased teamwork and communication between participants. This might be very challenging for online workshops: So far, the participants were perceived to be relatively passive in the live sessions (though this can be attributed to various reasons). However, using a virtual space like *gather* [263], where multiple teams can effectively share screens and communicate (based on proximity of the avatars) could help to foster a "workshop feeling". Randomly pairing participants for short "icebreaker sessions" with video chat (for example with *gatheround* [264]) might help participants to form connections even if they are shy in larger groups.

Independently of this, the majority of participants strongly support additional talks and lectures during the week (Figure 4.15a). If it is decided to follow the request, it would be advantageous to make the corresponding lecture slides collaboratively maintainable to avoid recurring efforts every year.

Finally, as another way to get participants more active, we have surveyed the interest in lightning talks given by participants. This idea was met with mixed reactions, and the fraction of decided proponents was relatively small: Such a session would probably not be worth the organizational effort.

4.2.10 Overall experiences

The drawback of our setup is its complexity: slightly more complex than the setup of the LHCb StarterKit or the material of the carpentries (mainly because it is not a stand-alone repository but integrates tightly with the rest of the software) and significantly more complex than using a

wiki system.

However, we think that this is more than justified by the advantages of our setup, many of which have been mentioned in the previous sections: By coupling the training material with the software and by performing unit tests, we ensure that all examples remain functional and keep the training material on the developers' agenda.

Most importantly, our material provides a very complete onboarding experience, even for newcomers who join "off-season". Lessons cover everything from basic physics knowledge, collaborative tools, and software prerequisites to submitting grid jobs. The lessons include almost 250 code snippets, 45 figures, and more than 50 overview boxes. More than 200 exercises (with more than 150 hints and 190 complete solutions) make the material engaging.

We have also received very positive feedback from newcomers (Figure 4.9). In the words of one of our participants:

> Very solid work regarding the textbook! Congrats! Everything was very clear which significantly minimized the need for guidance (...) Software Prerequisites was the highlight of the workshop as it summarized all the necessary tools that no one really spends time on explaining thoroughly to newcomers.

The entirety of our material is publicly available [265].

4.3 Coordination of Training Activities at the HEP Software Foundation

While the Belle II Training group focuses on teaching experiment-specific software, the capability of the collaboration just as much depends on its member's general knowledge of industry-standard tools and cross-experiment HEP-specific packages. For example, students without a basic understanding of Python will run into obstacles when writing steering files, no matter their understanding of basf2 modules. Furthermore, the scope of basf2 is limited and generally ends with the production of a flat table of event data information from which the physical results need to be extracted. Much of the individual analysis work thus depends on external software tools.

Luckily, these tools are either general-purpose or shared between different HEP experiments. Therefore, the corresponding education and training activities can also be shared and coordinated between different experimental collaborations. Such joint efforts not only increase the efficiency of software education for the entire field of HEP but create fertile cross-collaboration contact and broaden the horizon of everyone involved.

In this section, I describe activities that I helped coordinate as one of the conveners of the HEP Software Foundation in 2020 and have continued to support since then.

4.3.1 HSF and the Training and Careers working group

The *High Energy Physics Software Foundation* (HSF) [228] was founded between 2014 and 2015 [266]. Its key objective is the coordination and facilitation of common efforts in the field of software and computing for the HEP community. The HSF also supports the career development of computing and software experts and generally raises awareness for the importance of software in HEP.

As part of this effort, a *Training, Staffing, and Careers* working group was established right from the beginning. After several years, a series of whitepapers provided detailed roadmaps for HSF as a whole [267] and its training group [268] (among several other working groups). As of 2021, more than 1000 people have participated in events organized by the HSF training group [253].

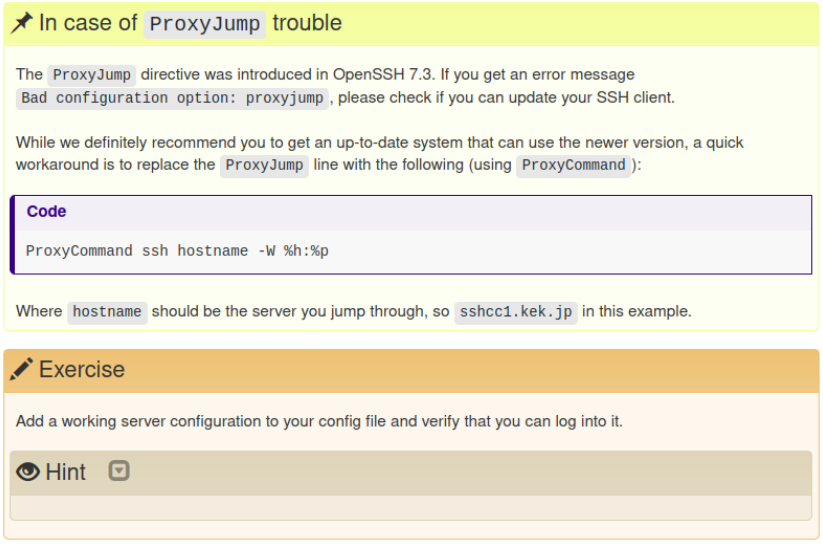


Figure 4.16: Styling of additional information, exercises, and hints with the Jekyll template of The Carpentries. Shown is a part of the SSH lesson that was initially developed as part of the Belle II StarterKit and is considered for adaptation within the HSF Training curriculum.

4.3.2 Training Material

All of the training material maintained by the HSF Training group is open source and hosted in a central GitHub organization [269].

The material is not entirely of a homogeneous format. For example the very popular C++ training material [270] is a collection of slides written in LaTeX beamer [271], and the training material for a K12 STEM teacher outreach program [272] uses notebooks that run in the Google Collaboratory [273].

Generally however, the HSF training group embraces the style of The Carpentries [249]: the training material is provided as self-contained web-pages, including verbose explanations, exercises and solutions. The web-pages are automatically built via *GitHubPages* [274] using *Jekyll* [275] (a general-purpose static page generator) and a template [276] adapted from The Carpentries [277]. All content is written in Markdown [278] with some extended syntax elements to provide additional meta-information and adjust formatting for exercises, solutions, and similar elements.

The entry-barrier for contributions is lower than in the setup of the Belle II training group (Section 4.2.5):

- Markdown is simpler and more widespread than reStructuredText
- The standalone repositories are lightweight and less complex
- The GitHub web interface allows to make most edits without a local clone of the repository
- The automatic builds via GitHub pages are available within less than a minute
- The incremental local builds only take a few seconds

However, the more sophisticated code inclusion and interlinking capabilities of sphinx would be missed in training modules that contain a large number of more extended code snippets (which is not an issue for most of the training modules offered).

An overview of all available training material is given in the HSF Training center [279] shown in Figure 4.17, which serves as the student-facing entry point.

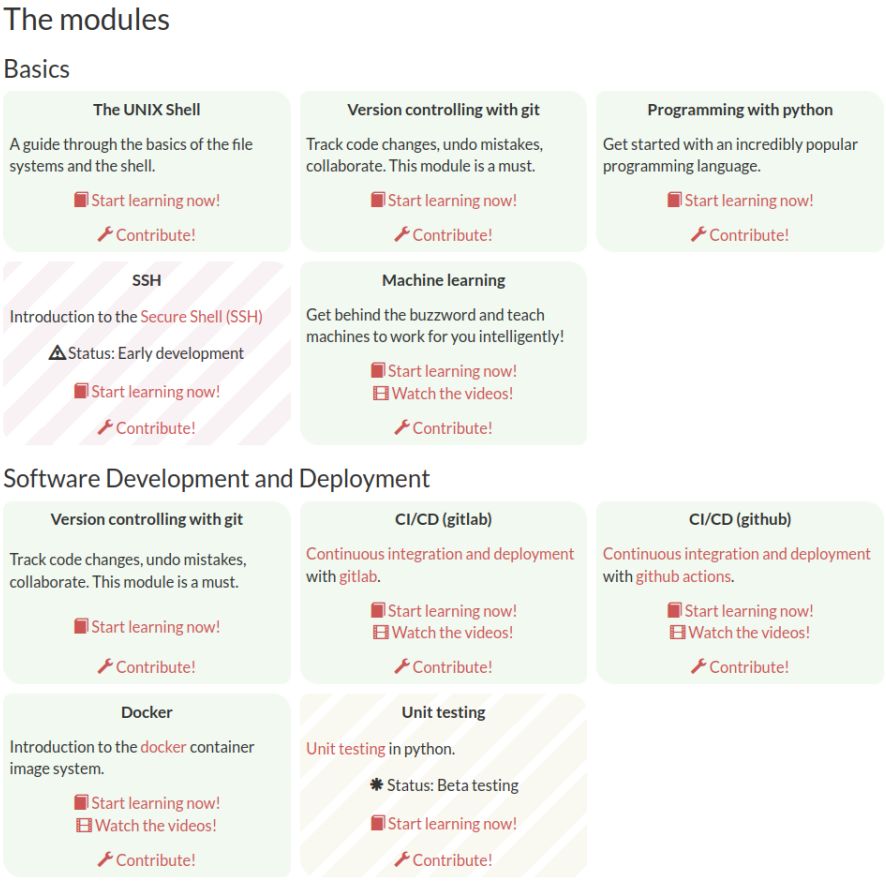


Figure 4.17: Screenshot of the HSF Training Center (truncated).

4.3.3 Training events

An important step in 2020 was the formalization of HSF Training events and their organization. We have compiled a thorough write-up of experiences and practical advice [280], laying the groundwork for scaling up our training activities in the future.

It is important to divide tasks between individuals at an early stage of the organization process. We distinguish between the following roles:

Instructors *Instructors* are subject-matter experts and develop the training material. If the training includes live lessons, they are the teachers; if the training is taught via videos, they record the videos. They are the primary academic drivers of HSF Training and gain experience in curriculum design.

Mentors All events of HSF Training are meant to be hands-on. To ensure that all participants make optimal progress and are not stalled by technical issues throughout the event, we need to provide them with assistance and support. *Mentors* work closely with participants and optimize the learning environment for individual participants. A sufficient mentor to participant ratio is required (typically, we calculate with one mentor per five students[3]). They are familiar with the subject matter but do not necessarily need to be experts. More importantly, they need to be open, motivate students to persevere, and acquire skills on the fly. They gain communication and pedagogy skills that complement their teaching portfolio.

[3] For online events, the $\sim 50\,\%$ attrition rate needs to be taken into account, see below.

Facilitators *Facilitators* are the primary organizers of the event. They arrange meeting venues, announce the event, oversee the registration process, and are the primary reference point for participants to communicate. They acquire soft skills qualifying them for leadership in the academic community.

All three roles are collectively referred to as *educators*. As of 2021, more than 100 educators have helped in the events of HSF Training.

In 2020, the COVID-19 pandemic made in-person training impossible, and HSF Training strategy had to be rapidly adjusted to virtual training events. Despite this transformation, seven events were organized throughout 2020:

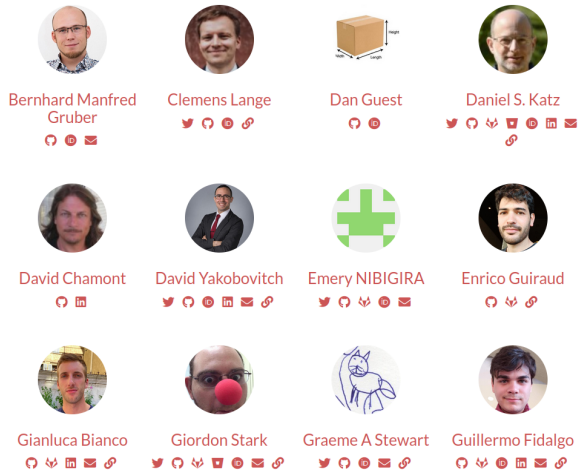
- *Virtual Pipelines Training* [281]
- *GitHub CI/CD Training* [282]
- *Virtual Docker Training* [283]
- *ML + GPU Training* [284]
- *HEP C++ Course and Hands-on Training* [285] (in collaboration with SIDIS [286])
- *Virtual Docker Training* [283]
- *Data Analysis for STEM teachers* [287]

Valuable lessons have been learned for both in-person and virtual training. In an effort to share our experience, HSF Training provides a detailed guide [280] for anyone organizing a training event.

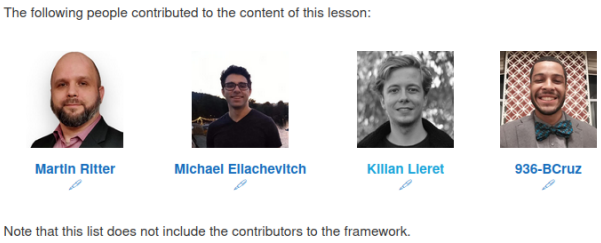
[4]While in-person events offer more opportunities for active and efficient engagement of participants and community building, they are generally more exclusive: Participants need sufficient funding and extra preparation time to arrange travel to the venue. Hosts have to book specially arranged/equipped rooms with multiple projectors and screens to show teaching materials and slides simultaneously. The space constraints typically limit the number of participants to a few dozen, and a long lead time is required for the logistics. Our in-person events have been managed by about five educators, which is necessary for the "hands-on" aspect to be successful. These educators also need to make a significant time commitment; they cannot just present their material and leave. Virtual events have a broader reach of participant attendance that is much higher than in-person events and enable a considerably more equitable service to the community. Because the teaching materials are fully preserved via lesson creation and YouTube videos beforehand, an inability to attend during the scheduled time does not considerably degrade learning. Finally, these video materials are captioned to be inclusive of those with hearing impairments. Captioning videos for a week-long event (∼$50/day) is considerably more economical than the cost of a hired sign language interpreter (∼$1000/day).

However, the disadvantage of virtual events is that it is difficult for educators and participants to interact closely – you just cannot recreate the in-person environment on Zoom. Educators and participants have to plan and act upon their spread across time zones in the best possible way. It is also challenging to keep everyone engaged and on the same page due to the pervasive culture of "multi-tasking" within HEP. Due to this issue, although initial registrations for these events are very high, the actual attendance is typically only 50% of those registered. In addition, the online experience is more prone to be distracted by other professional duties. However, it

[4]This and the next paragraph are near-verbatim copies of our paper [253], of which I am one of the primary authors.



(a) Educator profiles on the HSF Training Community page [291].



(b) Adding contributors to individual lessons via the allcontributors bot

Figure 4.18: Acknowledging contributors with profiles.

should be noted that this does not mean that there is a lesser degree of learning occurring at the training event. Tools like Mattermost, discord, and Slack have been effectively deployed for asynchronous communication, both during and after the event.

4.3.4 Community building

Creating training material, teaching, mentoring, and organizing training events requires time, persistence, and commitment. While IRIS-HEP [288] and FIRST-HEP [289] have kindly sponsored travel costs for educators, most of this work is done voluntarily.

[5]As the success of our mission thus depends crucially on the motivation of the community, we cultivate a strong sense of community ownership and pay special attention to acknowledging contributions of all kinds. We also encourage the participants in our training events to remain active or become more active, share feedback, and in particular, to sign up to be a mentor in one of the subsequent iterations of the same training module. If former participants do not yet feel confident about their mentoring skills, we offer to match them with a more senior mentor. In the same way, we encourage mentors to become instructors or facilitators and become more and more active in our organization. By actively engaging participants and educators throughout the training community, we can sustain and nurture a culture of intentional learning and grow our community in an organic fashion [290].

To help establish a sense of community while also acknowledging the contributions, we have created a community page [291] in which all educators are listed. Besides linking profiles and contact information, educators can choose to create a small profile with a self-introduction.

[5]This paragraph is an almost verbatim verbatim copy of our paper [253], of which I am one of the primary authors.

Furthermore, we are currently experimenting with using the *all contributors bot* [292] to easily add individual contributors to the lesson webpages (Figure 4.18).

4.3.5 Future

The HSF training group is still in the process of scaling up its training activities and extending the available training material. As part of my personal plans, I want to incorporate my lectures on Software Paradigms and Programming Patterns [293] into the curriculum. So far, I have taught them to more than 500 people at the inverted CERN School of Computing [294] and at LMU, but adaptation to the self-study optimized didactic style of HSF is required.

A profitable strategy to increase the number of contributions from the community is the organization of hackathons that bring together members to work together over a short period. I organized the first hackathon of this kind in 2020 [295] and this format has since been successfully repeated several times. For maximal participation rates, it is recommended to organize these events directly after (or as part of) another workshop that already attracts skilled community members.

At the same time, it is vital to establish and foster contacts with and between the training groups of the various HEP experiments. Therefore, in late 2021 I organized a round table workshop [296] to share experiences and to discuss training needs and plans for the coming years. Part of this workshop was also the completion of written profiles for all participating collaborations, including detailed contact information. Establishing such liaisons will help advertise training opportunities to students and find contributors for the various upcoming projects.

Chapter 5

Measuring $|V_{cb}|$ using hadronically tagged $B \to D^* \ell \nu_\ell$ decays

This chapter introduces a measurement of $|V_{cb}|$ using hadronically tagged $B \to D^* \ell \nu_\ell$ decays with 711 fb^{-1} of data at the $\Upsilon(4S)$ resonance recorded by the Belle detector. Fits of Asimov data, analyses of sideband data, and blind resolution studies show that the analysis preparation is well-advanced. The missing piece for the unblinding of the data and the completion of the $|V_{cb}|$ measurement is calibrating the tag side reconstruction efficiencies, which will be discussed in more detail in the following chapters.

5.1 Revisiting hadronically tagged $B \to D^* \ell \nu_\ell$ at Belle

The analysis introduced in this chapter is not the first study of hadronically tagged $B \to D^* \ell \nu_\ell$: In 2016, previous work by Saskia Falke used the Belle software framework and the FR for a very similar study. The effort was documented internally [297], and the first results were shown at the CKM conference [121]. However, the project was not fully brought to publication after Falke left the collaboration shortly afterward.

As part of my master's thesis [124], I repeated most steps of her analysis. Besides implementing additional cross-checks and correcting the calculation of several systematic uncertainties, I extended the analysis to separate the electron and muon channel. The lepton-flavor separated results can be used for cross-checks and, more importantly, can contribute to the search for new physics [298].

However, by 2018, the `b2bii` conversion package was well-established, and the hadronic FEI tagging showed impressive efficiency gains over the FR. Preparations had also started for the hadronically tagged reanalysis of $\mathcal{R}(D^{(*)})$ on Belle data with Belle II software (see Section 1.3). As $B \to D^* \ell \nu_\ell$ is the normalization channel of the $\mathcal{R}(D^{(*)})$ analysis, a dataset for the $|V_{cb}|$ measurement could be produced as a byproduct. Besides profiting from the efficiency gains by the FEI, this new $B \to D^* \ell \nu_\ell$ dataset also includes additional reconstruction channels and many additional corrections to the simulated dataset. In light of these developments, it was decided to abandon the previous analysis and instead focus on measuring $|V_{cb}|$ using the new dataset. While Markus Prim has been preparing the $B \to D^* \ell \nu_\ell$ signal extraction strategies, I have been working on calibrating the tagging efficiencies. The analysis is documented in a Belle-internal report [299].

5.1.1 Dataset and Reconstruction

As mentioned, the $B \to D^* \ell \nu_\ell$ dataset is produced with the framework of the $\mathcal{R}(D^{(*)})$ analysis [83, 84]. All corrections to MC are applied accordingly. For additional side studies (and possible future analysis), a separate sample of $B \to D \ell \nu_\ell$ decays is also reconstructed. The dataset

D^{*0}		D^{*+}		D^0		D^+	
$\mathbf{D^0\pi^0}$	**65**	$D^0\pi^+$	68	$K^-\pi^+$	4	$K^-\pi^+\pi^+$	9
		$D^+\pi^0$	31	$K^-\pi^+\pi^0$	14	$\mathbf{K^-\pi^+\pi^+\pi^0}$	**6**
				$K^-\pi^+\pi^+\pi^-$	8	$\mathbf{K^-\pi^+\pi^+\pi^+\pi^-}$	**1**
				$\mathbf{K^-\pi^+\pi^+\pi^-\pi^0}$	**4**	$\mathbf{K^0_S\pi^+}$	**2**
				$\mathbf{K^0_S\pi^0}$	**1**	$\mathbf{K^0_S\pi^+\pi^0}$	**7**
				$\mathbf{K^0_S\pi^+\pi^-}$	**3**	$\mathbf{K^0_S\pi^+\pi^+\pi^-}$	**3**
				$\mathbf{K^0_S\pi^+\pi^-\pi^0}$	**5**	$\mathbf{K^0_SK^+}$	**0**
				$\mathbf{K^-K^+}$	**0**	$\mathbf{K^+K^-\pi^+}$	**1**

Table 5.1: Reconstruction channels used in the $B \to D^* \ell \nu_\ell$ analysis. Decay modes highlighted in boldface were not considered for the previous $|V_{cb}|$ analysis. The numbers to the right of the reconstruction channel columns indicate the respective branching ratios in percent.

	$B^0 \to D^- \ell^+ \nu_\ell$			$B^+ \to \overline{D}^0 \ell^+ \nu_\ell$		
Analysis	$\ell = e^-$	$\ell = \mu^-$	$\ell = e^-, \mu^-$	$\ell = e^-$	$\ell = \mu^-$	$\ell = e^-, \mu^-$
New	2461	2120	4581	4789	4253	9042
Previous	1249	1061	2310	0	0	0
Gain	2.0	2.0	2.0	∞	∞	∞

Table 5.2: MC yields for the signal component in the new analysis compared to the previous analysis. *Gain* denotes the event yield ratio.

and MC corrections are discussed in more detail in Chapter 7 (from the point of view of the $B \to X \ell \nu_\ell$ dataset used for the tagging calibration).

While the previous analysis deemed the $B^+ \to \overline{D}^{*0} \ell^+ \nu_\ell$ reconstruction mode to be not sufficiently understood to include it the final results, the new analysis intends to use both B flavors. Including the $B^+ \to \overline{D}^{*0} \ell^+ \nu_\ell$ decays into the $|V_{cb}|$ measurement is expected to improve the sensitivity for $|V_{cb}|$ significantly: Because the extraction of $|V_{cb}|$ from the measured kinematic distributions hinges on the extrapolation to zero-recoil ($w = 1$, see Section 1.4.1), the amount of data at low w is particularly important. As can be seen from Figure 5.4b, the reconstruction efficiency of $B^0 \to D^{*-} \ell^+ \nu_\ell$ events in the first w bins is dwindling, while it is substantial for $B^+ \to \overline{D}^{*0} \ell^+ \nu_\ell$. Furthermore, the total reconstruction efficiency is higher for $B^+ \to \overline{D}^{*0} \ell^+ \nu_\ell$ ($\epsilon_{\text{tot}} \sim 1 \times 10^{-4}$) than for $B^0 \to D^{*-} \ell^+ \nu_\ell$ ($\epsilon_{\text{tot}} \sim 5 \times 10^{-5}$) as well (Table 5.2). The addition of this decay channel is thus a major improvement.

The reconstruction modes of D^* and D mesons are listed in Table 5.1. There as well, we include substantially more channels than the previous analysis.

In total, the added channels and the increased tagging efficiency of the FEI result in a twofold increase in efficiency for $B^0 \to D^{*-} \ell^+ \nu_\ell$ on MC (see Table 5.2).

5.1.2 Signal extraction strategy and determination of correlations

The form factor parameters and the $\mathcal{F}(1)|V_{cb}|$ normalization are extracted from a simultaneous fit to the four one-dimensional marginal distributions of w, $\cos\theta_\ell$, $\cos\theta_V$, and χ (see Figure 1.12 for the definition of the angles and Equation (1.4.3) for the definition of w).

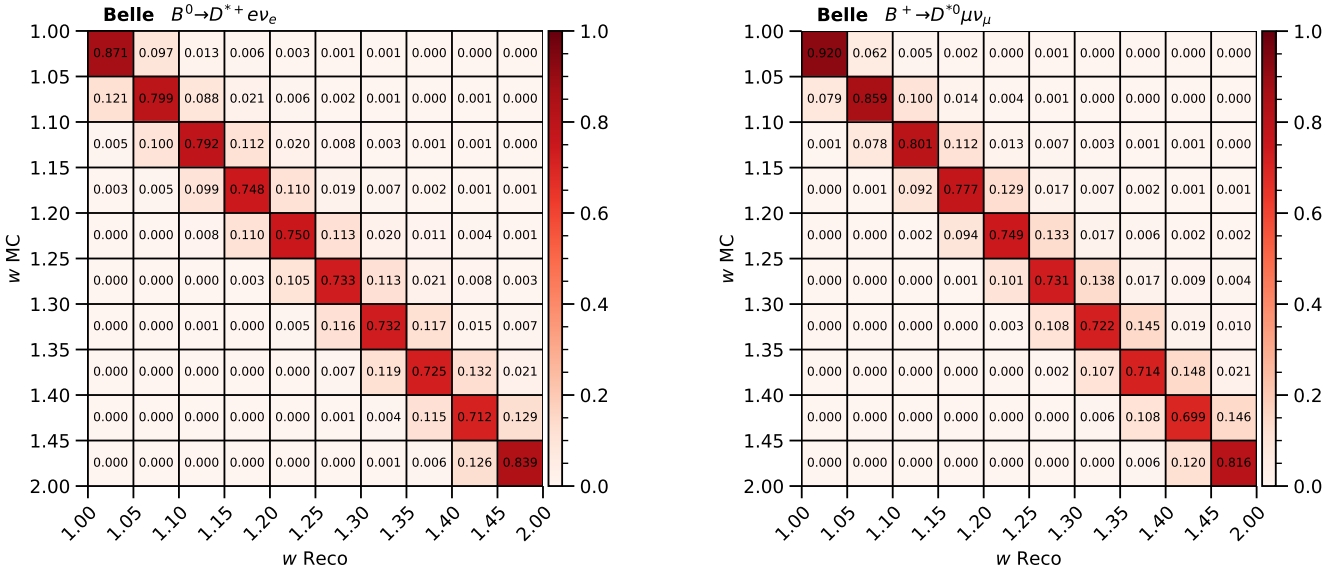


Figure 5.1: Migration matrices for two reconstruction channels. *w MC* denotes the true value of w. Figure from [299].

Therefore, differential signal yields $N_i^{(v)}$ ($v = w, \cos\theta_\ell \cos\theta_V, \chi$) have to be extracted for all four variables. To allow for the simple comparison of differential cross sections, we adopt the binning that has been used in previous measurements. Ten equidistant bins are chosen for the angular quantities in their respective ranges (-1 to 1 and 0 to 2π). For w, nine equidistant bins are chosen between $w = 1$ and $w = 1.45$ and an additional bin for $1.45 < w < 2.00$ is added. The range of the last bin is extended to recover all events that exceed the physical range of w due to resolution effects (the previous analysis limited the range to $w < 1.504$).

To measure $N_i^{(v)}$, the events of data and MC are projected into bins $b_i^{(v)}$ based on the measured values v_{reco}.

Due to detector effects, the true value v_{true} of v for an event can be different from the corresponding measured value v_{reco}. Therefore, an event can be counted towards $N_i^{(v)}$ while it should actually contribute to $N_j^{(v)}$ ($j \neq i$) based on v_{true}. This resolution effect is called *migration*. Using v_{true} and v_{reco} from the MC dataset, we can quantify this effect as a *migration matrix*:

$$(\mathcal{M}^v)_{ij} = \frac{\#\{v_{\text{reco}} \in b_i^{(v)} \wedge v_{\text{true}} \in b_j^{(v)}\}}{\#\{v_{\text{true}} \in b_j^{(v)}\}}, \tag{5.1.1}$$

where $\#\{\dots\}$ counts[1] all events that fulfill a certain condition. Two examples of such a migration matrices are shown in Figure 5.1. Strategies to take migration into account are summarized in Sections 5.1.3 and 5.1.4.

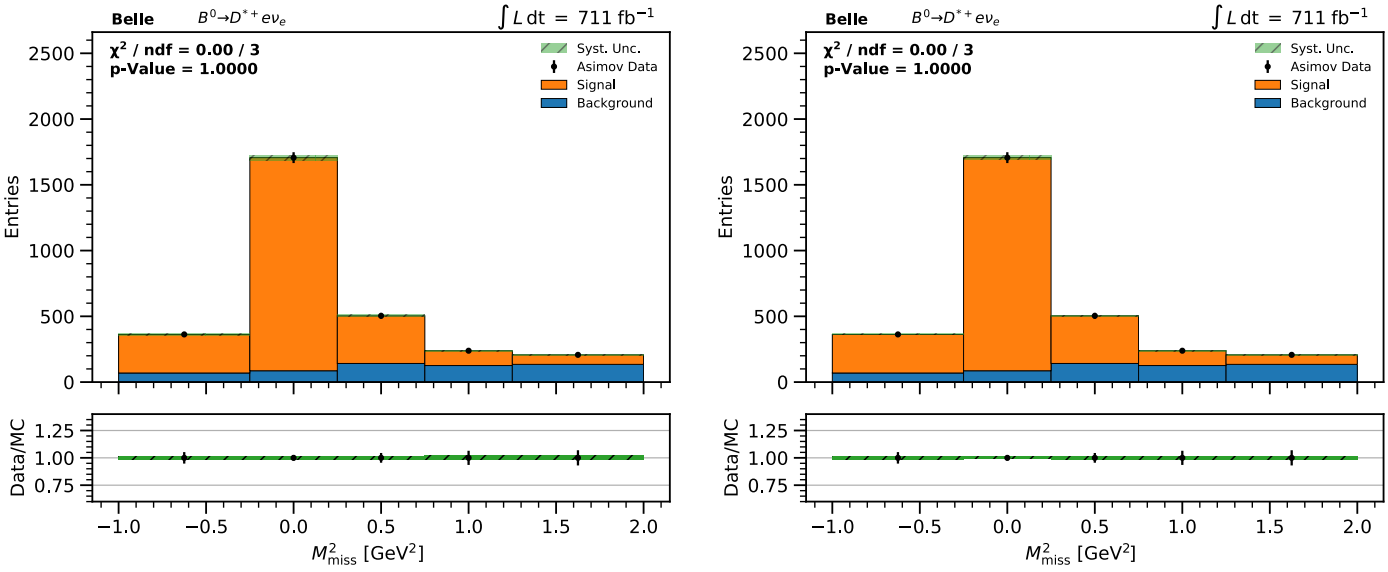
To measure $N_i^{(v)}$ based on the events in $b_i^{(v)}$, an extended maximal likelihood fit is performed (very similar to the setup described in Section 9.2.1). The *missing mass*

$$m_{\text{miss}}^2 := (p_{\text{B}}^\mu - p_{\text{D}^*}^\mu - p_\ell^\mu)^2 \tag{5.1.2}$$

is used as fit variable.

For correctly reconstructed $\text{B} \to \text{D}^* \ell \nu_\ell$ decays, this variable peaks at $m_{\text{miss}}^2 = p_\nu^2 = 0$. As in the previous analysis, the signal definition only checks that the lepton is correctly reconstructed and

[1] If not specified otherwise, *counting* is always meant to include event weights on MC.



(a) m^2_{miss} distribution before fit.

(b) m^2_{miss} distribution after fit.

Figure 5.2: Fit of $B^0 \to D^{*-} \ell^+ \nu_\ell$ with Asimov data. Figure from [299].

stems from a $B \to D^* \ell \nu_\ell$ decay but allows incorrect reconstructions of the D^*. This exception is made because these events also peak at $m^2_{\text{miss}} = 0$, and while the corresponding peak is somewhat broader, they cannot be separated well without introducing strongly anticorrelated event yields. The drawback is that the incorrect D^* reconstructions generally lead to higher migration between bins of the kinematic variables.

The signal extraction strategy has been fully implemented. To validate the implementation, we perform fits on *Asimov data* where the data yield of each bin is set to the corresponding MC expectation value. An example of such a fit is shown in Figure 5.2.

Because the four marginal distributions are measured with the same data events, they are correlated. Furthermore, systematic uncertainties are shared, and the fit templates use the same MC data. These three sources of correlation are investigated separately using different methods. Two examples of the resulting correlation matrices are shown in Figure 5.3.

5.1.3 Measuring $|V_{cb}|$

The principal method of measuring $|V_{cb}|$ from the measured differential cross sections has already been explained in Sections 1.4.1 and 1.5: We use an analytic expression of the differential decay rate that depends on form factors parameters and $|V_{cb}|$ and fit it to the experimentally measured event yields.

5.1.3.1 Reconstruction efficiencies

In order to compare decay rates with the measured event yields, we need to account for the reconstruction efficiencies: Only one in every $\sim 10^4$ signal B decays is reconstructed. Moreover, this reconstruction efficiency depends on both kinematic variables and the reconstruction modes.

In principle, we have a straightforward method of determining these efficiencies: On MC, we know the exact number of B decays (N_{BB}), as well as the assumed differential decay rates for all processes. We can thus directly compare the expected yield (based on the decay rates and N_{BB}) with the reconstructed yield to determine the efficiency.

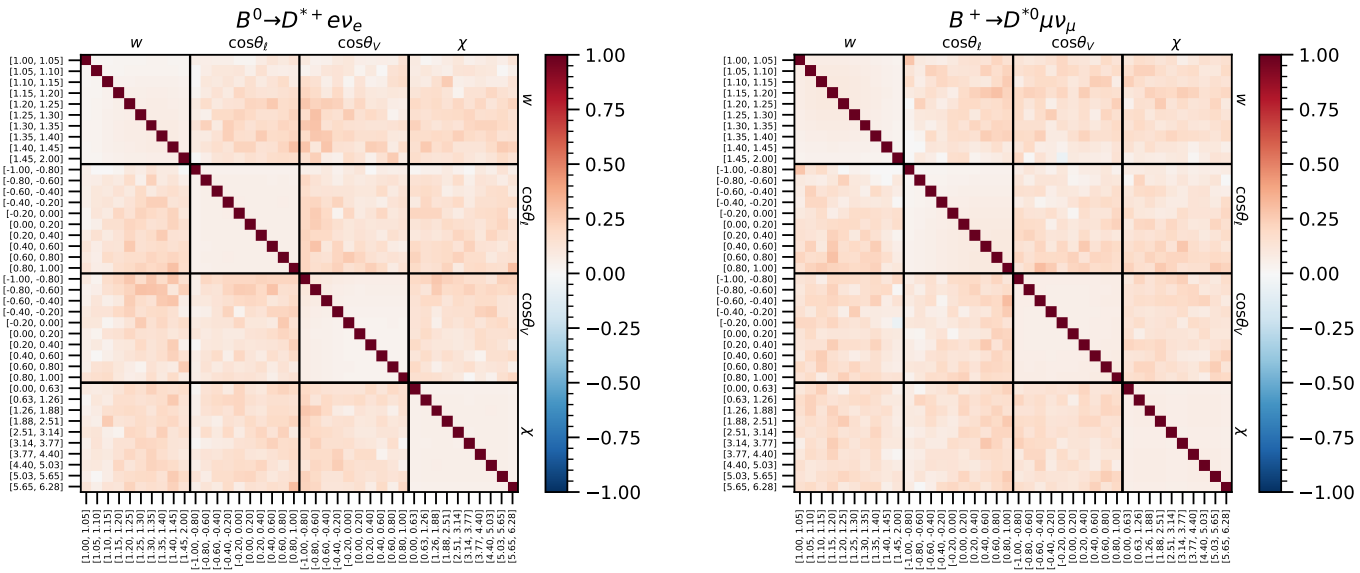


Figure 5.3: Correlation matrices for two reconstruction modes. Figure from [299].

However, this method assumes that the efficiencies of all involved algorithms are equal when applied to recorded data and MC simulation. Unfortunately, this is not always the case. As a simple example, [300] shows that π^0 candidates are less frequently reconstructed in data than in MC. The direction of this deviation is also typical: Because the reconstruction algorithms are usually developed and tuned based on their performance on MC, they are often less efficient on real data.

To account for these differences, we apply *efficiency corrections* to the MC dataset. Several well-established corrections are introduced in Chapter 7. From experience, these corrections sufficiently cover all reconstruction steps on the signal side.

This leaves the efficiency on the tag side, that is, of the FEI. Correcting the efficiency of the FEI is referred to as *tagging calibration*. Compared to the signal side, the number of considered decay channels is orders of magnitude larger. Moreover, the use of multivariate techniques trained on a multitude of variables at different stages of the reconstruction chain leads to a vast number of possible sources of bias. This makes the calibration of the FEI significantly more challenging.

Chapter 6 breaks down the calibration approach presented in this thesis.

5.1.3.2 Migration

We also need to account for migration between the bins of the measured rates. Two options are possible:

1. We apply the migration matrix $\mathcal{M}$ from Equation (5.1.1) to the theoretical prediction of the bin contents ($\mathbf{N}^{(v)}_{\text{pred}}$). The result can then be compared to the measured bin contents ($\mathbf{N}^{(v)}_{\text{meas}}$) as

$$\chi^2 = (\mathbf{N}^{(v)}_{\text{meas}} - \mathcal{M}\mathbf{N}^{(v)}_{\text{pred}})^\intercal \Sigma^{-1} (\mathbf{N}^{(v)}_{\text{meas}} - \mathcal{M}\mathbf{N}^{(v)}_{\text{pred}}) + \chi^2_{\text{NP}}, \tag{5.1.3}$$

where Σ is the covariance matrix of $\mathbf{N}^{(v)}_{\text{meas}}$ and χ^2_{NP} constrains additional nuisance parameters. Minimizing χ^2 as a function of $\mathcal{F}(1)|V_{\text{cb}}|$ and the form factor parameters leads to the optimal values. In our prepared fit for $|V_{\text{cb}}|$, we use a variant of Equation (5.1.3) that uses the information from all four marginalized kinematic distributions (avoiding using the

normalization information multiple times). Several toy studies have validated our statistical treatment.

2. We convert N_i^{meas} into differential decay rates Γ_i^{meas} that can be directly compared with Γ_i^{pred} (in a similar way as just outlined). This approach requires to *unfold* the measurement, that is, to "undo" the effect of the migration. Naively, this could be achieved by inverting the migration matrix, though numerical instabilities in the matrix inversion can amplify the statistical uncertainties.

The first method is conceptually simpler and numerically stable. Hence, this method will be used for the $|V_{cb}|$ value presented in the analysis. However, experiment-independent physical quantities are required to combine the results with other measurements. Consequently, the unfolding of the decay rates will also be part of the analysis. Generally, it is important to preserve or publish all quantities required for the fits to allow for future combined fits with other measurements or analyses with different physical models (be it for form factors or models of new physics).

5.1.4 Unfolded spectra

Unfolding techniques are subject to a bias-variance tradeoff, where bias usually means bias towards MC quantities. The already mentioned technique of inverting the migration matrix is an unbiased estimator. However, small singular values of the matrix can lead to considerable variance.[2] Conversely, unfolding bin-by-bin by using the ratio $\#\{v_{\text{true}} \in b_i^{(v)}\}/\#\{v_{\text{reco}} \in b_i^{(v)}\}$ as correction factor has a small variance but significant bias.

A middle way between these extremes are *regularization* techniques. For example, *SVD unfolding* [301] effectively dampens small singular values of the migration matrix, which reduces the variance while introducing relatively little bias.[3] This method has been used in the previous analysis and is also planned for our work. A comparison of the three methods mentioned can be found in [302, 303].

Figure 5.4 shows the conversion of the fitted signal yields to unfolded decay rates on Asimov data. Figure 5.4a demonstrates that the fitted signal yields perfectly agree with the signal yields in MC, validating the background subtraction procedure. In Figure 5.4b, the event yields of Figure 5.4a have been unfolded and are compared with the signal yields in MC binned in v_{true}. The agreement between both sets of histograms validates the unfolding procedure. Finally, the reconstruction efficiencies can be used to convert the event counts into the differential decay rates (Figure 5.4c). As expected, the decay rates of all four subsamples agree with the theoretical expectation used in the generation of the MC dataset.[4]

We can also see that the results for charged B mesons have significantly smaller uncertainties in the first w bin because of the significantly larger reconstruction efficiency at low w.

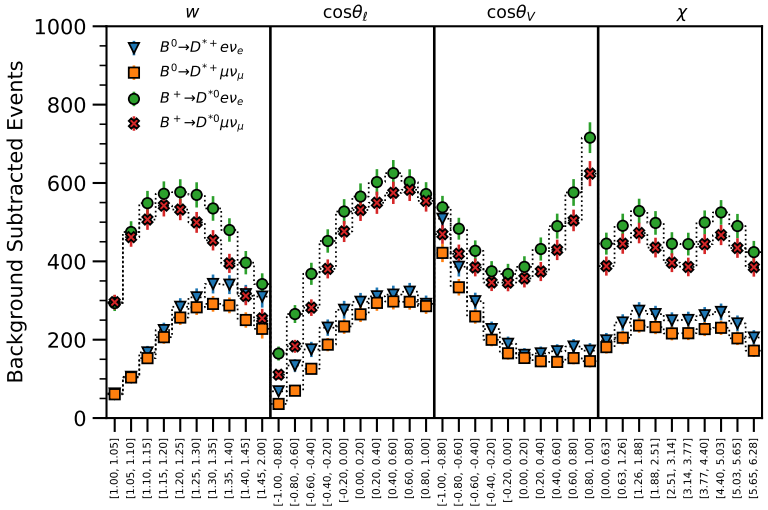
5.1.5 Sideband and resolution studies

Several analyses have previously observed discrepancies in the m_{miss}^2 resolution of hadronically tagged semileptonic decays [83, 304, 305].

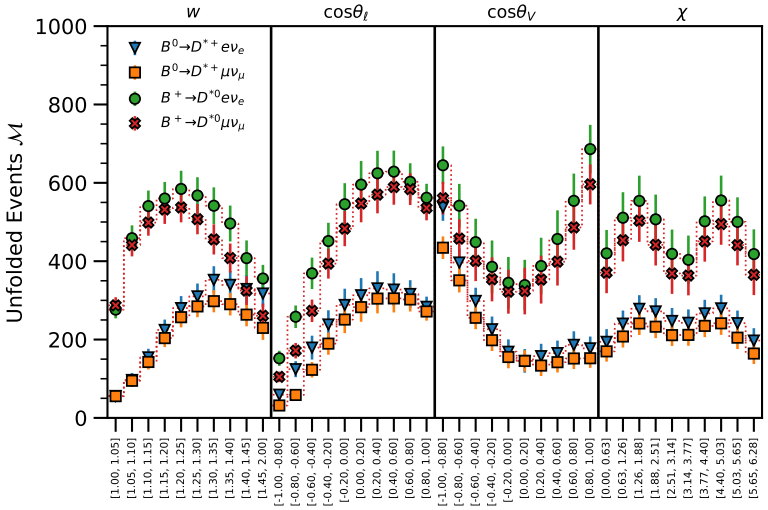
[2] Nonetheless, this is not a deficiency of the method: the matrix inversion method has minimal variance among all unbiased estimators.

[3] SVD unfolding can also be introduced by rephrasing the matrix inversion problem $A\mathbf{w} = \mathbf{b}$ as a minimization problem of $(A\mathbf{w} - \mathbf{b})^\intercal (A\mathbf{w} - \mathbf{b})$ and introducing an additional term that rewards small bin-by-bin variations of ratios to the MC expectation.

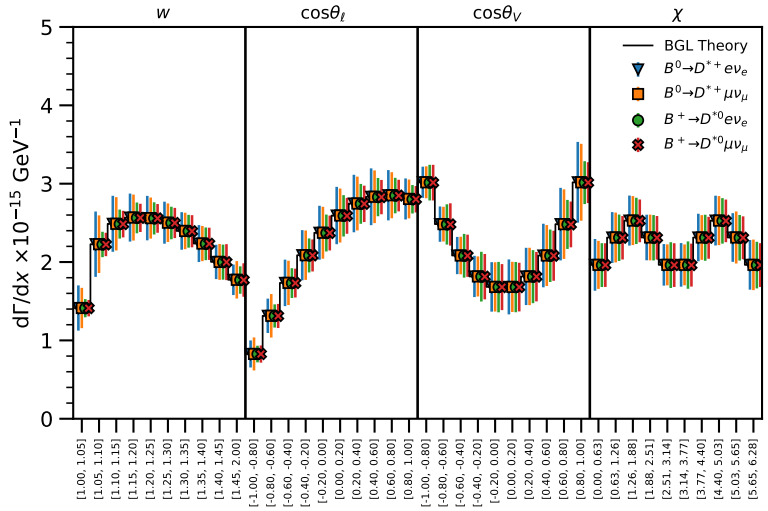
[4] Technically, a different form factor model is used in MC generation, and the form factors are updated thereafter, see Section 7.5.



(a) Fitted signal yields. The dotted black lines of Figure 5.4a show the MC expectation based on v_{reco}.



(b) Unfolded signal yields. The dotted red lines Figure 5.4b show the MC expectation based on v_{true}.



(c) Differential decay rates. The black line shows the theoretical expectation using the BGL form factor model.

Figure 5.4: Unfolding and efficiency correction of fitted signal yields on Asimov data. Figure from [299].

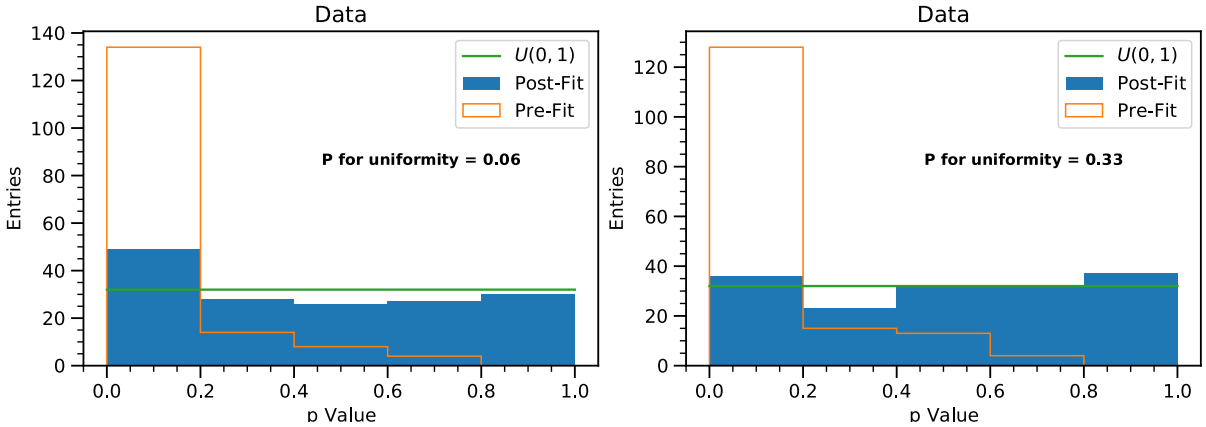
To investigate a possible issue (without unblinding the data of the signal region), we have checked the m^2_{miss} resolution for the $m^{\text{tag}}_{\text{bc}} < 5.27\,\text{GeV}$ sideband. No resolution issue was apparent. Furthermore, we have confirmed good data MC agreement for $m^2_{\text{miss}} > 1\,\text{GeV}^2$.

After these promising results, we have gone one step further and evaluated the p-value distributions for our actual signal region fits on data. In total, 160 fits are performed for the four different data samples and ten bins of each kinematic variable. No information other than the p-values was extracted such that the analysis remains blinded. If the MC model of m^2_{miss} in the signal region is accurate within the assigned uncertainties, a uniform p-value distribution is expected.

The result is shown in Figure 5.5a (blue post-fit histogram): an excess of low p-values is clearly visible. We also perform a statistical test for the uniformity of the distribution by comparing the five bin contents in the p-values with a uniform distribution (green line) using a χ^2 test. A p-value of only 6 % shows that there are indeed statistically significant deviations from the expectation.

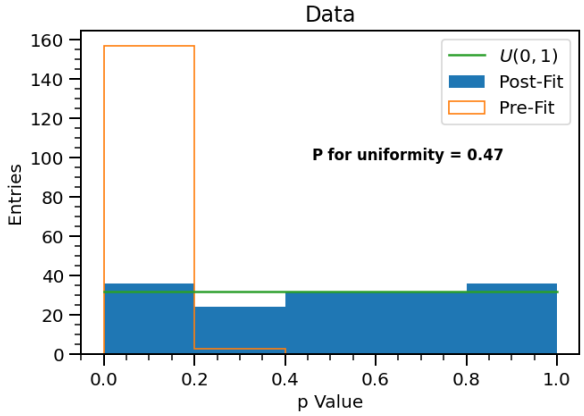
To test if this is indeed a problem with the modeling of the resolution, we (blindly) apply the resolution correction that has been developed in [83] based on Laplace distributions. Indeed, Figure 5.5b shows a good p-value for the uniformity test in this case.

Figure 5.5c confirms that we can further improve the uniformity by applying the tagging calibration presented in this thesis. This finding underlines the point of Section 1.3: The calibration factors do not only impact the overall normalization but can have a significant impact on the measured kinematic distributions. Therefore, the tagging calibration could be one of the puzzle pieces explaining the resolution differences observed in multiple tagged analyses.



(a) Without resolution correction.

(b) With resolution correction.



(c) With resolution correction and tagging calibration.

Figure 5.5: Applying a resolution correction significantly improves the goodness of fit. Shown are the p-values of the 160 fits that are performed for the evaluation of the kinematic distributions (no other quantity but the p-value has been extracted from the fits to avoid unblinding the analysis). The p-value distribution is then compared with a flat distribution with a χ^2 test. The p-value of this test for uniformity is denoted *P for uniformity*. Figures from [299].

Chapter 6

Tagging calibration

The Full Event Interpretation (Section 4.1.1) reconstructs B meson decays by employing multivariate classifiers in multiple stages. Differences between recorded data and the MC simulation can lead to efficiency differences between both datasets. Therefore a data-driven calibration is required (Section 5.1.3.1).

This chapter introduces a calibration using an analysis of inclusive $B \to X\ell\nu_\ell$ decays. Crucial assumptions that ensure the efficacy and transferability of the calibration factors are discussed.

6.1 Calibration factors

All calibration studies share the same approach: reconstructing $\Upsilon(4S) \to B_{tag}B_{sig}$, where B_{tag} is reconstructed by the FEI. If the reconstruction efficiency of the B_{sig} is well-understood and adequately modeled in MC, the difference in event yields between MC and data can be attributed to the FEI. The calibration factors are then calculated as

$$\epsilon = \frac{N^{\text{Data}}}{N^{\text{MC}}}, \tag{6.1.1}$$

where N denotes the respective event yields. By applying ϵ as event weights to the MC, the efficiency difference towards data is corrected.

It is crucial to exclude any efficiency differences and model dependencies on the signal side from the calibration factors. If, as in this thesis, $B_{sig} \to X\ell\nu_\ell$ is used for calibration, and the branching ratio of $X\ell\nu_\ell$ is assumed to be 1 % higher than the (unknown) real value, then ϵ will be 1 % too low. In the case of branching ratios, this can be readily handled by assigning appropriate multiplicative uncertainties.

When reconstructing any decay on the signal side, incorrect reconstructions occur, and the dataset contains a fraction of decays other than the decay of interest. These *background decays* are generally more difficult to describe than the *signal decays*. For this reason, background subtraction techniques are employed, usually involving fits to kinematic distributions. The calibration factor is then defined as

$$\epsilon^{\text{sig}} = \frac{N_{\text{sig}}^{\text{Data}}}{N_{\text{sig}}^{\text{MC}}}, \tag{6.1.2}$$

where $N_{\text{sig}}^{\text{Data}}$ is measured with background subtraction and $N_{\text{sig}}^{\text{MC}}$ can be directly calculated using the generator-level variables on MC.

Furthermore, the calibration factors depend on additional variables, such as the output classifier $\mathcal{P}_{\text{FEI}}$ and the FEI reconstruction mode C_{FEI}. If the distributions of $\mathcal{P}_{\text{FEI}}$ and C_{FEI} are significantly different between the B_{sig} decay used to determine the calibration factors and the dataset they

are applied to, then separate calibration factors must be extracted. For this, data and MC are projected into bins based on these variables, and Equation (6.1.2) is applied for each of them. Because of the required granular calibration, using a decay mode with relatively large branching fractions is necessary. The calibration presented in this note uses $B \to X\ell\nu_\ell$ decays with a branching ratio of around 11 %.

For the previous hadronically tagged $|V_{cb}|$ analysis using the Belle dataset, the FR was used instead of the FEI (see Section 5.1). As both algorithms are conceptually similar, the same calibration issue arises for the FR. The calibration [306] was performed using inclusive $B_{sig} \to X\ell\nu_\ell$. The lepton momentum in the B rest frame, $|\vec{p}_\ell^{\,*}|$, was used as the fitting variable.

The same calibration approach has also been used for the hadronic FEI applied to Belle II data [307]. However, only three bins in $\mathcal{P}_{FEI}$ are chosen because of the still small dataset, and no separation of the different reconstruction modes other than by B flavor is possible.

The hadronic FEI was first calibrated for Belle data in [308], which also used the $B \to X\ell\nu_\ell$ decay on the signal side. The missing mass m^2_{miss} was used as a fit variable. However, separate calibration factors were only calculated for the different decay channels, not for $\mathcal{P}_{FEI}$. Based on the results of this thesis, it is to be expected that this simplification has resulted in a significant bias.[1]

This thesis presents the most complete calibration of the hadronic FEI on Belle data to date. Besides measuring separate calibration factors for both C_{FEI} and $\mathcal{P}_{FEI}$ bins, numerous side-studies probe core assumptions that have not been tested before.

6.2 Calibration assumptions

This section formalizes the requirements to determine and apply the calibration factors that have been introduced in Section 6.1.

Assuming that

i) The reconstruction efficiency (on both real data and simulation) factorizes between signal and tag side, i.e., $\epsilon = \epsilon_{FEI}\epsilon_{sig}$,

ii) The efficiency of the signal side reconstruction is the same between data and MC, i.e., $\epsilon_{sig}^{MC} = \epsilon_{sig}^{Data}$,

iii) The branching fraction of the signal side decay $B_{sig} \to \dots$ is correct in MC, i.e., $BR^{gen}(B_{sig} \to \dots) = BR^{real}(B_{sig} \to \dots)$,

the required FEI calibration factor can be calculated as

$$\begin{aligned}\epsilon_{cal} &:= \frac{N^{\text{Data reco}}_{B_{tag}\to\dots}}{N^{\text{MC reco}}_{B_{tag}\to\dots}} = \frac{\epsilon^{Data}_{FEI} BR^{real}(B \to \text{had.}) N^{real}_{BB}}{\epsilon^{MC}_{FEI} BR^{gen}(B \to \text{had.}) N^{gen}_{BB}} \overset{\text{ii}}{=} \frac{\epsilon^{Data}_{FEI}\epsilon^{Data}_{sig} BR^{real}(B \to \text{had.}) N^{real}_{BB}}{\epsilon^{MC}_{FEI}\epsilon^{MC}_{sig} BR^{gen}(B \to \text{had.}) N^{gen}_{BB}} = \\ &\overset{\text{i}}{=} \frac{\epsilon^{Data} BR^{real}(B \to \text{had.}) N^{real}_{BB}}{\epsilon^{MC} BR^{gen}(B \to \text{had.}) N^{gen}_{BB}} \overset{\text{iii}}{=} \frac{\epsilon^{Data} BR^{real}(B \to \text{had.}) BR^{real}(B_{sig} \to \dots) N^{real}_{BB}}{\epsilon^{MC} BR^{gen}(B \to \text{had.}) BR^{gen}(B_{sig} \to \dots) N^{gen}_{BB}} = \\ &= \frac{N^{\text{Data reco}}}{N^{\text{MC reco}}},\end{aligned} \tag{6.2.1}$$

[1] This can, for example, be seen in Table 10.1 that compares the weighted average of calibration factors for different $\mathcal{P}_{FEI}$ and C_{FEI} distributions. Based on the comparably small variation between decay channels in Figure 10.1a, the observed differences are mostly due to the different $\mathcal{P}_{FEI}$ distributions (shown in Figure 6.1).

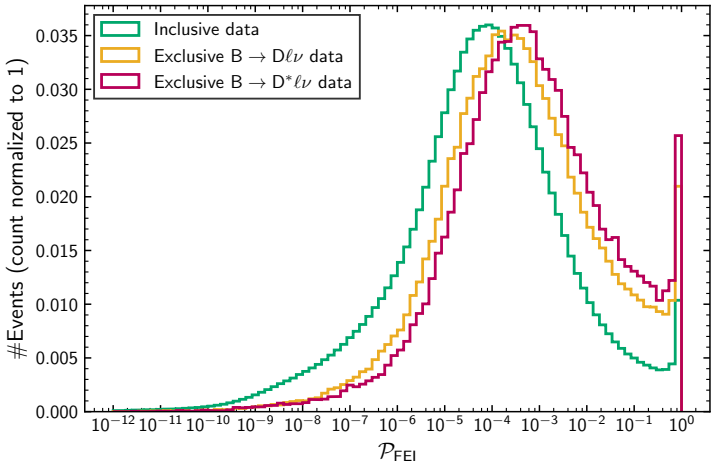


Figure 6.1: Distribution of the FEI classifier output in the inclusive dataset and in exclusive datasets. MC distributions and details for high $\mathcal{P}_{\text{FEI}}$ are shown in Figure B.1.

where N_{BB} is the number of B meson pairs produced, $N^{X\text{ reco}}_{\text{B}_{\text{tag}}\to\dots}$ ($X = \text{Data}, \text{MC}$) is the number of reconstructed tag decays if no signal side is reconstructed, and $N^{X\text{ reco}}$ is the number of background subtracted reconstructed $\Upsilon(4\text{S}) \to \text{B}_{\text{tag}}(\to \text{had.})\text{B}_{\text{sig}}(\to \text{signal})$ events. In particular, while the calibration factor is usually referred to as *efficiency correction*, the branching ratios and N_{BB} are implicitly calibrated as well. This also means that no systematic uncertainty on N_{BB} needs to be added in the $|V_{cb}|$ analysis (this was done incorrectly in the past).

Of course, condition iii can never be fulfilled strictly. In practice, we rather assume that $\text{BR}^{\text{gen}}(\text{B}_{\text{sig}} \to \dots)$ is equal to $\text{BR}^{\text{real}}(\text{B}_{\text{sig}} \to \dots)$ within an uncertainty of $\Delta\text{BR}^{\text{gen}}(\text{B}_{\text{sig}} \to \dots)$ and thus introduce a multiplicative uncertainty of $1 + \Delta\text{BR}^{\text{gen}}(\text{B}_{\text{sig}} \to \dots)/\text{BR}^{\text{gen}}(\text{B}_{\text{sig}} \to \dots)$ in Equation (6.2.1). In fact, this will be the largest source of systematic uncertainty on ϵ_{cal}.

Similarly, condition ii might only be fulfilled approximately within uncertainties. For $\text{B} \to X\ell\nu_\ell$, the reconstruction efficiency on the signal side ϵ_{sig} comes down to the reconstruction of the charged lepton. Any discrepancy between $\epsilon^{\text{MC}}_{\text{sig}}$ and $\epsilon^{\text{Data}}_{\text{sig}}$ should be corrected by the PID corrections and its uncertainties (see Section 7.4).

As mentioned, the calibration factor can depend on the reconstructed tag decay channel and on the classifier outputs of the FEI. Because the number of reconstructed tags in a certain decay channel and in particular of the classifier output depends on the signal side reconstruction (Figures 6.1 and 6.2), a series of calibration factors needs to be extracted for these different subsets. This weakens condition i, because it now only needs to hold *separately* for each of the subsamples that are used to calculate the different calibration factors. To test i, we perform several side-studies that take close look at how the signal side might further influence the FEI performance (see for example Appendix B.2).

6.3 Calibration factors and the $|V_{cb}|$ measurement

$\text{B} \to \text{D}^*\ell\nu_\ell$ decays are one of the major components of the inclusive $\text{B} \to X\ell\nu_\ell$ channel used for tagging calibration. Therefore, this section addresses the treatment of possible correlations between the calibration factors and the $\text{B} \to \text{D}^*\ell\nu_\ell$ measurement.

Assuming the factorization of signal and tag efficiencies, $\text{BR}^{\text{meas}}(\text{B} \to \text{D}^*\ell\nu_\ell)$ is measured as

$$\text{BR}^{\text{meas}}(\text{B} \to \text{D}^*\ell\nu_\ell) = \frac{N^{\text{Data}}_{\text{B}\to\text{D}^*\ell\nu_\ell}}{2\epsilon^{\text{Data}}_{\text{sig}}\epsilon^{\text{Data}}_{\text{tag}}\text{BR}^{\text{real}}(\text{B} \to \text{had.})N^{\text{real}}_{\text{BB}}}, \tag{6.3.1}$$

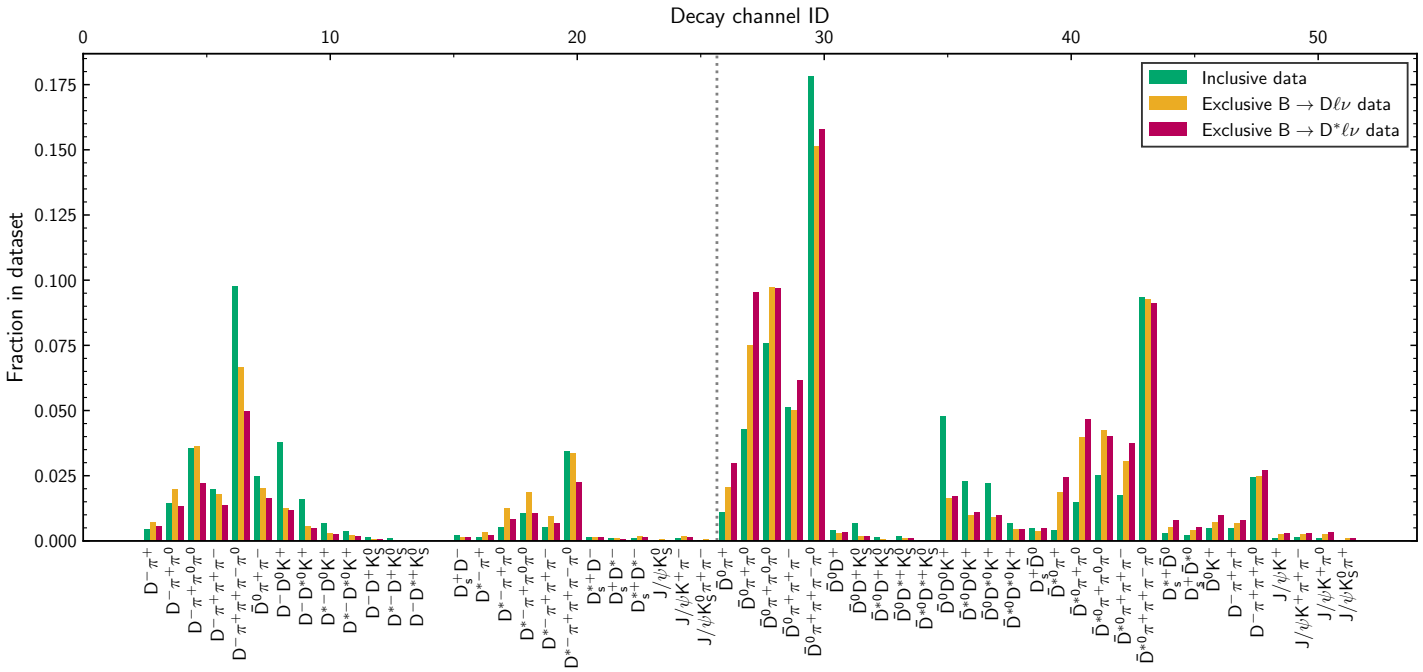


Figure 6.2: Frequency of reconstruction modes in the inclusive $B \to X\ell\nu_\ell$ dataset and in the exclusive datasets. The dashed line separates B^0 modes from B^+ modes. The decay channels are ordered by the decay channel ID, which is also tabulated in Table 9.1.

where $N^{\text{Data}}_{B\to D^*\ell\nu_\ell}$ is the number of reconstructed and background subtracted $\Upsilon(4S) \to (B \to D^*\ell\nu_\ell)(B \to \text{had.})$ events.

From MC we can infer

$$\epsilon^{\text{MC}}_{\text{sig}} = \frac{N^{\text{MC}}_{B\to D^*\ell\nu_\ell}}{2\epsilon^{\text{MC}}_{\text{tag}}\text{BR}^{\text{gen}}(B \to D^*\ell\nu_\ell)\text{BR}^{\text{gen}}(B \to \text{had.})N^{\text{gen}}_{B\bar{B}}}. \tag{6.3.2}$$

Assuming that $\epsilon^{\text{Data}}_{\text{sig}} \overset{!}{=} \epsilon^{\text{MC}}_{\text{sig}}$ and defining the calibration factor as in the previous section, we have

$$\begin{aligned}\text{BR}^{\text{meas}}(B \to D^*\ell\nu_\ell) &= \frac{1}{\epsilon_{\text{cal}}}\frac{N^{\text{Data}}_{B\to D^*\ell\nu_\ell}}{N^{\text{MC}}_{B\to D^*\ell\nu_\ell}}\text{BR}^{\text{gen}}(B \to D^*\ell\nu_\ell) = \\ &= \frac{N^{\text{MC}}_{B\to X\ell\nu_\ell}}{N^{\text{Data}}_{B\to X\ell\nu_\ell}}\frac{N^{\text{Data}}_{B\to D^*\ell\nu_\ell}}{N^{\text{MC}}_{B\to D^*\ell\nu_\ell}}\text{BR}^{\text{gen}}(B \to D^*\ell\nu_\ell),\end{aligned} \tag{6.3.3}$$

where $N^{\text{Data}}_{B\to X\ell\nu_\ell}$ is the number of reconstructed and background subtracted $\Upsilon(4S) \to (B \to X\ell\nu_\ell)(B \to \text{had.})$ events.

As mentioned, there exists statistical overlap between $N^{\text{Data/MC}}_{B\to X\ell\nu_\ell}$ and $N^{\text{Data/MC}}_{B\to D^*\ell\nu_\ell}$. However, only around 0.1 % of the $B \to X\ell\nu_\ell$ events are considered in the $|V_{cb}|$ analysis. This is due to the significantly larger reconstruction efficiency of $B \to X\ell\nu_\ell$ that only reconstructs the lepton. Furthermore, the $X\ell\nu_\ell$ channel also includes $B \to D\ell\nu_\ell$ and $B \to D^{**}\ell\nu_\ell$ decays, both of which have significant branching ratios. Therefore, the statistical overlap is neglected.

The branching fraction of $\text{BR}(B \to D^*\ell\nu_\ell)$ in $N^{\text{MC}}_{B\to D^*\ell\nu_\ell}$ is set up to cancel with $\text{BR}^{\text{gen}}(B_{\text{sig}} \to \dots)$ (which therefore does not need any uncertainty). Thus, there is no shared systematic uncertainty with $N^{\text{MC}}_{B\to X\ell\nu_\ell}$ or $N^{\text{Data}}_{B\to X\ell\nu_\ell}$ (which depends only slightly on the $\text{BR}(B \to D^*\ell\nu_\ell)$ via the signal fit template).

Chapter 7

The Monte Carlo Dataset

Almost all analysis strategies in experimental particle physics rely on the comparison of experimental data to simulated Monte Carlo (MC) data. MC data needs to be specifically produced to match the experimental conditions of the recorded data. Therefore, MC specific to the Belle detector configurations and beam conditions is used in this analysis. However, several years have passed since the final MC production of the Belle experiment, and our general understanding of the relevant physics has improved significantly. Reproducing the required amount of MC data is unfeasible because of the immense computation power requirements. Therefore, we instead apply a series of corrections to the last MC production of the Belle experiment, including the production of several custom samples of specific decay modes.

7.1 Monte Carlo Generation

The MC dataset is generated in two steps:

Event Generation For the generation of MC, physical events (collisions, particle decays, and kinematic properties) are sampled according to their branching ratios and models of decay kinematics. Belle uses the EvtGen package [309], a generator that specializes in the modeling of the physical processes involved in B meson decays. PYTHIA [310] is used to model the fragmentation and hadronization of quarks and PHOTOS [311] describes radiative photons from final state radiation. The information at this stage is called *truth* or *generator level* information.

Detector simulation In the second step, the interactions of the generated particles with the Belle detector are simulated. This step makes MC production computationally expensive and time-consuming because the material interactions and the feedback given from the various sub-detectors need to be described accurately. Detector simulation in Belle is performed by the Geant 3 package [312, 313].

After detector simulation, MC data has the same form as actual recorded data and can be analyzed with the exact same tools, allowing side-by-side comparisons between MC and data at every step of the analysis.

7.2 Monte Carlo Samples

The MC dataset that is being used in the tagging calibration is made up of several different samples:

Generic b → c events This sample, produced by the Belle collaboration, is the most important sample for the $|V_{cb}|$ analysis and the tagging calibration. It contains $B\overline{B}$ events where both B mesons decay via a b → c transition, such as $B \to D^{(*)}\ell\nu_\ell$ and $B \to D^{**}\ell\nu_\ell$ decays.

The sample is split by B flavor and denoted *Generic Charged* (B^+ pairs) and *Generic Mixed* (B^0 pairs).

Continuum events As already mentioned in Section 3.1, continuum events are non-$B\overline{B}$ background, i.e., e^-e^+ collisions that do not result in a $\Upsilon(4S)$ particle but in $q\bar{q}$ pairs with $q = c, u, d, s$. They are covered by two samples of Belle MC: the *Charm* sample ($q = c$) and the *UDS* sample ($q = u, d, s$).

b → uℓν events $B\overline{B}$ decays involving a b → u transition are considerably less frequent than b → c transitions. However, they become important in the tail of the $|\vec{p}_\ell^{\,*}|$ distribution. We use the samples produced by the Belle collaboration.

$B \to D^{**}\ell\nu_\ell$ events Semileptonic events with orbitally excited charm mesons are an important part of the $B \to X\ell\nu_\ell$ dataset used in the calibration. They also form an important background in the $|V_{cb}|$ analysis. The four different resonances are collectively referred to as D^{**} mesons and listed in Table 7.1. The different decays are depicted in Figure 7.1.

Compared to $B \to D^{(*)}\ell\nu_\ell$ decays, $B \to D^{**}\ell\nu_\ell$ decays are considerably less understood, and the models to describe them are still evolving. Since the generation of the Belle MC, the values of masses and decay widths have changed considerably (see [83, Table 2.2]). Furthermore, Belle MC does not consider several decay modes of the D_1, which have since been measured. While form factors can be corrected by applying weights, updating these key properties requires the reproduction of the MC. This analysis uses the following privately produced samples:

- $B \to D_0^*\ell\nu_\ell$,
- $B \to D_1\ell\nu_\ell$,
- $B \to D_1'\ell\nu_\ell$,
- $B \to D_2^*\ell\nu_\ell$,
- $B \to D_1(\to D\pi\pi)\ell\nu_\ell$.

Gap events While the $B \to D^{(*)(*)}\ell\nu_\ell$ decays make up around 90 % of the $B \to X\ell\nu_\ell$ branching ratio, the remaining 10 % are less well understood, and the description Belle MC is outdated. We consider two different sets of decays to fill this *gap*. Because the decays involve D^{**} mesons or are either not considered or incorrectly modeled in Belle MC, the following privately produced samples are being used:

- $B \to D\eta\ell\nu_\ell$,
- $B \to D\pi\pi\ell\nu_\ell$,
- $B \to D^*\eta\ell\nu_\ell$,
- $B \to D^*\pi\pi\ell\nu_\ell$,
- $B \to D_0^*(\to D\eta)\ell\nu_\ell$,
- $B \to D_0^*(\to D^*\pi\pi)\ell\nu_\ell$,
- $B \to D_0^*(\to D\pi\pi)\ell\nu_\ell$,
- $B \to D_1'(\to D^*\eta)\ell\nu_\ell$,
- $B \to D_1'(\to D^*\pi\pi)\ell\nu_\ell$,
- $B \to D_1'(\to D\pi\pi)\ell\nu_\ell$.

	$s_l^{\pi_l}$	J^P	m [GeV]	Γ [GeV]
D_0^*	$\frac{1}{2}^+$	0^+	2.3	0.27 / 0.22
$\mathrm{D}_1' = D_1^*$	$\frac{1}{2}^+$	1^+	2.4	0.38
D_1	$\frac{3}{2}^+$	1^+	2.4	0.03
D_2^*	$\frac{3}{2}^+$	2^+	2.5	0.05

Table 7.1: Quantum numbers of D, D*, and D** mesons [314]. The displayed masses and decay widths are the values used in the privately produced MC samples. $s_l^{\pi_l}$ denotes the spin and parity of the light degrees of freedom. If two values are given, they refer to the values of neutral and charged mesons. D_1' and D_1^* are alternative notations for the same meson.

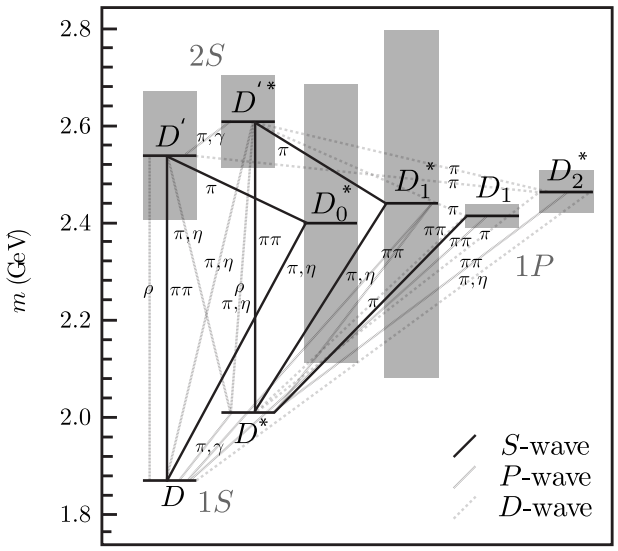


Figure 7.1: Orbitally excited charm mesons and their strong decays. Gray bands indicate the decay widths, showing the clear separation between the broad states D_0^*, D_1' and the narrow states D_1 and D_2^*. Lines indicate decays between the mesons. The line style and opacity distinguishes between the orbital momenta of the partial waves. The particles next to the lines are additional decay products. Figure from [314].

In all samples except the continuum sample, one of the B mesons is decaying generically (b $\to$ c transition), whereas the other is forced to decay in the specified channel(s).

In addition to the samples described here, the $|V_{\mathrm{cb}}|$ analysis includes the rare MC sample (containing B decays of small branching ratios, such as $\mathrm{B}^+ \to \ell^+ \nu_\ell \gamma$). This sample is negligible in the tagging calibration. Furthermore, the tuples used in the $|V_{\mathrm{cb}}|$ analysis (originally produced for the analysis presented in [83, 84]) include updates to semitauonic decays. The respective branching ratios are generally smaller than for semileptonic decays with light leptons. Tau leptons are also not included in the signal definition of the tagging calibration (see Section 9.2.4). For this reason, they are not corrected in the tagging calibration (the description in the generic Belle MC is used).

The amount of MC events is often expressed in the unit of *streams*, where one stream is intended to correspond to the integrated luminosity of the recorded data. Since this correspondence is only approximately fulfilled, it is corrected by assigning appropriate event weights. We also update the ratio of charged and neutral B meson pairs to the current world average.

Because the tagging calibration is systematically limited, only one stream of the Belle MC samples is used. The $|V_{\mathrm{cb}}|$ analysis considers a much tighter signal selection and thus uses all

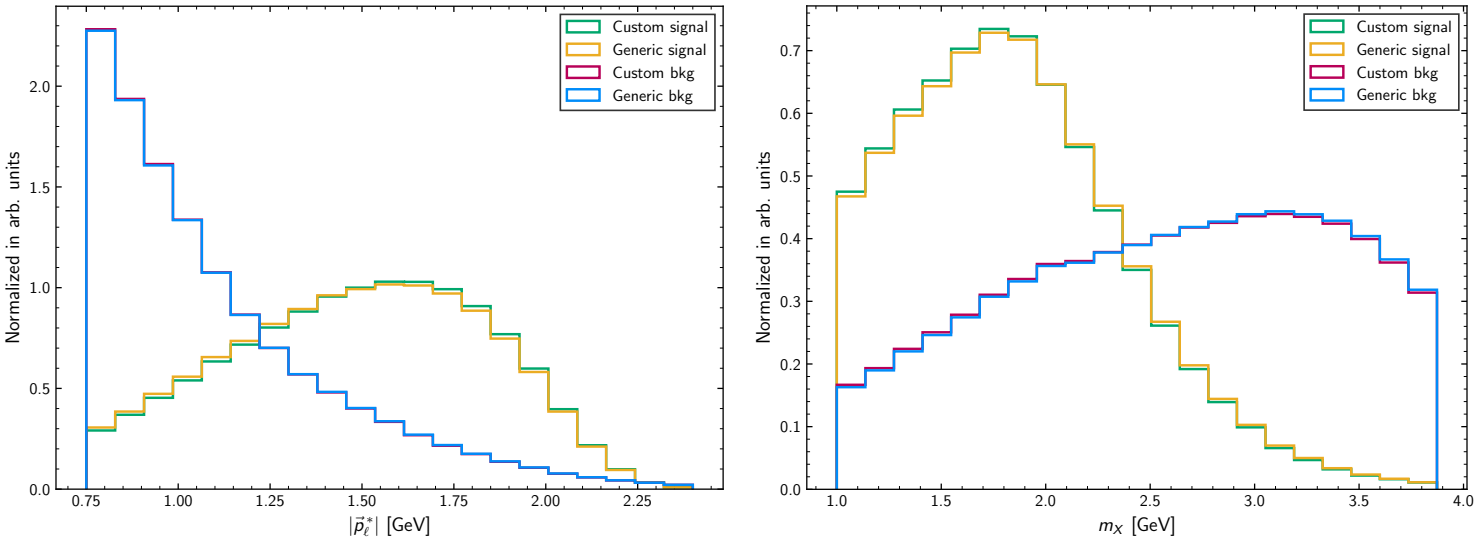


Figure 7.2: Distributions of the signal and background components for generic MC and the updated MC. The observable m_X is introduced in Section 9.2.5, the signal and background components in Section 9.2.4. *Custom* refers to the updated MC.

available streams (10 for generic and 6 for continuum events).

All samples are combined into one dataset that is subsequently compared with data. The $B \to D^{**}\ell\nu_\ell$ and gap samples are intended to *replace* the corresponding parts of generic $b \to c$ MC, which must thus be removed. In particular, this includes all events featuring generic $B \to D^{**}\ell\nu_\ell$, $B \to D^{(*)}\pi\ell\nu_\ell$ (Goity Roberts processes), or $B \to D^{(*)}(2S)\ell\nu_\ell$ decays on signal or tag side.

As can be seen in Figure 7.2, the effect of the MC corrections on the overall shape of the kinematic distributions is relatively small (though more pronounced in the $B \to D^{**}\ell\nu_\ell$ distributions, see Figure 7.3). Instead, the more significant impact of our corrections is on the normalizations. For example, the difference in the overall normalizations of the dataset when comparing the two different gap models (of equal total branching ratio) is of order 3 %.

Several additional technical details regarding the samples are mentioned in Appendix D.1.1.

7.3 Branching ratios

This section discusses updates of the branching ratios of decays in the MC dataset. Generally, corrections are only applied based on the signal side decays; any discrepancy on the tag side is to be included in the calibration factors (therefore, it is crucial that any analysis that uses the calibration factors does not apply any other tag side weights either).

The most important update is that of the inclusive branching ratio to [99, Equation (206)]

$$\mathrm{BR}(B \to X_c\ell\nu_\ell) = (10.65 \pm 0.16)\%. \tag{7.3.1}$$

Note that this update corresponds to an update of the non-semileptonic decay ratios for consistency:

$$w_{\mathrm{NSL}} = \frac{1 - \mathrm{BR}(B \to X_c\ell\nu_\ell)|_{\mathrm{updated}}}{1 - \mathrm{BR}(B \to X_c\ell\nu_\ell)|_{\mathrm{current}}} = \begin{cases} 1.010 & \text{charged} \\ 1.007 & \text{mixed.} \end{cases} \tag{7.3.2}$$

However, as we only reweight signal side decays and only very few hadronic decays are incorrectly reconstructed to contribute there, this weight is neglected.

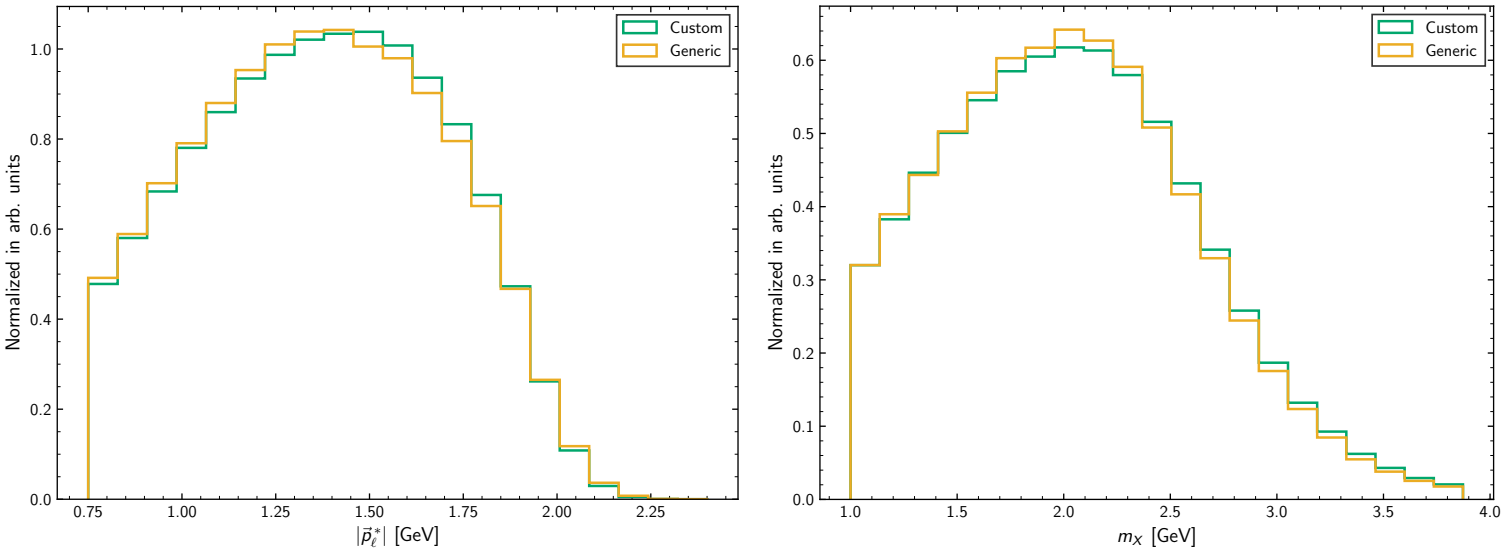


Figure 7.3: D^{**} decays in generic MC and in the dedicated samples (denoted *Custom*, already includes the form factor weights introduced in Section 7.5). The observable m_X is introduced in Section 9.2.5, the signal and background components in Section 9.2.4.

All updated branching ratios are listed in Table 7.2. They will be introduced in detail below.

7.3.1 Isospin averages and $B \to D^{(*)}\ell\nu_\ell$

To a good approximation, the strong force is invariant under SU(2) isospin transformations [315, p. 377][316, p. 87]: if two baryons or mesons only differ by an up quark exchanged by a down quark, then their properties associated with the strong force will be approximately equal. This relation applies to B^0 and B^+ mesons as well as $D^{(*)0}$ and $D^{(*)+}$ mesons. Because the $B \to D^{(*)}\ell\nu_\ell$ decay is a $b \to c$ transition (and thus does not involve the u and d quarks at leading order), it is a reasonable assumption that this decay obeys the isospin symmetry, i.e., that the only difference in branching ratios is due to the different lifetimes τ_{B^0} and τ_{B^+} of the B^0 and B^+ meson. The advantage of this assumption is that we can then combine the measurements of $\mathrm{BR}(B^0 \to D^{(*)+}\ell^-\bar{\nu}_\ell)$ and $\mathrm{BR}(B^+ \to \overline{D}^{(*)0}\ell^+\nu_\ell)$ using the well known values of τ_{B^0} and τ_{B^+}, thereby reducing the uncertainty on both branching ratios.

Formally, we make the assumption that

$$\begin{aligned}\mathrm{BR}(B^0 \to D^{(*)-}\ell^+\nu_\ell) &= \Gamma(B \to D^{(*)}\ell\nu_\ell)\tau_{B^0}\\ \mathrm{BR}(B^+ \to \overline{D}^{(*)0}\ell^+\nu_\ell) &= \Gamma(B \to D^{(*)}\ell\nu_\ell)\tau_{B^+},\end{aligned} \tag{7.3.3}$$

therefore,

$$\mathrm{BR}(B^0 \to D^{(*)-}\ell^+\nu_\ell) \cdot \frac{\tau_{B^+}}{\tau_{B^0}} = \mathrm{BR}(B^+ \to \overline{D}^{(*)0}\ell^+\nu_\ell), \tag{7.3.4}$$

where [99, Equation (40)]

$$\frac{\tau_{B^+}}{\tau_{B^0}} = 1.076 \pm 0.004. \tag{7.3.5}$$

Decay	B^0	B^+
Golden modes		
$B \to D^* \ell \nu_\ell$	5.11 ± 0.11	5.50 ± 0.11
$B \to D \ell \nu_\ell$	2.24 ± 0.07	2.41 ± 0.07
D^{**} modes		
$B \to D_1 \ell^+ \nu_\ell$	0.62 ± 0.10	0.66 ± 0.11
$B \to D_0^* \ell^+ \nu_\ell$	0.39 ± 0.07	0.42 ± 0.08
$B \to D_1' \ell^+ \nu_\ell$	0.39 ± 0.08	0.42 ± 0.09
$B \to D_2^* \ell^+ \nu_\ell$	0.29 ± 0.03	0.29 ± 0.03
Non-resonant gap		
$B \to D \pi\pi \ell^+ \nu_\ell$	0.06 ± 0.08	0.06 ± 0.09
$B \to D^* \pi\pi \ell^+ \nu_\ell$	0.20 ± 0.10	0.22 ± 0.10
$B \to D \eta \ell \nu_\ell$	0.40 ± 0.40	0.40 ± 0.40
$B \to D^* \eta \ell \nu_\ell$	0.40 ± 0.40	0.40 ± 0.40
Resonant gap		
$B \to D_0^*(\to D\pi\pi)\ell^+ \nu_\ell$	0.03 ± 0.03	0.03 ± 0.03
$B \to D_0^*(\to D^*\pi\pi)\ell^+ \nu_\ell$	0.10 ± 0.10	0.11 ± 0.11
$B \to D_1'(\to D\pi\pi)\ell^+ \nu_\ell$	0.03 ± 0.03	0.03 ± 0.03
$B \to D_1'(\to D^*\pi\pi)\ell^+ \nu_\ell$	0.10 ± 0.10	0.11 ± 0.11
$B \to D_0^*(\to D\eta)\ell^+ \bar{\nu}_\ell$	0.40 ± 0.40	0.40 ± 0.40
$B \to D_1'(\to D^*\eta)\ell^+ \bar{\nu}_\ell$	0.40 ± 0.40	0.40 ± 0.40

Table 7.2: Overview of the B branching ratios used in the analysis. All values are specified in percent. The non-resonant and resonant gap model are alternative models describing the same total branching ratio. In the parallel table in [83], $B \to D_1(\to D\pi\pi)\ell^+\nu_\ell$ is listed separately (rather than being included in $B \to D_1 \ell^+ \nu_\ell$).

With this, we can calculate the isospin average as

$$\begin{aligned} \mathrm{BR}(B^+ \to \overline{D}^{(*)0} \ell^+ \nu_\ell)\Big|_{\text{IS avg}} = \\ w_{B^0} \cdot \mathrm{BR}(B^0 \to D^{(*)-} \ell^+ \nu_\ell) \cdot \frac{\tau_{B^+}}{\tau_{B^0}} + w_{B^+} \cdot \mathrm{BR}(B^+ \to \overline{D}^{(*)0} \ell^+ \nu_\ell) \end{aligned} \tag{7.3.6a}$$

$$\mathrm{BR}(B^0 \to D^{(*)-} \ell^+ \nu_\ell)\Big|_{\text{IS avg}} = \frac{\tau_{B^0}}{\tau_{B^+}} \mathrm{BR}(B^+ \to \overline{D}^{(*)0} \ell^+ \nu_\ell)\Big|_{\text{IS avg}}, \tag{7.3.6b}$$

where $w_{B^0}, w_{B^+} \in \mathbb{R}^+$ are weights for the average $(w_{B^0} + w_{B^+})$. We choose inverse-variance weighting [317, Equation (4.3)] to minimize the uncertainty of the average, i.e., set

$$w_{B^0} = \frac{\sigma_0^{-2}}{\sigma_0^{-2} + \sigma_+^{-2}}, \qquad w_{B^+} = \frac{\sigma_+^{-2}}{\sigma_0^{-2} + \sigma_+^{-2}}, \tag{7.3.7}$$

where σ_0 (σ_+) denotes the uncertainty on $\mathrm{BR}(B^0 \to D^{(*)-} \ell^+ \nu_\ell)$ ($\mathrm{BR}(B^+ \to \overline{D}^{(*)0} \ell^+ \nu_\ell)$).

BR	Exp. Avg. [%]	IS Avg. [%]
$\mathrm{BR}(\mathrm{B}^0 \to \mathrm{D}^- \ell^+ \nu_\ell)$	$2.31 \pm 0.18_{\mathrm{stat}} \pm 0.36_{\mathrm{syst}}$	2.24 ± 0.07
$\mathrm{BR}(\mathrm{B}^+ \to \overline{\mathrm{D}}^0 \ell^+ \nu_\ell)$	$2.35 \pm 0.03_{\mathrm{stat}} \pm 0.09_{\mathrm{syst}}$	2.41 ± 0.07
$\mathrm{BR}(\mathrm{B}^0 \to \mathrm{D}^{*-} \ell^+ \nu_\ell)$	$5.06 \pm 0.02_{\mathrm{stat}} \pm 0.12_{\mathrm{syst}}$	5.11 ± 0.11
$\mathrm{BR}(\mathrm{B}^+ \to \overline{\mathrm{D}}^{*0} \ell^+ \nu_\ell)$	$5.66 \pm 0.07_{\mathrm{stat}} \pm 0.21_{\mathrm{syst}}$	5.50 ± 0.11

Table 7.3: Isospin averages for $\mathrm{BR}(\mathrm{B} \to \mathrm{D}^{(*)} \ell \nu_\ell)$. Abbreviations: *Exp. Avg.* (experimental average, see text), *IS Avg.* (isospin average).

	Belle	Rwght.		Belle	Rwght.
$\mathrm{D}_1^+ \to \mathrm{D}^{*+}\pi^0$	33.33	19.97	$\mathrm{D}_1^0 \to \mathrm{D}^{*0}\pi^0$	33.33	19.97
$\mathrm{D}_1^+ \to \mathrm{D}^{*0}\pi^+$	66.67	39.94	$\mathrm{D}_1^0 \to \mathrm{D}^{*+}\pi^-$	66.67	39.94
$\mathrm{D}_1^+ \to \mathrm{D}^+\pi^+\pi^-$	—	17.19	$\mathrm{D}_1^0 \to \mathrm{D}^0\pi^+\pi^-$	—	17.19
$\mathrm{D}_1^+ \to \mathrm{D}^+\pi^0\pi^0$	—	11.45	$\mathrm{D}_1^0 \to \mathrm{D}^0\pi^0\pi^0$	—	11.45
$\mathrm{D}_1^+ \to \mathrm{D}^0\pi^+\pi^0$	—	11.45	$\mathrm{D}_1^0 \to \mathrm{D}^+\pi^-\pi^0$	—	11.45
$\mathrm{D}_2^{*+} \to \mathrm{D}^{*+}\pi^0$	10.30	13.12	$\mathrm{D}_2^{*0} \to \mathrm{D}^{*+}\pi^-$	26.35	20.90
$\mathrm{D}_2^{*+} \to \mathrm{D}^{*0}\pi^+$	20.90	26.25	$\mathrm{D}_2^{*0} \to \mathrm{D}^{*0}\pi^0$	10.30	13.12
$\mathrm{D}_2^{*+} \to \mathrm{D}^+\pi^0$	22.90	20.21	$\mathrm{D}_2^{*0} \to \mathrm{D}^+\pi^-$	45.90	40.42
$\mathrm{D}_2^{*+} \to \mathrm{D}^0\pi^+$	45.90	40.42	$\mathrm{D}_2^{*0} \to \mathrm{D}^0\pi^0$	22.90	20.21

Table 7.4: Updates of branching ratio of D^{**} decay modes. All branching ratios are shown in percent. *Rwght.* denotes the updated branching ratio. Table identical in [83].

Table 7.3 shows the resulting isospin averaged branching ratios. The input branching ratios were taken from [99], Table 67, 68, 70 and 71.

7.3.2 $\mathrm{B} \to \mathrm{D}^{**} \ell \nu_\ell$

The branching ratios of the four 1P states of the D meson (referred to as D^{**} mesons) are more challenging to measure. In most cases, only *partial* branching ratios (B branching ratios that consider a specific decay mode of the D^{**}) are available. To obtain the total branching ratio, we use additional measurements of ratios of D^{**} branching ratios. In the following, we apply the procedure described in [318]. Central to the calculations will be the assumption of isospin symmetry (which is assumed in all results presented), in particular of the isospin factors for two and three-body pion modes:

$$f_\pi = \frac{\mathrm{BR}(\mathrm{D}^{**} \to \mathrm{D}^{(*)-}\pi^+)}{\mathrm{BR}(\mathrm{D}^{**} \to D^{(*)}\pi)} = \frac{2}{3}, \tag{7.3.8a}$$

$$f_{\pi\pi} = \frac{\mathrm{BR}(\mathrm{D}^{**} \to \mathrm{D}^{(*)-}\pi^+\pi^-)}{\mathrm{BR}(\mathrm{D}^{**} \to \mathrm{D}^{(*)}\pi\pi)} = \frac{1}{2} \pm \frac{1}{6}. \tag{7.3.8b}$$

7.3.2.1 $B \to D_1 \ell^+ \nu_\ell$

$B \to D_1 \ell^+ \nu_\ell$ decays either via $D^* \pi$ or $D \pi \pi$. We use the well measured partial branching ratio [99, Table 74]

$$\mathrm{BR}(B \to D_1(\to D^{*+}\pi^-)\ell^+\nu_\ell) = (0.281 \pm 0.010_{\mathrm{stat}} \pm 0.015_{\mathrm{syst}})\% \tag{7.3.9}$$

together with the ratio

$$f_{D_1} = \frac{\mathrm{BR}(D_1 \to D^{*-}\pi^+)}{\mathrm{BR}(D_1 \to D^0\pi^+\pi^-)} = 2.32 \pm 0.54, \tag{7.3.10}$$

which can be computed from [319, p. 23], and Equations (7.3.5), (7.3.8a) and (7.3.8b) to obtain[1]

$$\mathrm{BR}(B^0 \to D_1\ell^+\nu_\ell) = (0.62 \pm 0.10)\% \quad \text{and} \quad \mathrm{BR}(B^+ \to D_1\ell^+\nu_\ell) = (0.66 \pm 0.11)\%. \tag{7.3.11}$$

The branching ratios of the D_1 decay modes are shown in Table 7.4.

7.3.2.2 $B \to D_0^* \ell^+ \nu_\ell$

The D_0^* meson is assumed to only decay to $D\pi$, so this case is straightforward. We use the average of Belle and BaBar measurements from [99, Table 74]

$$\mathrm{BR}(B \to D_0^*(\to D^+\pi^-)\ell^+\nu_\ell) = (0.28 \pm 0.03_{\mathrm{stat}} \pm 0.04_{\mathrm{syst}})\% \tag{7.3.12}$$

together with Equations (7.3.5) and (7.3.8a) to obtain

$$\mathrm{BR}(B^0 \to D_0^*\ell^+\nu_\ell) = (0.39 \pm 0.07)\% \quad \text{and} \quad \mathrm{BR}(B^+ \to D_0^*\ell^+\nu_\ell) = (0.42 \pm 0.08)\%. \tag{7.3.13}$$

7.3.2.3 $B \to D_1' \ell^+ \nu_\ell$

D_1' is assumed to only decay to $D^*\pi$, so this case is simple as well: $\mathrm{BR}(B \to D_1'(\to D^{*+}\pi^-)\ell^+\nu_\ell)$ has been measured by DELPHI [320], Belle [321] and BaBar [322], all listed in [99], table 76. Because the Belle measurement is incompatible, we use the variance weighted average of only the other two measurements:

$$\mathrm{BR}(B^+ \to D_1'(\to D^{*+}\pi^-)\ell^+\nu_\ell) = (0.28 \pm 0.06)\%. \tag{7.3.14}$$

Together with Equations (7.3.5) and (7.3.8a) we obtain

$$\mathrm{BR}(B^0 \to D_1'\ell^+\nu_\ell) = (0.39 \pm 0.08)\% \quad \text{and} \quad \mathrm{BR}(B^+ \to D_1'\ell^+\nu_\ell) = (0.42 \pm 0.09)\%. \tag{7.3.15}$$

7.3.2.4 $B \to D_2^* \ell^+ \nu_\ell$

The D_2^* is assumed to only decay to $D^{(*)}\pi$. This is not entirely accurate, because $D_2^* \to D\pi\pi$ contributes to a smaller degree as well, but this process is not included in Belle nor in Belle II MC at the moment (see Section 7.3.3). Therefore, for this thesis, $B \to D_2^*\ell^+\nu_\ell$ refers to $B \to D_2^*(\to D^{(*)}\pi)\ell^+\nu_\ell$. Using

$$\mathrm{BR}(B^- \to D_2^{*0}(\to D^{*+}\pi^-)\ell^-\bar{\nu}_\ell) = (0.077 \pm 0.006_{\mathrm{stat}} \pm 0.004_{\mathrm{syst}})\% \tag{7.3.16}$$

from [99], table 75 and

$$f_{D_2^*} = \frac{\mathrm{BR}(D_2^* \to D^-\pi^+)}{\mathrm{BR}(D_2^* \to D^{*-}\pi^+)} = 1.54 \pm 0.15 \tag{7.3.17}$$

computed from [319, p. 23] and Equations (7.3.5) and (7.3.8a), we obtain

$$\mathrm{BR}(B^0 \to D_2^*\ell^+\nu_\ell) = (0.27 \pm 0.03)\% \quad \text{and} \quad \mathrm{BR}(B^+ \to D_2^*\ell^+\nu_\ell) = (0.29 \pm 0.03)\%. \tag{7.3.18}$$

The branching ratios of the D_2^* decay modes are shown in Table 7.4.

[1] Implementation note: in the calibration framework we follow the nomenclature of [83]. There, $B \to D_1\ell^+\nu_\ell$ refers to the decay with $D_1 \to D^*\pi$. The decay $B \to D_1(\to D\pi\pi)\ell^+\nu_\ell$ is considered separately (and contains the prefix `Gap` in the code though it is not technically part thereof).

7.3.3 Decays not considered in the MC

The decays $\mathrm{D}_\mathrm{s}^{(*)}\mathrm{K}\ell^+\nu_\ell$ have been measured by Belle [323]. They are currently not included in the Belle or Belle II MC. Furthermore, we have already mentioned that the same applies to $\mathrm{B} \to \mathrm{D}_2^*(\to \mathrm{D}\pi\pi)\ell^+\nu_\ell$. However, because the sum of all three branching ratios but barely exceeds 0.1 %, this is negligible.

7.3.4 The gap

One of the challenges that the decay channel $\mathrm{B} \to X\ell\nu_\ell$ poses is the so-called gap: The sum of all branching ratios of well-measured exclusive decay modes of $\mathrm{B} \to X\ell\nu_\ell$ does not match the branching ratio measured by inclusive analyses. This difference of branching ratios is called the "gap."

In order to obtain a simulation sample that comes close to the branching ratio observed in data, this tension needs to be resolved ("filling the gap"). This means adjusting the branching ratios of decays already included in the simulation or adding new decays.

Because the understanding of the gap has evolved since the time that Belle MC was produced, we replace the gap component of Belle MC using custom produced samples.

Two different gap models are considered for the tagging calibration:

Non-resonant gap (abbreviated as *NRG*, also referred to as *old gap* in [83, 84]): The gap is filled with $\mathrm{B} \to \mathrm{D}^{(*)}\eta\ell\nu_\ell$ and $\mathrm{B} \to \mathrm{D}^{(*)}\pi\pi\ell\nu_\ell$.

Resonant gap (abbreviated as *RG*, also referred to as *new gap*, *alternative gap* in [83, 84]). The gap is again filled with decays to $\mathrm{D}^{(*)}\eta\ell\nu_\ell$ and $\mathrm{D}^{(*)}\pi\pi\ell\nu_\ell$, but via a D_0^* or D_1' resonance:

- $\mathrm{B} \to \mathrm{D}_0^*(\to \mathrm{D}\eta)\ell\nu_\ell$,
- $\mathrm{B} \to \mathrm{D}_0^*(\to \mathrm{D}^{(*)}\pi\pi)\ell\nu_\ell$,
- $\mathrm{B} \to \mathrm{D}_1'(\to \mathrm{D}^*\eta)\ell\nu_\ell$,
- $\mathrm{B} \to \mathrm{D}_1'(\to \mathrm{D}^{(*)}\pi\pi)\ell\nu_\ell$.

The two modes involving the η make up the majority of the sample.

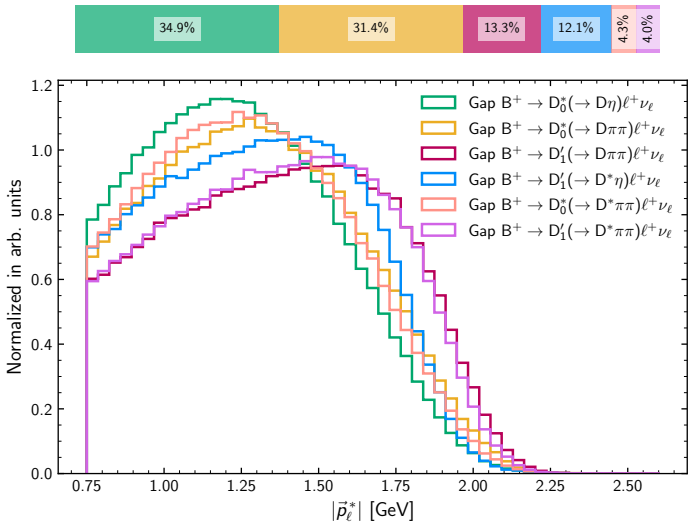
Figure 7.4 compares the $|\vec{p}_\ell^{\,*}|$ distributions of the different gap modes. Both gap descriptions differ considerably: The non-resonant gap peaks at a significantly lower $|\vec{p}_\ell^{\,*}|$ value than the resonant gap (Figure 7.4c). Because of the $|\vec{p}_\ell^{\,*}|$ cut, this leads to differences in event yields at the percent level. Though the effect on the total signal shape is relatively small because the gap only constitutes $\sim 10\,\%$ of the signal component, a difference in the total signal shapes can be perceived as well (Figure 7.4e). For this reason, the calibration will be performed with both of these gap types, and a systematic uncertainty based on the resulting differences will be assigned (see Section 9.4).

Besides the different shapes, the normalizations of the two gaps in the dataset are also different: The total amount of events in the non-resonant gap is only 83 % of that of the resonant gap. This difference is likely caused by the bias towards lower $|\vec{p}_\ell^{\,*}|$ values in the non-resonant gap which means that fewer events pass the $|\vec{p}_\ell^{\,*}| > 0.75\,\mathrm{GeV}$ cut.

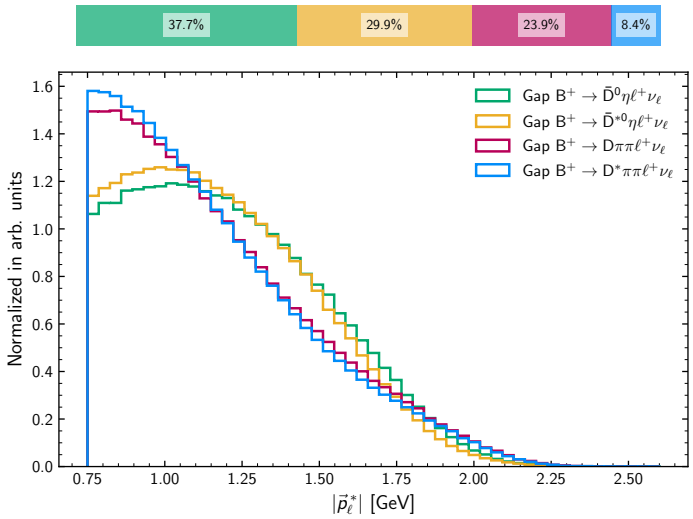
7.3.4.1 Non-resonant gap

We first consider $\mathrm{B} \to \mathrm{D}^{(*)}\pi\pi\ell^+\nu_\ell$. BaBar [324] has measured the ratios

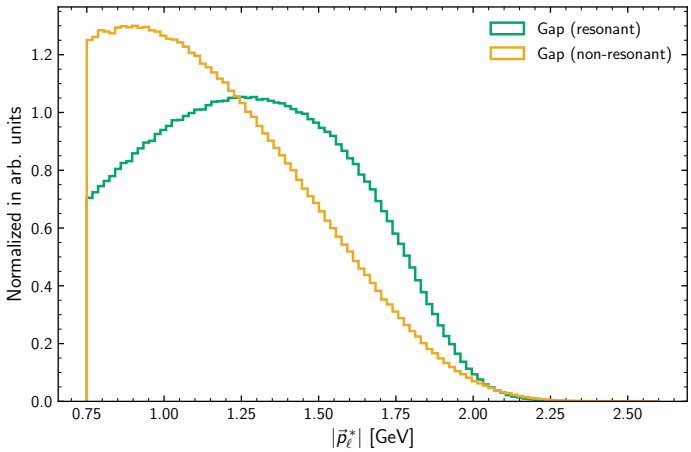
$$\frac{\mathrm{BR}(\mathrm{B} \to \mathrm{D}\pi\pi\ell^+\nu_\ell)}{\mathrm{BR}(\mathrm{B} \to \mathrm{D}\ell\nu_\ell)} = 0.067 \pm 0.010_{\mathrm{stat}} \pm 0.008_{\mathrm{syst}}, \tag{7.3.19a}$$



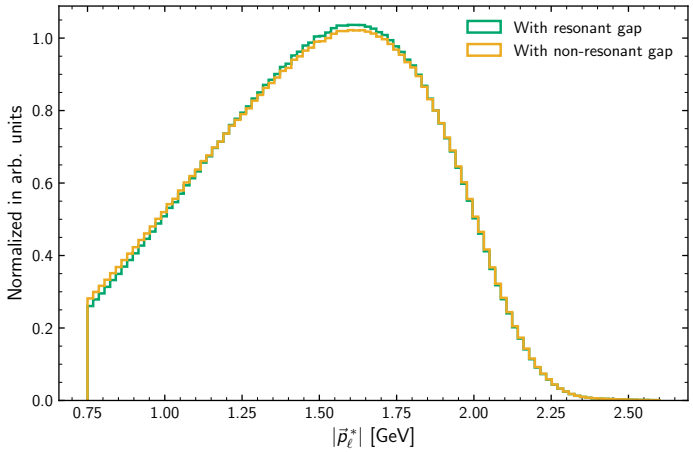
(a) The modes of the resonant gap.



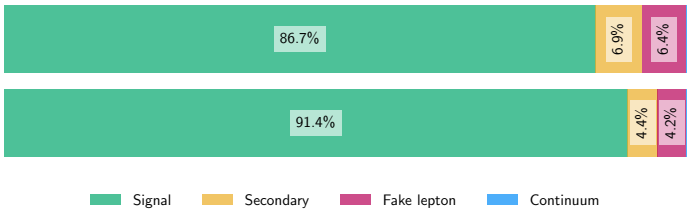
(b) The modes of the non-resonant gap.



(c) Comparing the resonant with the non-resonant gap.



(d) Complete signal component with resonant or non-resonant gap.



(e) Fit components of the resonant gap model (top) and the non-resonant gap model (bottom).

Figure 7.4: The two different gap models vs $|\vec{p}^*_\ell|$. The color bars on top of Figures 7.4a and 7.4b show the relative normalizations of the different processes within the respective model. The fit components are introduced in Section 9.2.4.

$$\frac{\mathrm{BR}(\mathrm{B} \to \mathrm{D}^{*}\pi\pi\ell^{+}\nu_{\ell})}{\mathrm{BR}(\mathrm{B} \to \mathrm{D}^{*}\ell\nu_{\ell})} = 0.019 \pm 0.005_{\mathrm{stat}} \pm 0.004_{\mathrm{syst}}. \tag{7.3.19b}$$

However, this measurement of $\mathrm{B} \to \mathrm{D}\pi\pi\ell^{+}\nu_{\ell}$, this includes contributions from $\mathrm{D}_1(\to \mathrm{D}\pi\pi)\ell^{+}\nu_{\ell}$, which were calculated in Section 7.3.2.1 and are removed to avoid double counting. Using the $\mathrm{B} \to \mathrm{D}^{(*)}\ell\nu_{\ell}$ branching ratios (Section 7.3.1), we obtain

$$\mathrm{BR}(\mathrm{B}^{0} \to \mathrm{D}\pi\pi\ell^{+}\nu_{\ell}) = (0.06 \pm 0.08)\%, \quad \mathrm{BR}(\mathrm{B}^{+} \to \mathrm{D}\pi\pi\ell^{+}\nu_{\ell}) = (0.06 \pm 0.09)\%, \tag{7.3.20a}$$

$$\mathrm{BR}(\mathrm{B}^{0} \to \mathrm{D}^{*}\pi\pi\ell^{+}\nu_{\ell}) = (0.20 \pm 0.10)\%, \quad \mathrm{BR}(\mathrm{B}^{+} \to \mathrm{D}^{*}\pi\pi\ell^{+}\nu_{\ell}) = (0.22 \pm 0.10)\%. \tag{7.3.20b}$$

We now consider the difference between the inclusive branching ratio from Equation (7.3.1) and the sum of all exclusive decays specified so far[2]. This difference is divided between $\mathrm{B} \to \mathrm{D}^{(*)}\eta\ell\nu_{\ell}$ with a 50 %-50 % split between both modes. Because this is purely guesswork, a 100 % uncertainty is assumed. Thus,

$$\begin{aligned} \mathrm{BR}(\mathrm{B}^{0} \to \mathrm{D}\eta\ell\nu_{\ell}) &= \mathrm{BR}(\mathrm{B}^{0} \to \mathrm{D}^{*}\eta\ell\nu_{\ell}) = (0.40 \pm 0.40)\%, \\ \mathrm{BR}(\mathrm{B}^{+} \to \mathrm{D}\eta\ell\nu_{\ell}) &= \mathrm{BR}(\mathrm{B}^{+} \to \mathrm{D}^{*}\eta\ell\nu_{\ell}) = (0.40 \pm 0.40)\%. \end{aligned} \tag{7.3.21}$$

7.3.4.2 Resonant gap

The resonant gap model has identical final states with the non-resonant gap model, and we assume the same branching ratios for them. The only difference is that we consider them as resonant decays with a D_0^{*} or D_1' intermediate state. For the $\mathrm{D}^{(*)}\pi\pi\ell^{+}\nu_{\ell}$ final state, we divide the branching ratio equally between the D_0^{*} and D_1' resonance and assign 100 % uncertainty. Using Equations (7.3.20a) and (7.3.20b), we obtain

$$\mathrm{BR}(\mathrm{B}^{0} \to \mathrm{D}_0^{*}(\to \mathrm{D}\pi\pi)\ell^{+}\nu_{\ell}) = \mathrm{BR}(\mathrm{B}^{0} \to \mathrm{D}_1'(\to \mathrm{D}\pi\pi)\ell^{+}\nu_{\ell}) = (0.03 \pm 0.03)\%, \tag{7.3.22a}$$

$$\mathrm{BR}(\mathrm{B}^{+} \to \mathrm{D}_0^{*}(\to \mathrm{D}\pi\pi)\ell^{+}\nu_{\ell}) = \mathrm{BR}(\mathrm{B}^{+} \to \mathrm{D}_1'(\to \mathrm{D}\pi\pi)\ell^{+}\nu_{\ell}) = (0.03 \pm 0.03)\%, \tag{7.3.22b}$$

$$\mathrm{BR}(\mathrm{B}^{0} \to \mathrm{D}_0^{*}(\to \mathrm{D}^{*}\pi\pi)\ell^{+}\nu_{\ell}) = \mathrm{BR}(\mathrm{B}^{0} \to \mathrm{D}_1'(\to \mathrm{D}^{*}\pi\pi)\ell^{+}\nu_{\ell}) = (0.10 \pm 0.10)\%, \tag{7.3.22c}$$

$$\mathrm{BR}(\mathrm{B}^{+} \to \mathrm{D}_0^{*}(\to \mathrm{D}^{*}\pi\pi)\ell^{+}\nu_{\ell}) = \mathrm{BR}(\mathrm{B}^{+} \to \mathrm{D}_1'(\to \mathrm{D}^{*}\pi\pi)\ell^{+}\nu_{\ell}) = (0.11 \pm 0.11)\%. \tag{7.3.22d}$$

We choose D_0^{*} as resonance for the $\mathrm{D}^{*}\eta\ell^{+}\bar{\nu}_{\ell}$ final state and D_1' for $\mathrm{D}\eta\ell^{+}\bar{\nu}_{\ell}$. This means that the resonant version of Equation (7.3.21) is

$$\mathrm{BR}(\mathrm{B}^{0} \to \mathrm{D}_0^{*}(\to \mathrm{D}\eta)\ell^{+}\bar{\nu}_{\ell}) = \mathrm{BR}(\mathrm{B}^{0} \to \mathrm{D}_1'(\to \mathrm{D}^{*}\eta)\ell^{+}\bar{\nu}_{\ell} = 0.40 \pm 0.40, \tag{7.3.23a}$$

$$\mathrm{BR}(\mathrm{B}^{+} \to \mathrm{D}_0^{*}(\to \mathrm{D}\eta)\ell^{+}\bar{\nu}_{\ell}) = \mathrm{BR}(\mathrm{B}^{+} \to \mathrm{D}_1'(\to \mathrm{D}^{*}\eta)\ell^{+}\bar{\nu}_{\ell} = 0.40 \pm 0.40. \tag{7.3.23b}$$

7.3.5 D branching ratios

We also update the D branching ratios based on [325]. The updated branching ratios are listed in Table 7.5.

7.4 Particle identification correction

Particle identification (PID) is the assignment of a particle hypothesis to an objected reconstructed to the raw detector data (see Section 3.4). This identification is far from perfect and the particle hypothesis is oftentimes incorrect. Misidentified particles are referred to as *fakes* (e.g., *fake electron* for an object mistakenly identified as an electron). If this misidentification behaves identically in MC and data, we have nothing to worry about: the issue will cancel out in the physical results. Unfortunately, most algorithms are (necessarily) trained on MC data and thus perform slightly better on it. For this reason, corrections are necessary.

[2] We have included the ignored decays from Section 7.3.3 in the sum of exclusive modes – however, this barely makes any difference.

Decay	BR [%]
$\rho^+ K^-$	11.30
$\bar{K}^0 \pi^+ \pi^- \pi^-$	10.40
$K^- \pi^+ \pi^- \pi^-$	8.86
$K^{*-} \rho^+$	6.79
$a_1^+ K^-$	4.33
$K^{*-} \pi^+$	3.95
$K^- \pi^+$	3.95
$K^- e^+ \nu_e$	3.54
$K^- \mu^+ \nu_\mu$	3.41
$K^- \pi^+ \omega$	3.10
$\bar{K}^0 \pi^+ \pi^- \pi^- \pi^- \pi^-$	2.68
$\bar{K}^0 \pi^+ \pi^- \pi^- \pi^-$	2.60
$\bar{K}^0 \pi^-$	2.48
$\bar{K}^{*0} \pi^+ \pi^-$	2.40
$\omega \bar{K}^0$	2.22
$K^{*-} e^+ \nu_e$	2.15
$\bar{K}^{*0} \pi^-$	1.95
$\bar{K}^{*0} \eta$	1.90
$\bar{K}^0 \eta'$	1.90
$K^{*-} \mu^+ \nu_\mu$	1.89
$K^- \pi^+ \pi^+ \pi^-$	1.81
$\bar{K}^0 \pi^- e^+ \nu_e$	1.44
$\bar{K}^0 \pi^- \mu^+ \nu_\mu$	1.44
$\bar{K}^{*0} \pi^+ \pi^- \pi^-$	1.30
$\rho^0 \bar{K}^0$	1.26
$K^- \pi^+ \pi^-$	1.15
$\bar{K}^{*0} \omega$	1.10
$\bar{K}^{*0} \pi^- \pi^-$	1.07
$\pi^+ \pi^- \pi^- \pi^-$	1.02
$\bar{K}^0 \eta$	1.02
$\rho^+ \pi^-$	1.01
$\bar{K}^{*0} \rho^0$	1.01

(a) D^0 branching ratios

Decay	BR [%]
$\bar{K}^0 \mu^+ \nu_\mu$	8.76
$\bar{K}^0 e^+ \nu_e$	8.73
$K^- \pi^+ \pi^+$	7.09
$a_1^+ \bar{K}^0$	6.95
$\rho^+ \bar{K}^0$	6.40
$\bar{K}^{*0} e^+ \nu_e$	5.40
$\bar{K}^{*0} \mu^+ \nu_\mu$	5.27
$\bar{K}^{*0} \pi^- \pi^+$	4.72
$\bar{K}_1'^0 \pi^+$	3.80
$K^{*+} \bar{K}^0$	3.40
$\bar{K}^0 \pi^+$	3.12
$\phi \pi^+ \pi^-$	2.30
$\bar{K}^0 \pi^+ \pi^-$	1.70
$\bar{K}_0^{*0} \pi^+$	1.52
$\pi^+ \pi^+ \pi^- \pi^-$	1.16
$\bar{K}_1^{0\prime} \mu^+ \nu_\mu$	1.06
$\bar{K}^{*0} \pi^+$	1.04
$\bar{K}^0 \bar{K}^0 K^+$	1.02

(b) D^+ branching ratios

Table 7.5: D branching ratios. Only branching ratios exceeding 1 % are shown. The smaller branching ratios are shown in Table D.1.

Because we want to include all tag side effects in the calibration factors that we determine in this study, we only need to consider the signal side, that is, only the single lepton that we reconstruct.

We divide the issue into two parts:

1. *Lepton identification efficiency*: Is there a difference in the reconstruction efficiency of true leptons between data and MC (given true leptons, how many do we reconstruct as leptons)?
2. *Fake rate*: How often are non-leptons reconstructed as leptons, and is this number different between data and MC?

The first question is studied in [326]. Correction factors are determined in bins of the lab frame momentum $|\vec{p}_\ell|$ and the polar angle θ_ℓ from the sample $e^-e^+ \to \gamma\gamma \to \ell^-\ell^+$. To account for any effects of the hadronic environment (which is not present in the two-photon study), a smaller dataset of the inclusive decay $B \to XJ/\psi(\to \ell^+\ell^-)$ is studied. Systematic differences between the correction factors determined with both methods are assigned as an uncertainty. The two SVD configurations are considered separately, and a systematic uncertainty for further run period dependencies is added.

For the fake rate, we use the results of [327]. Again, correction factors are determined in bins of $|\vec{p}_\ell|$ and θ_ℓ. Furthermore, different selections for the leptons are distinguished.

To include the uncertainty and correlations of these corrections, we sample 20 variations from a Gaussian distribution describing the correction factors within uncertainties. For each bin in $(|\vec{p}_\ell|, \theta_\ell)$ we set

$$w_i = w_i^0 + e_i^{\text{stat}} r_i^{\text{stat}} + e_i^{\text{syst}} r^{\text{syst}}, \tag{7.4.1}$$

where w_i^0 is the nominal correction weight and e_i^{stat} and e_i^{syst} are the statistical and systematic uncertainties. r_i^{stat} and r^{syst} are both drawn from normal distributions. Note that we use separate random numbers for each bin in the case of the statistical uncertainties (uncorrelated between bins), while we share it in case of the systematic uncertainty (assumed to be 100 % correlated between bins).

This procedure is done separately for both corrections, for both lepton types, and for the particles that can be misidentified to be leptons.

The effect of the PID correction for correctly reconstructed leptons and fake leptons is shown in Figure 7.5. The effects on the signal component of the fit variable distributions are shown in Figure 7.6 (fit variables and the signal component will be introduced in Section 9.2).

It has been shown that the impact of the lepton fake rate on the fitted calibration factors is vanishing (fake leptons are not included in the signal definition and constitute only a small part of the total dataset). For this reason, they are not separately listed in the tables of uncertainties in Chapter 10.

7.5 Form Factor corrections

In the time since the production of the Belle MC, the description of the hadronic form factors has improved, requiring additional correction weights in MC.

The form factor corrections only affect the shape of kinematic distributions, not the branching ratio. However, because the kinematic variables are correlated with the efficiency of the event selections (such as the cuts involving the lepton momentum), the form factor corrections also impact the normalization (see Figure 7.7).

In contrast, Figure 7.8 shows that the impact on the shape of $|\vec{p}_\ell^{\,*}|$ is barely perceivable.

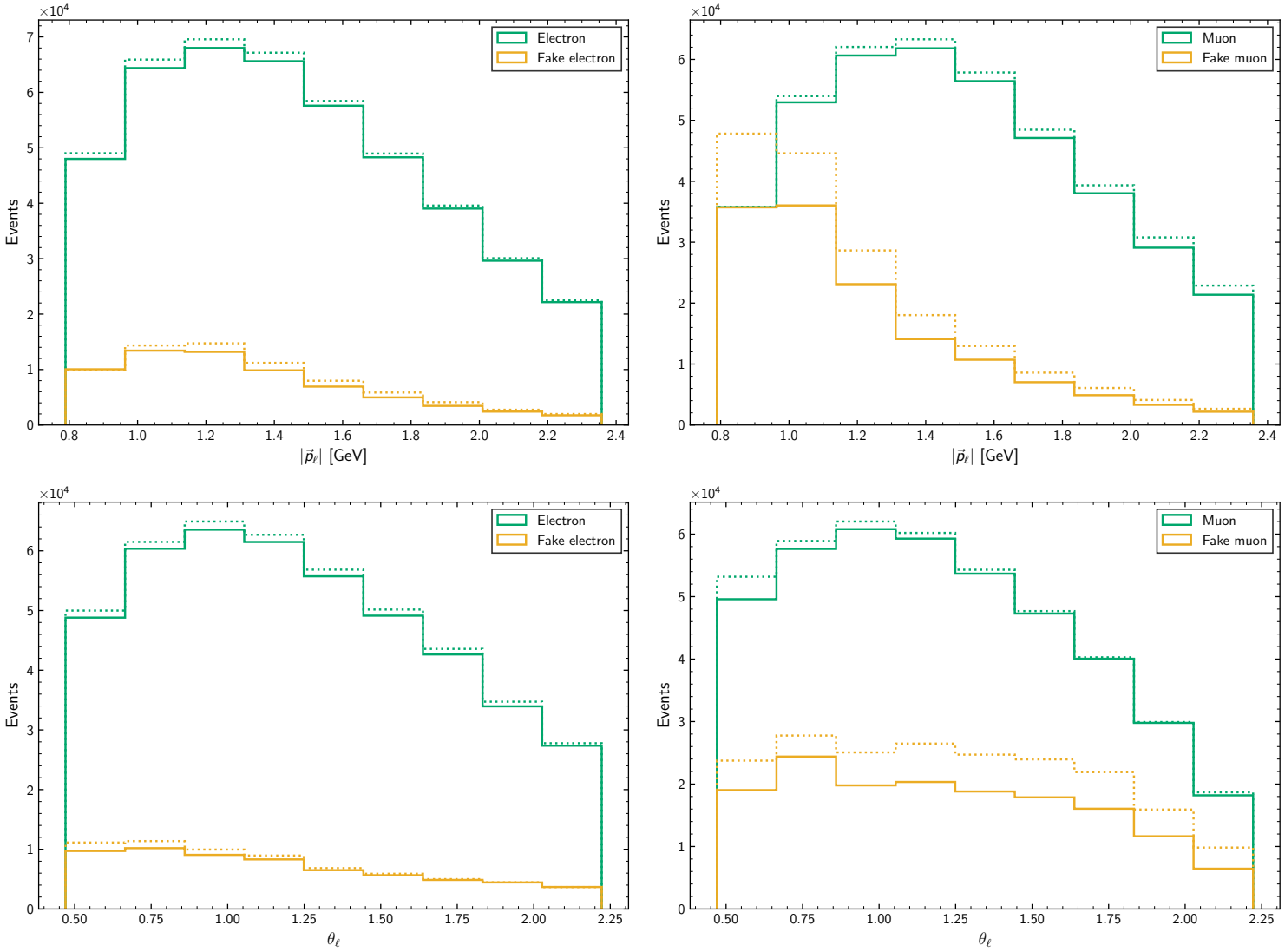


Figure 7.5: PID corrections for electrons and muons. Dotted lines show the $|\vec{p}_\ell|$ and θ_ℓ distributions without PID correction weights, and solid lines include them.

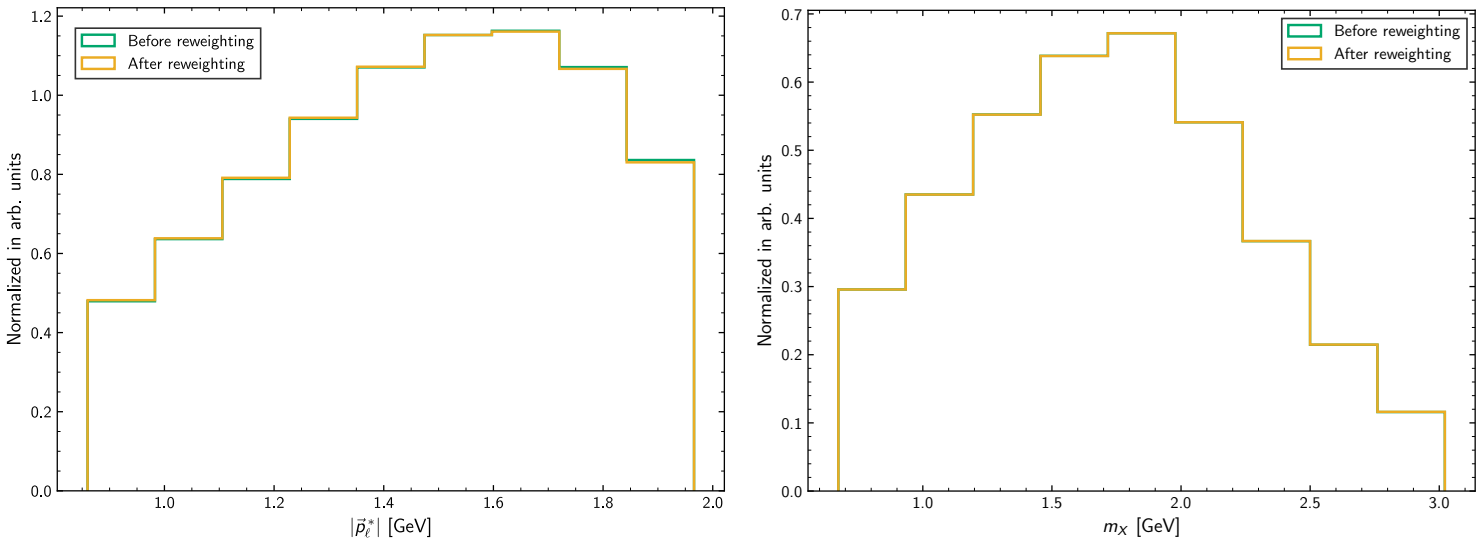


Figure 7.6: Impact of PID reweighting on the fit variable candidates $|\vec{p}^{\,*}_\ell|$ and m_X (defined in Section 9.2.5).

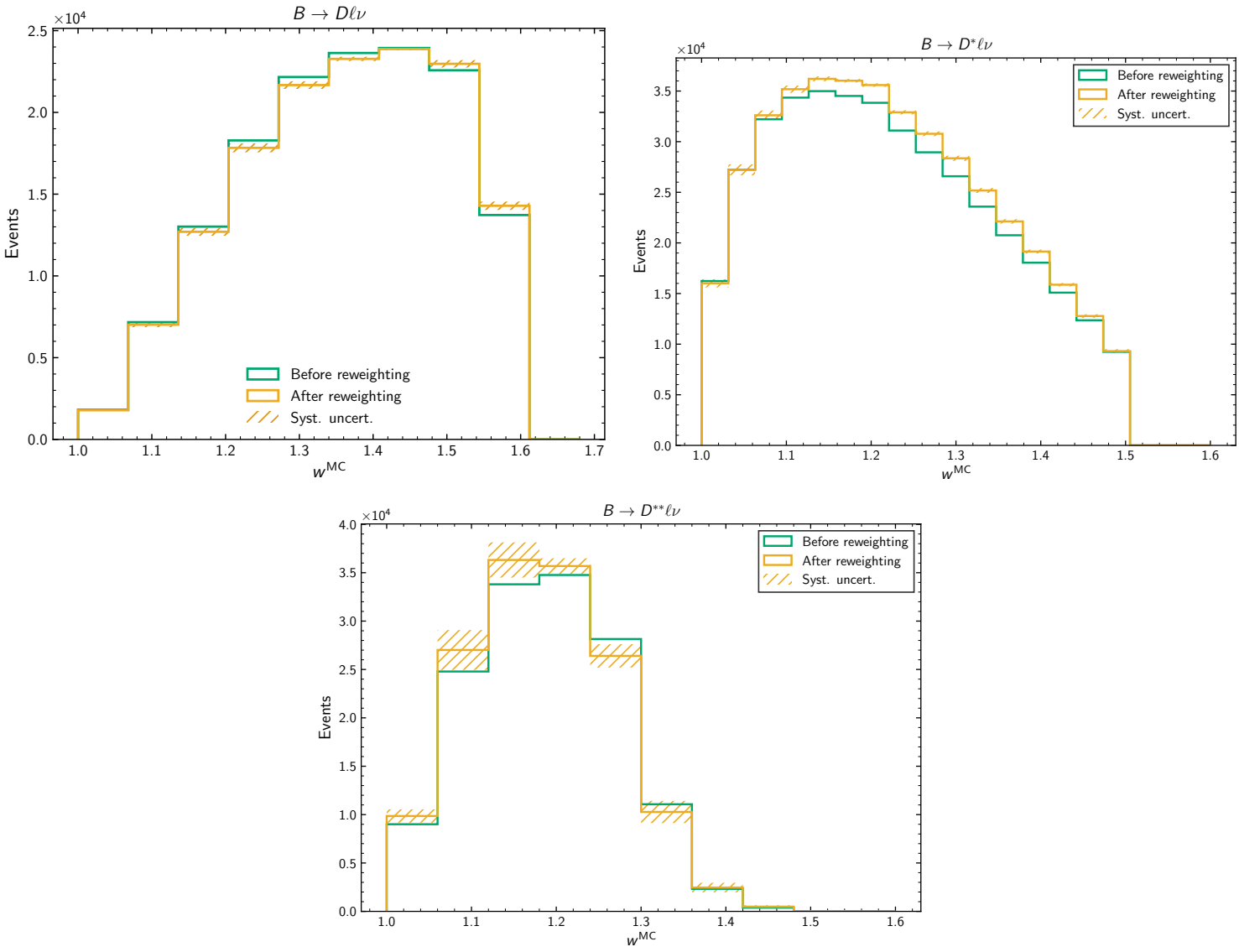


Figure 7.7: Form factor reweighting of $B \to D\ell\nu_\ell$, $B \to D^*\ell\nu_\ell$, and $B \to D^{**}\ell\nu_\ell$.

The following two sections describe the calculation of the correction weights for the different form factors.

7.5.1 $B \to D^{(*)}\ell\nu_\ell$

Our MC dataset uses the $B \to D^{(*)}\ell\nu_\ell$ decays produced in the generic sample. In this sample, CLN was used to model the form factors. Besides using outdated parameters for this model, recent analyses prefer the BGL parameterization (see discussion in Section 1.4.2). Therefore, we update the description by applying event weights. Because the analytical forms of the differential decay rates are known, the weights can simply be calculated as

$$w_{\text{CLN}\to\text{BGL}}(w, \cos\theta_\ell, \cos\theta_V, \chi) := \frac{\frac{\mathrm{d}\Gamma_{\text{BGL}}(w, \cos\theta_\ell, \cos\theta_V, \chi)}{\mathrm{d}w\, \mathrm{d}\cos\theta_\ell\, \mathrm{d}\cos\theta_V\, \mathrm{d}\chi}}{\frac{\mathrm{d}\Gamma_{\text{CLN}}(w, \cos\theta_\ell, \cos\theta_V, \chi)}{\mathrm{d}w\, \mathrm{d}\cos\theta_\ell\, \mathrm{d}\cos\theta_V\, \mathrm{d}\chi}}, \tag{7.5.1}$$

where the dependency on the helicity angles is only relevant for $B \to D^*\ell\nu_\ell$. Note that all variables are to be evaluated on generator level. The implementation of both form factor parameterizations is provided by the `eFFORT` package [328].

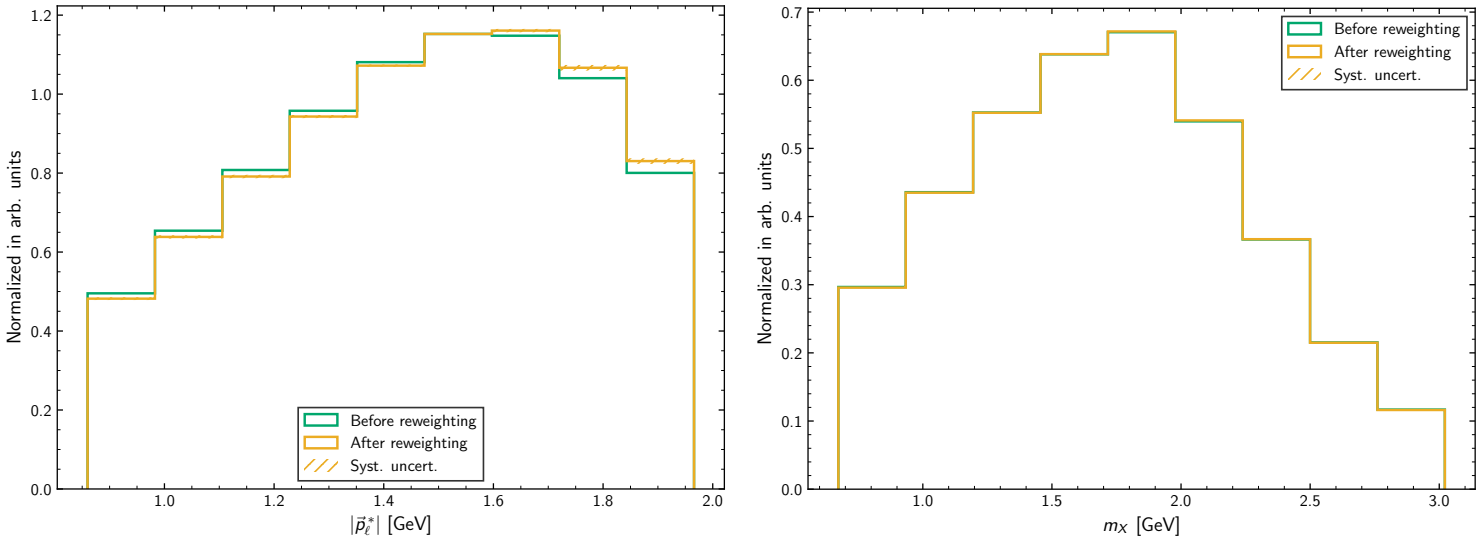


Figure 7.8: Impact of form factor reweighting on fit variable candidates $|\vec{p}^{\,*}_\ell|$ and m_X (defined in Section 9.2.5).

The parameters for BGL are taken from a fit of $(1,1,2)$-BGL to untagged $\mathrm{B}^0 \to \mathrm{D}^{*-}\ell^+\nu_\ell$ Belle data [106, Table V]. To describe the uncertainties on this parameterization, we use the covariance matrix of the parameters, diagonalize it, and consider up and down variations of its eigenvalues to obtain independent up and down variations of the correction weights. These variations are then used in the fit model to calculate covariance matrices of the fit variable bin contents and, finally, the uncertainty on the fit result.

7.5.2 $\mathrm{B} \to \mathrm{D}^{**}\ell\nu_\ell$

We use privately produced dedicated samples to describe $\mathrm{B} \to \mathrm{D}^{**}\ell\nu_\ell$ events (see Section 7.2). Unfortunately, only the ISGW2 form factor parameterization [329] is implemented in the Belle software framework used to produce the MC (production of Belle MC with the Belle II framework is not easily possible). The more recent LLSW parameterization [330] that is for example used at Belle II includes additional $\mathcal{O}(\Lambda_{\mathrm{QCD}}/m_Q)$ corrections that are not considered in ISGW2. Therefore, we apply correction weights to our $\mathrm{B} \to \mathrm{D}^{**}\ell\nu_\ell$ samples to change their form factor model from LLSW to ISGW2.

Instead of employing different analytical expressions for the decay rates and calculating weights as in Equation (7.5.1), we use the results of [331]. In this study, the same MC samples are produced once with the ISGW2 model and once with the LLSW model (using the Belle II software framework).

Formula Equation (7.5.1) is approximated with histograms of the MC data:

$$\hat{w}_{\mathrm{LLSW}\to\mathrm{ISGW2}}(i_w, i_{\cos\theta_\ell}, i_{\cos\theta_V}) = \frac{N^{\mathrm{LSSW}}_{i_w, i_{\cos\theta_\ell}, i_{\cos\theta_V}}}{N^{\mathrm{ISGW2}}_{i_w, i_{\cos\theta_\ell}, i_{\cos\theta_V}}}, \tag{7.5.2}$$

with the histogram bin contents N for the two parameterizations and i_x the bin in the (generator level) variable x. Note that the form factors for the decays involving different D^{**} resonances depend only on a subset of these three variables: w (D_0^*), w and $\cos\theta_\ell$ (D_1', D_2^*), or w, $\cos\theta_\ell$ and $\cos\theta_V$ (D_1). The weights $\hat{w}_{\mathrm{LLSW}\to\mathrm{ISGW2}}(i_w, i_{\cos\theta_\ell}, i_{\cos\theta_V})$ are then translated to a continuous analytical expression $w_{\mathrm{LLSW}\to\mathrm{ISGW2}}(w, \cos\theta_\ell, \cos\theta_V)$ using an interpolation with radial basis functions.

7.6 Additional corrections for $B \to D^* \ell \nu_\ell$

In the analysis of $B \to D^* \ell \nu_\ell$, several additional efficiency corrections are applied (they are not necessary for $X \ell \nu_\ell$, because only a lepton is reconstructed):

Slow pion efficiency Particles with $p_T < 200\,\text{MeV}$ do not reach the detector sub-systems necessary for a successful identification and are always assumed to be pions. The reconstruction efficiency also differs between MC and data, particularly for very low p_T, and must be corrected. For this, the correction factors derived from $B^0 \to D^{*-}\pi^+$ and $B^+ \to \overline{D}^{*0}\pi^+$ in [332] are used.

K^0_S efficiency The K^0_S reconstruction efficiency is corrected using weights from a study of $D^* \to D^0(\to K^0_S \pi^+\pi^-)\pi$ [333]

Neutral pion efficiency The π^0 reconstruction efficiency is corrected with results from a study of $\tau^- \to \pi^0\pi^0\nu_\tau$ [300]

Hadron efficiency and fake rate Similar to the lepton efficiency and fake rate discussed in Section 7.4, the efficiency and fake rates of kaon and pion reconstruction needs to be corrected. This is done using an analysis of $D^{*+} \to D^0(\to K^-\pi^+)\pi^+$ [334].

Chapter 8

Reconstruction of $B \to X\ell\nu_\ell$ events

This chapter describes the reconstruction of hadronically tagged $B \to X\ell\nu_\ell$ decays used for the tagging calibration.

8.1 Reconstruction

The reconstruction comprises three steps:

1. Identification of a single lepton candidate,
2. Reconstruction of B_{tag} candidates with the FEI,
3. Combination of a B_{tag} candidate with a lepton candidate to form a candidate of the decay $\Upsilon(4S) \to B_{tag}(\to \text{had.})\, B_{sig}(\to X\ell\nu_\ell)$.

Several details and implementation notes are listed in Appendix D.1.3.

8.1.1 Charged lepton reconstruction

As mentioned in Section 3.4, charged particles are identified from tracks. On MC, the helix parameters of the fitted tracks and the track momenta are subjected to resolution corrections (*smearing*) [335, 336] in order to improve the agreement with the tracking performance on data.

Tracks need to pass the *impact parameter* requirements

$$\mathrm{d}r < 2\,\mathrm{cm} \quad \text{and} \quad |\mathrm{d}z| < 4\,\mathrm{cm}, \tag{8.1.1}$$

where both variables are defined based on the point (*POCA*) of the trajectory that is closest to the interaction point (*IP*). $\mathrm{d}r$ is the distance between POCA and the IP in the x-y-plane and $|\mathrm{d}z|$ is the distance between POCA and the IP in the z-direction.

Furthermore, we require

$$|\vec{p}_\mathrm{T}| > 0.1\,\mathrm{GeV} \quad \text{and} \quad |\vec{p}_\ell| > \begin{cases} 0.3\,\mathrm{GeV} & \ell = \mathrm{e} \\ 0.6\,\mathrm{GeV} & \ell = \mu, \end{cases} \tag{8.1.2}$$

where $|\vec{p}_\mathrm{T}|$ is the transverse momentum of the lepton. These selection criteria ensure that detector sub-systems used for PID are reached.

To reduce the number of fake leptons, we use the PID likelihoods (see Section 7.4):

$$\mathcal{L}_\ell > \begin{cases} 0.6 & \ell = \mathrm{e} \\ 0.9 & \ell = \mu. \end{cases} \tag{8.1.3}$$

For muons we also require hits in the KLM (`muBelleQuality=1`). The cuts on the likelihoods and track parameters are not subject to optimization because they are a prerequisite to using the PID correction values [326, 327]. They are identical in [83].

Electrons forced on a curved trajectory emit bremsstrahlung photons (*BSP*s), resulting in energy loss and change of direction. To avoid reconstructing the emitted photons as independent particles and to exclude bias to the electron momentum, BSP-corrected electron candidates are formed by adding the BSP four-momenta to the four-momentum of electron candidates. The BSP-corrected electron candidates compete with the uncorrected candidates in the best candidate selection (Section 8.2). For the BSP-correction, all photons in a cone with an opening angle of 2° around the electron trajectory at the POCA that satisfy

$$\frac{|\vec{p}_\gamma|}{|\vec{p}_\mathrm{e}|} < 0.4 \quad \text{and} \quad E_\gamma < 0.4\,\mathrm{GeV} \tag{8.1.4}$$

are considered to be BSPs. Additional details are found in [83].

Finally, we place the following requirement for the center-of-mass momentum of the lepton candidate:

$$|\vec{p}^{\,*}_\ell| > 0.75\,\mathrm{GeV}. \tag{8.1.5}$$

This cut reduces background and data volume for the B → $X\ell\nu_\ell$ decay. Furthermore, fake and secondary leptons are known to be insufficiently modeled for low lepton momenta.

Duplicated tracks

Tracks of low transverse momentum are strongly bent in the magnetic field. These *curling* trajectories are occasionally incorrectly reconstructed as multiple tracks. Duplicated curling tracks are removed by a procedure described in [83]. While this step is negligible for the high momentum signal leptons, it is used in Section 8.1.2.

8.1.2 Tag side reconstruction and whole event selection

If no leptons that satisfy the conditions of the previous section are found, the event is discarded. We also exclude events with more than 18 tracks fulfilling the impact parameter requirements of Equation (8.1.1), or more than 17 additional photons in the event.

In the next step, the FEI reconstructs B meson candidates.

Finally, we apply the selections

$$m_\mathrm{bc}^\mathrm{tag} > 5.27\,\mathrm{GeV} \quad \text{and} \quad -0.15 < \Delta E_{\mathrm{B_{tag}}}/\mathrm{GeV} < 0.1 \tag{8.1.6}$$

that reduce background from incorrectly reconstructed $\mathrm{B_{tag}}$ mesons and are identically applied in the $|V_\mathrm{cb}|$ and $\mathcal{R}(\mathrm{D}^{(*)})$ analysis. Both variables have been defined in Equation (4.1.1). A small subtlety regarding the $m_\mathrm{bc}^\mathrm{tag}$ cut is discussed in Section 8.2.2.

Lepton candidates are then combined with $\mathrm{B_{tag}}$ candidates to form candidates for $\Upsilon(4\mathrm{S}) \to \mathrm{B_{tag}}(\to \text{had.})\,\mathrm{B_{sig}}(\to X\ell\nu_\ell)$. However, note that the X particle is not reconstructed explicitly, and no selection is applied based on its properties.

Rest of Event

Some variables (discussed as fit variables and used in side-studies) use kinematic properties of the X particle. For these variables, we combine all particles of the event that are not associated with the lepton or the $\mathrm{B_{tag}}$. This is referred to as the *rest of event* (ROE). To avoid including background noise, additional selection criteria are applied to the ROE particles (*ROE mask*):

- All charged particles are required to fulfill the impact parameter requirements of Equation (8.1.1). Furthermore, we consolidate duplicated curling low transverse momentum tracks (see Section 8.1.1).

- Photons that are reconstructed from ECL clusters are subjected to energy and azimuthal angle requirements: Only photons that satisfy one of the following three conditions are considered[1]

$$\begin{aligned} &E_\gamma > 0.50\,\text{GeV} \quad \text{and} \quad 32.2^\circ < \theta_\gamma < 128.7^\circ \quad \text{(Barrel region)}, \quad \text{or} \\ &E_\gamma > 0.10\,\text{GeV} \quad \text{and} \quad 12.4^\circ < \theta_\gamma < 31.4^\circ \quad \text{(Forward endcap region)}, \quad \text{or} \\ &E_\gamma > 0.15\,\text{GeV} \quad \text{and} \quad 130.7^\circ < \theta_\gamma < 155.1^\circ \quad \text{(Backward endcap region)}. \end{aligned} \tag{8.1.7}$$

8.1.3 Reconstruction in the $|V_{cb}|$ analysis

The $|V_{cb}|$ analysis introduced in Section 1.4 reconstructs the signal side in the $B \to D^* \ell \nu_\ell$ mode. In addition to the charged leptons, charged kaons and pions are reconstructed with the same selection criteria of Section 8.1.1 except for the $|\vec{p}^*_\ell|$ cut. The kaon and pion likelihood must satisfy $\mathcal{L} > 0.1$. All photon candidates must satisfy Equation (8.1.7).

The considered decay changes of the D^* decay have been shown in Table 5.1. Several requirements are placed on the intermediate particles:

- K^0_S mesons are reconstructed from V0s and need to satisfy $\Delta M/\sigma_M < 3$, where $\Delta M := |M_{\text{reco}} - M_{\text{real}}|$ is the difference between the reconstructed mass of a candidate and the nominal mass of a K^0_S, and σ_M is the resolution of M_{reco}. Furthermore, the standard quality requirements described in [337] need to be fulfilled.

- Neutral pions need to fulfill $E > 0.1\,\text{GeV}$, $\Delta M/\sigma_M < 3$.

- D mesons must not exceed a CMS momentum of 3 GeV. Furthermore, a series of ΔM cuts are applied depending on the reconstruction modes.

- D^* candidates need to pass channel dependent $M_{D^*} - M_D$ and pion momentum selections.

Reconstructed $B \to D^* \ell \nu_\ell$ decays need to have a successfully fitted decay vertex. B_{tag} candidates are reconstructed and selected as described for $B \to X \ell \nu_\ell$. However, after B_{tag} and B_{sig} candidates have been combined to $\Upsilon(4S)$ candidates, no additional track in the ROE is allowed. While the B_{tag} selection in the $B \to X \ell \nu_\ell$ reconstruction is almost entirely independent from the signal side selection, this *completeness constraint* significantly affects the selection of the B_{tag} candidates.

8.2 Best Candidate Selection

[2]About four $\Upsilon(4S)$ candidates per event are reconstructed. A best candidate selection (BCS) selects the most promising candidate for every event.

It is performed in three steps:

[1]This standard requirement is available as the `goodBelleGamma` variable.

[2]The numbers and plots in this section correspond to a small sample of generic $b \to c$ and continuum MC and data without the continuum suppression introduced in Section 8.3. With the exception of Figure 8.2 (which considers candidates rather than events), the results include all event weights on MC introduced in Sections 7.4 and 7.5.

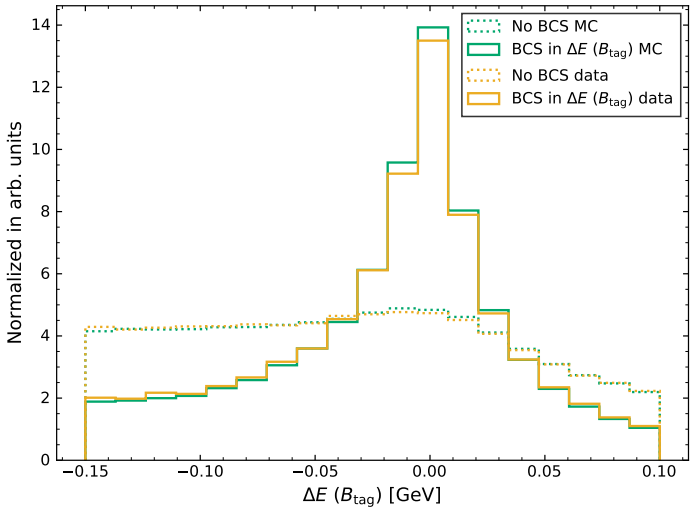


Figure 8.1: BCS in $|\Delta E_{\mathrm{B_{tag}}}|$.

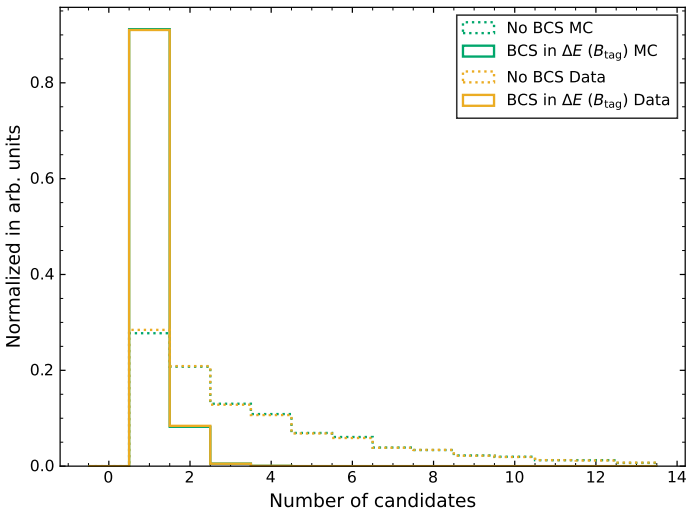
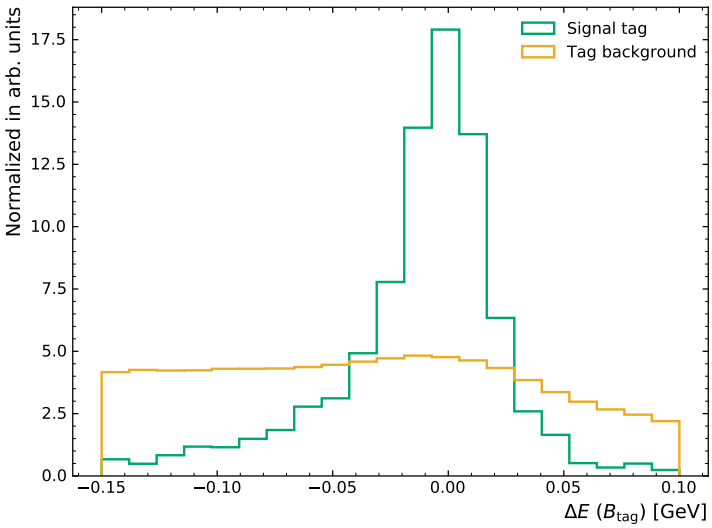
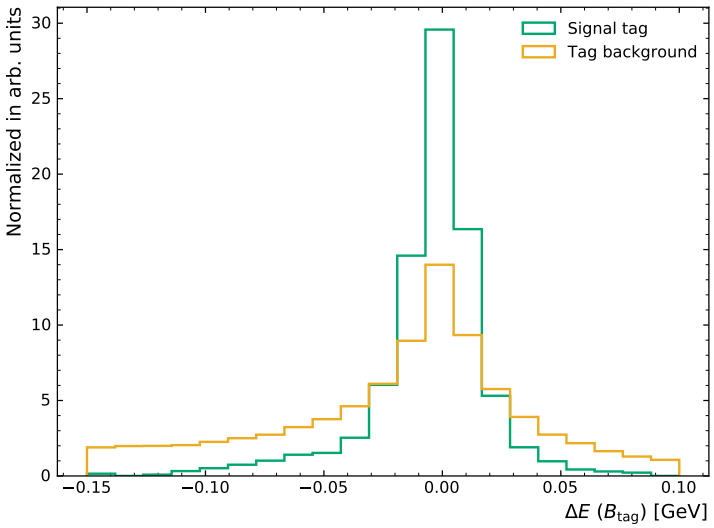


Figure 8.2: Number of candidates before and after BCS in $|\Delta E_{\mathrm{B_{tag}}}|$.



(a) Before BCS in $|\Delta E_{\mathrm{B_{tag}}}|$.



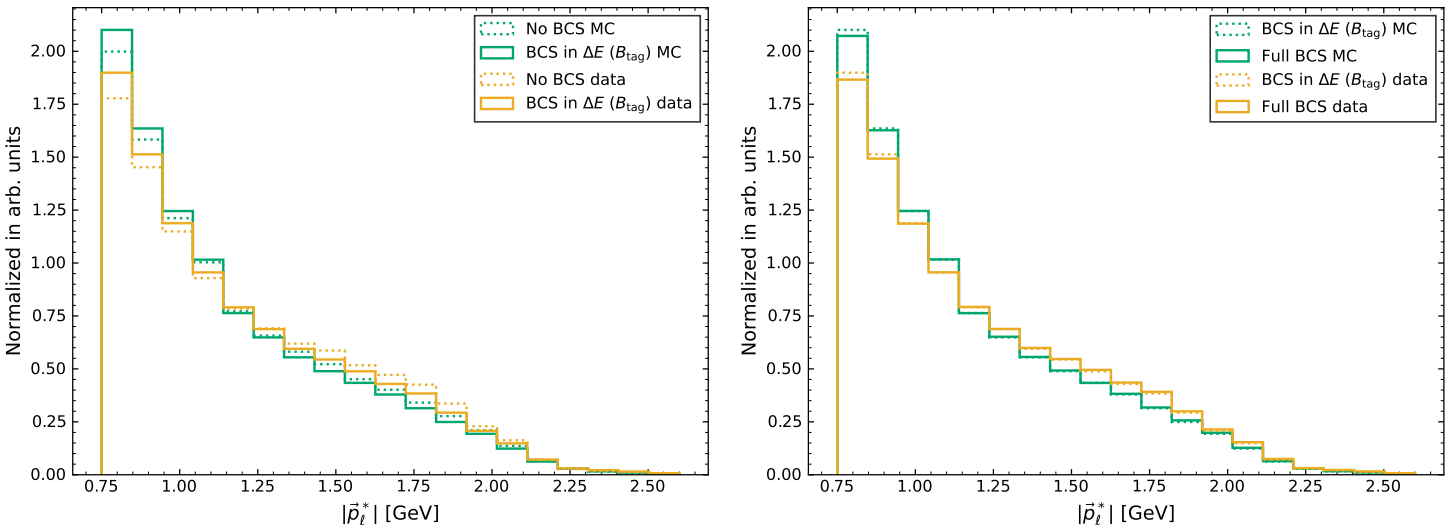
(b) After BCS in $|\Delta E_{\mathrm{B_{tag}}}|$.

Figure 8.3: $\Delta E_{\mathrm{B_{tag}}}$ distribution of correctly and incorrectly constructed $\mathrm{B_{tag}}$ mesons (denoted "Signal tag" and "Tag background" in the figure).

1. The $\Upsilon(4S)$ candidates with the lowest value of $|\Delta E_{\mathrm{B_{tag}}}|$ are selected. This selection is already performed online to reduce disk space requirements. The number of candidates per event is reduced to ≈ 1.1 (Table 8.1). The distributions of the numbers of candidate per event are very similar in data and MC and are shown in Figure 8.2. The $\Delta E_{\mathrm{B_{tag}}}$ distribution before and after this step is shown in Figure 8.1 and separately for correctly and incorrectly reconstructed $\mathrm{B_{tag}}$ mesons in Figure 8.3.

2. If multiple candidates remain, the $\Upsilon(4S)$ candidate with the highest $|\vec{p}^{\,*}_\ell|$ is selected. This reduces the number of candidates to ≈ 1.003 (Table 8.1). The $|\vec{p}^{\,*}_\ell|$ distributions before and after the second step are shown in Figure 8.4b.

3. If multiple candidates remain, a random selection is applied.



(a) $|\vec{p}^*_\ell|$ before and after BCS in $|\Delta E_{\mathrm{B_{tag}}}|$.

(b) $|\vec{p}^*_\ell|$ after BCS in $|\Delta E_{\mathrm{B_{tag}}}|$ and after full BCS.

Figure 8.4: BCS in $|\vec{p}^*_\ell|$

Dataset	#Cand	#Cand after ΔE BCS	#Cand after $|\vec{p}^*_\ell|$ BCS
MC	4.0079	1.09407	1.00309
Data	3.9754	1.09602	1.00376

Table 8.1: Number of candidates per event after the first and second stage of the BCS.

8.2.1 Best candidate selection in the $|V_{\mathrm{cb}}|$ analysis

In contrast, the BCS of the $|V_{\mathrm{cb}}|$ analysis first considers the total energy stored in ECL clusters that are not associated to the $\Upsilon(4S)$ candidate or to any charged track in the event. This observable is denoted $\mathrm{n}E^{\mathrm{ECL}}_{\mathrm{extra}}$ and can be used for completeness constraints. By selecting the $\Upsilon(4S)$ candidate with minimal $\mathrm{n}E^{\mathrm{ECL}}_{\mathrm{extra}}$, incorrectly reconstructed events are suppressed.

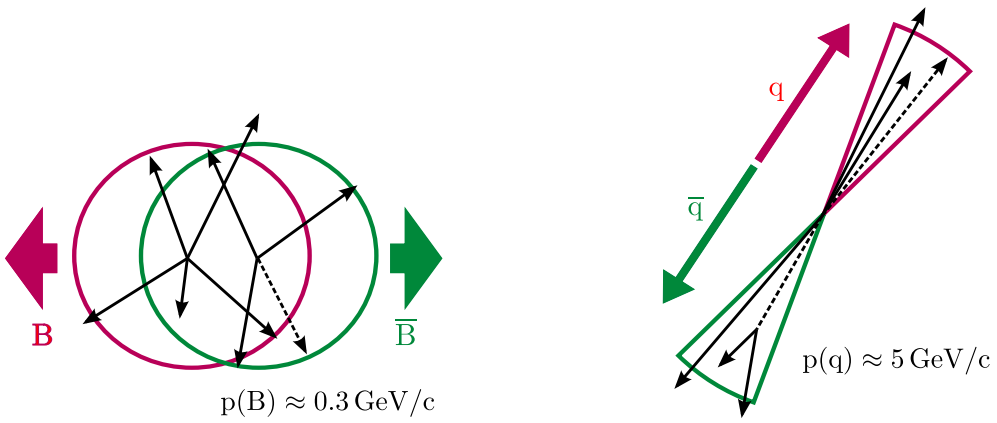
If multiple candidates remain, the candidate with lowest $|\Delta E_{\mathrm{B_{tag}}}|$ is chosen. If multiple candidates still remain, a random selection is performed.

This BCS selection is more effective in suppressing incorrectly reconstructed $\mathrm{B_{tag}}$ mesons than the one used for $\mathrm{B} \to X\ell\nu_\ell$ (note that $\mathrm{n}E^{\mathrm{ECL}}_{\mathrm{extra}}$ is not usable there). Together with the track completeness constraint (Section 8.1.3), this further separates the composition of $\mathrm{B_{tag}}$ mesons of the $\mathrm{B} \to \mathrm{D}^*\ell\nu_\ell$ dataset and the $\mathrm{B} \to X\ell\nu_\ell$ dataset and explains the different distributions of $\mathcal{P}_{\mathrm{FEI}}$ observed in Figure 6.1.

8.2.2 Small subtlety regarding order of operations of BCS and $m^{\mathrm{tag}}_{\mathrm{bc}}$ cut

This section describes a detail that is relevant for the understanding of a comparison between the inclusive and the exclusive dataset in Appendix C.3 (see discussion in Appendix D.6.1).

To retain a small amount of $m^{\mathrm{tag}}_{\mathrm{bc}}$ sideband, only a $m^{\mathrm{tag}}_{\mathrm{bc}} > 5.265$ cut is applied in the first reconstruction steps of the analysis. These first steps are referred to as *online analysis* and permanently discard all events that are excluded by selections. It has high practical value to apply the $|\Delta E_{\mathrm{B_{tag}}}|$ BCS as part of the online analysis to reduce the disk storage volume. Therefore, the order of operations is as follows:



(a) B$\overline{\mathrm{B}}$ event (b) Continuum event

Figure 8.5: Comparing the event kinematics of $\Upsilon(4S)$ decays and continuum events in the center of mass reference frame. The high momentum back-to-back jets of continuum events can be clearly distinguished from the lower momentum isotropic decay products of B decays. Figures adapted from [338].

(1) Require $m_{\mathrm{bc}}^{\mathrm{tag}} > 5.265$ (online),

(2) BCS in $|\Delta E_{\mathrm{B_{tag}}}|$ (online),

(3) Require $m_{\mathrm{bc}}^{\mathrm{tag}} > 5.27$ (offline),

(4) BCS in $|\vec{p}_\ell^{\,*}|$ (offline),

(5) Random BCS (offline),

where *offline* refers to selection steps later in the analysis. It is important to highlight that the BCS and cut-based selection steps do not commute.

8.3 Continuum suppression

Continuum events (Section 7.2) are typically the dominant source of background at B-factories because of their large cross section (twice the cross section of the $\Upsilon(4S)$ at $\sqrt{s} = 10.58\,\mathrm{GeV}$!). Fortunately, they are relatively easy to separate from B$\overline{\mathrm{B}}$ events: Due to the low mass difference between the $\Upsilon(4S)$ and the B$\overline{\mathrm{B}}$ system, the B mesons are produced almost at rest in the $\Upsilon(4S)$ center of mass system. In this reference system, the momenta of the B decay products are thus distributed relatively isotropically (Figure 8.5a). In contrast, the light quark pairs of the continuum events are produced with large momenta and decay in two back-to-back jets of light hadrons (Figure 8.5b). For this reason, various observables based on the momenta of the decay products have been developed.

Following the $|V_{\mathrm{cb}}|$ measurement selection, we first apply a very soft $R_2^{\mathrm{tag}} < 0.6$ cut (variable introduced below). We then apply a boosted decision tree (BDT) that has been trained to distinguish continuum and B$\overline{\mathrm{B}}$ events and cut on its output $\mathcal{L}_{\mathrm{CS}} > 0.5$ (Figure 8.8). The input variables of this multivariate classifier are introduced below.

To avoid bias on the $\mathrm{B_{tag}}$ selection, we do not train the BDT on the B → $X\ell\nu_\ell$ sample. Instead, we apply the BDT trained on the tag side of the B → $\mathrm{D}^*\ell\nu_\ell$ tuple and thus perform precisely the same continuum suppression as the $|V_{\mathrm{cb}}|$ analysis. As a drawback, the performance of the BDT is worse than one would usually expect had it been trained on the same dataset. The training uses the `FastBDT` implementation [241, 242] and is described in [83, 84].

The following input variables are used as input to the training:

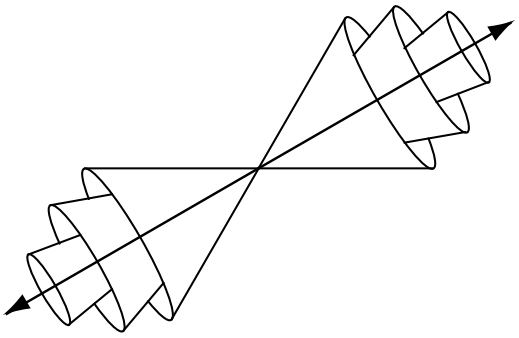


Figure 8.6: The first three Cleo cones. The axis points in the $\pm\vec{T}$ direction. Figure adapted from [340].

Thrust variables The thrust T is one of the earliest (and simplest) variables for continuum suppression. It is defined as [339]

$$T = \max_{\vec{T} \text{ with } |\vec{T}|=1} \frac{\sum_{i=1}^{N} \left|\vec{T}\vec{p}_i\right|}{\sum_{i=1}^{N} |\vec{p}_i|}, \tag{8.3.1}$$

T reaches its maximum of 1 if all momenta are parallel (back-to-back jets of continuum) and tends towards its minimum for isotropic momentum distributions (B decays). This observable is considered for both the signal and tag side. The vector $\vec{T}$ where the maximum in Equation (8.3.1) is attained is called the *thrust axis*. The angle between the thrust axis of the tag side and the signal side is called θ_{TBTO} and is another discriminatory variable: The back-to-back jet structure of continuum events produces a peak at $\cos\theta_{\mathrm{TBTO}} = 1$, while the distribution remains flat for $\mathrm{B\overline{B}}$ events. The angle between the tag side thrust and the beam (z) axis, $\theta_{\mathrm{TB}z}$, is used as another training variable.

Cleo cones For this observable [340], concentric pairs of cones parallel to $\pm\vec{T}$ with opening angle θ_j^{CC} are defined (illustrated in Figure 8.6). The observables are then given by the sum of momenta between two adjacent cones,

$$\mathrm{CleoCone}_j := \sum_{i \,\mathrm{s.t.}\, \theta_{j-1}^{\mathrm{CC}} < \pm\theta_i < \theta_j^{\mathrm{CC}}} |\vec{p}_i|, \tag{8.3.2}$$

where θ_i is the angle of the i-th particle relative to $\vec{T}$ and $\theta_0 := 0$. Traditionally, nine cones are chosen (θ steps of 10°).

Fox-Wolfram Moments This set of observables H_k is defined as [341, 342]

$$H_k = \sum_{i,j=1}^{N} |\vec{p}_i||\vec{p}_j| P_k(\cos\theta_{ij}), \tag{8.3.3}$$

where θ_{ij} is the angle between the momenta of a pair of particles and P_k the k-th Legendre polynomial. In this analysis, we only use the *reduced Fox Wolfram Moment* $R_2 = H_2/H_0$ (in addition to the use as a training variable, a very soft cut is applied).

KSFW Moments The Kakuno Super Fox Wolfram (KSFW) Moments are extended variants of the Fox Wolfram Moments using input from the B decay and the remaining particles in the event. In addition to the momenta, they also consider information about the electrical charges. In total, there are 16 observables, explained in detail in [187, p. 114].

All input variables are shown in Figures D.1 to D.6.

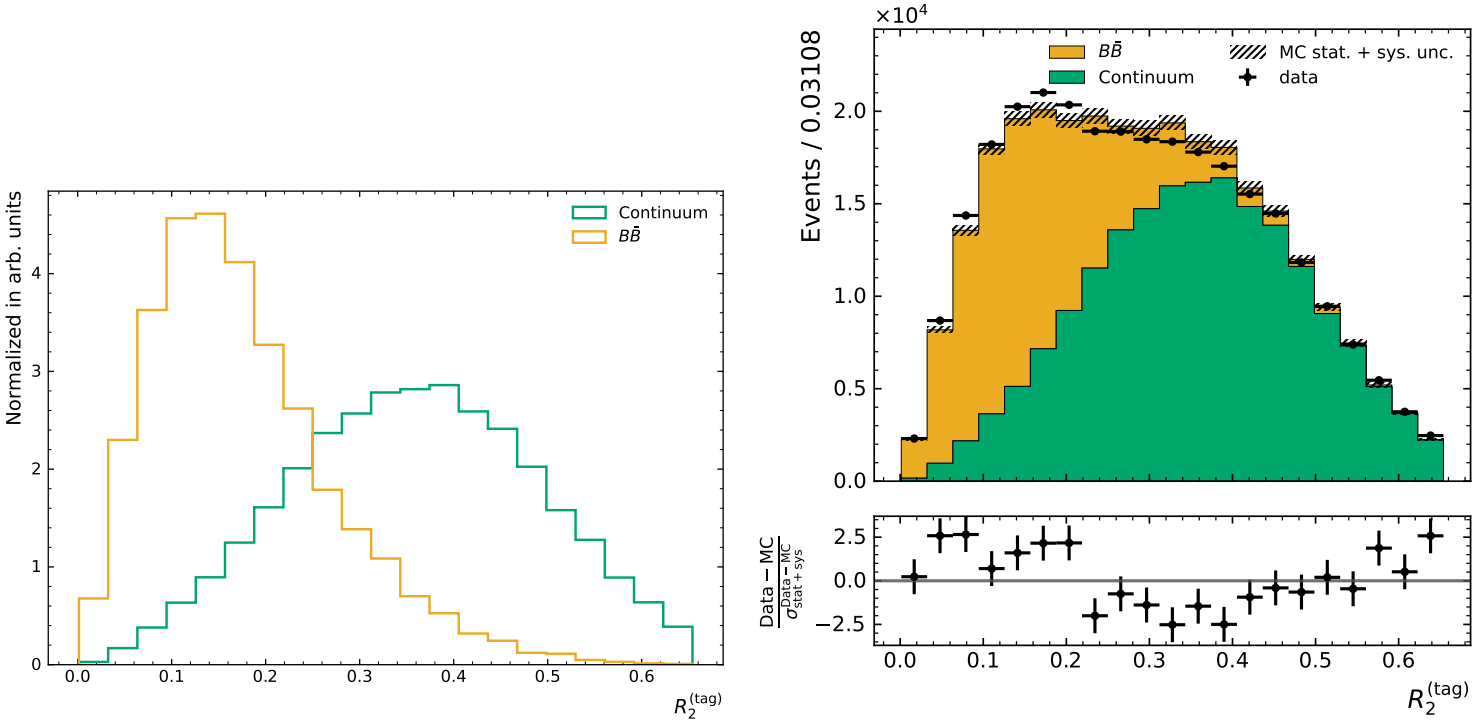


Figure 8.7: The reduced Fox Wolfram moment R_2^{tag}. We apply a $R_2^{\text{tag}} < 0.6$ selection. The variable is also used in the training of the continuum suppression BDT. The data-MC plot does not include calibration and is generated with 10 % of data.

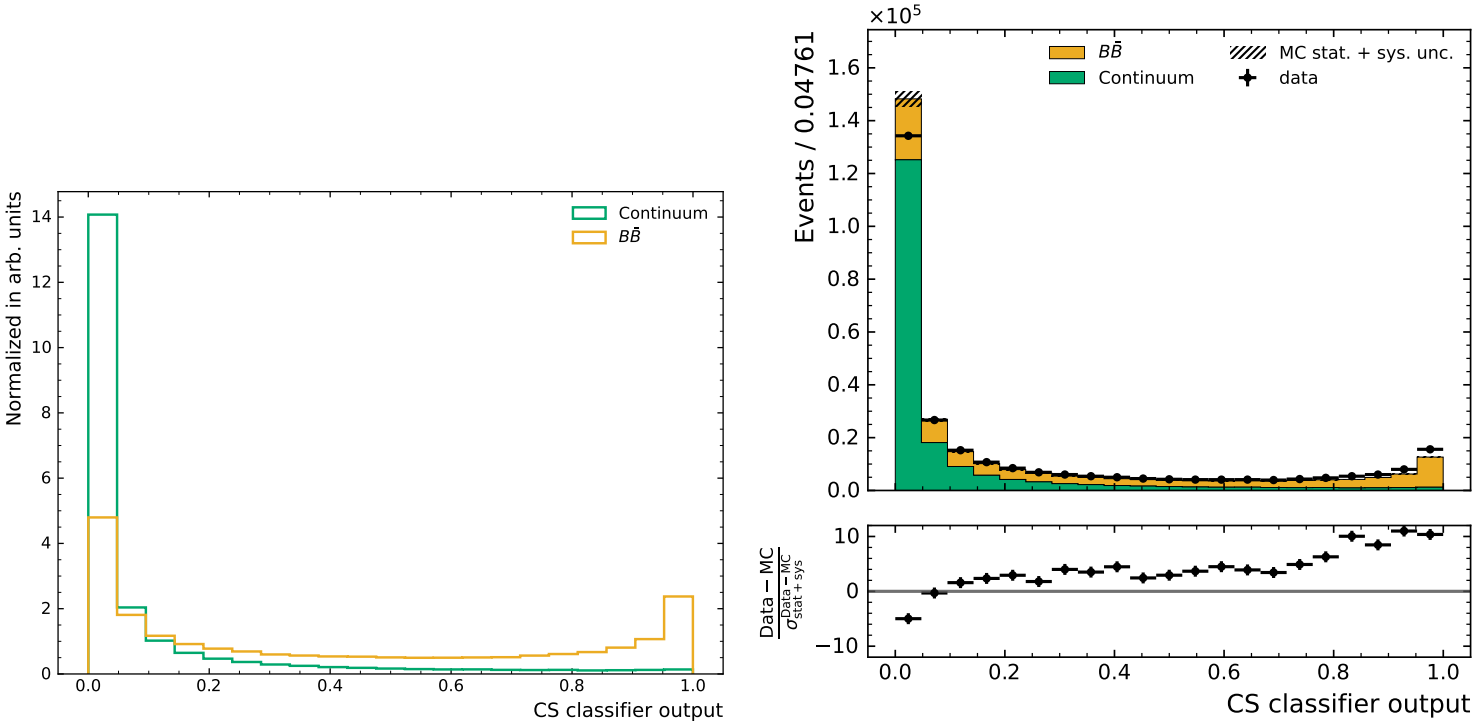


Figure 8.8: Output of continuum suppression BDT $\mathcal{L}_{\text{CS}}$. We apply a $\mathcal{L}_{\text{CS}} > 0.5$ selection. The data-MC plot does not include calibration and is generated with 10 % of data.

Chapter 9

Calibration procedure

As introduced in Chapter 6, the calibration factors are calculated using fits to kinematic distributions of hadronically tagged $B \to X\ell\nu_\ell$ decays. This chapter goes into the details of the background subtraction and the calculation of the uncertainties on the calibration factors.

Separate calibration factors are calculated for disjoint subsets of the data, which we call *categories*. A two-dimensional maximum likelihood fit with three fit components is used for background subtraction. $|\vec{p}_\ell^{\,*}|$ is used as the primary fit variable. In some categories, a simplified one-dimensional fit configuration is used instead. The calibration is performed separately for the dataset with the resonant and non-resonant gap models. We then take the average of both results and assign an appropriate uncertainty based on the difference.

9.1 Categorization

As explained in Chapter 6, the calibration factors are calculated separately in bins of the FEI classifier output $\mathcal{P}_{\text{FEI}}$ and FEI decay channel C_{FEI}. Each of the selections $(C_{\text{FEI}} \in C) \wedge (p_0 < \mathcal{P}_{\text{FEI}} < p_1)$ for which we determine a separate calibration factor is called a *category*. Here, C denotes a set of hadronic decay channels and $0 \leq p_0 < p_1 \leq 1$.

In the FEI training used in this analysis, 54 reconstruction modes are distinguished. They are listed in Table 9.1. Note that the reconstruction modes are only described to the level of the first daughter particle (no $D^{(*)}$ decay channels are distinguished). However, based on the side-study described in Appendix D.4, we know that the different $D^{(*)}$ reconstruction modes can have a significant impact on the calibration factors and should therefore ideally be considered separately.

The categorization is a bias-variance tradeoff: If we use too many categories, we reduce bias but increase the statistical uncertainty of the calibration factors (this is discussed in more detail in Appendix D.3.1). In contrast, if we use very few categories, we reduce statistical fluctuations but might not be able to transfer the calibration factors to the exclusive dataset.

The different categories are numbered and denoted C_i $(i = 0, \ldots, N_{\text{cat}} - 1)$. They are sorted by C_{FEI} first (arbitrary but consistent order, see Table 9.1), ascending $\mathcal{P}_{\text{FEI}}$ second.

The remainder of this section describes the categorization that has proven to be reasonable to us and is used for the results presented in this thesis. The categories will be made apparent in the results (see e.g., Figure 10.1a or Figure 10.1a). Different categorization schemes will be compared in Appendix B.1, which also assigns a systematic uncertainty for the remaining bias.

The categorization used in our results

We consider each decay channel separately. For the sub-decay channels, we consider:

- The decay channel of the first daughter particle if the first daughter particle is a D meson,

- The decay channel of the first daughter particle and its first daughter if the first daughter particle is a D^* meson.

The order of the particles is specified in Table 9.1 and for the $D^{(*)}$ reconstruction modes in Table D.2. Out of this list of sub-decay channels, only those with a total weight exceeding a predefined value N_{sdc} are considered separately. For simplicity, we will still denote this decay channel (with or without specified sub-decay channel or collection thereof) as C_{FEI}.

For a given C_{FEI}, the bins in $\mathcal{P}_{\text{FEI}}$ are chosen as quantiles in the corresponding $B \to X\ell\nu_\ell$ data. The more data we have for a decay channel, the more quantiles in $\mathcal{P}_{\text{FEI}}$ we want to choose. We define the number of quantiles as

$$n_{\text{quantiles}}(n) = \begin{cases} \left\lceil \frac{n}{c_1} \right\rceil, & n \leq N_1 \\ \max\left\{ \left\lceil \frac{n}{N_1} \right\rceil, \left\lfloor \frac{n}{c_2} \right\rfloor \right\} & n > N_1, \end{cases} \tag{9.1.1}$$

where n is the number of data events for C_{FEI} and $c_1 < c_2$, and N_1 are constants in $\mathbb{N}$. The idea of this case distinction is to choose at least *some* bins in $\mathcal{P}_{\text{FEI}}$ for C_{FEI} with very few events, even if it comes at the cost of statistical uncertainty. The max in the second case ensures that the number of quantiles rises monotonically with the number of events. For the results presented in the following, we use

$$N_{\text{sdc}} = 20 \times 10^4, \quad N_1 = 3 \times 10^4, \quad c_1 = 1 \times 10^4, \quad c_2 = 3 \times 10^4. \tag{9.1.2}$$

The relatively high value of N_{sdc} means that sub-decay channels are only considered for the decay channels with the highest weights. Otherwise, $\mathcal{P}_{\text{FEI}}$ subdivisions will be preferred. This is reasonable because the variance of the calibration factors in $\mathcal{P}_{\text{FEI}}$ seems to be larger than that between sub-decay channels.

Appendix B.1 calculates calibration factors for different categorization parameters and compares their weighted average when applied to the $B \to D^*\ell\nu_\ell$ dataset. An uncertainty is assigned to the observed differences.

9.2 Background subtraction

This section describes the use of maximum likelihood fits to measure the $B \to X\ell\nu_\ell$ signal yield from the reconstructed $B \to X\ell\nu_\ell$ decays.

9.2.1 Maximum likelihood fits

To determine the calibration factors, binned extended maximum likelihood fits with several components are performed. For this, we use the `TemplateFitter` package [343] developed and described in [83, 331].

This means maximizing the likelihood given by

$$\mathcal{L} = \prod_{i=1}^{N} \mathcal{P}(n_i; \nu_i) \times \prod_{k=1}^{K} \mathcal{G}_k, \tag{9.2.1}$$

where the first product runs over the bins in the fit variables and the second over the fitted components. Both terms are described in the following. $P(n_i; \nu_i)$ describes a Poisson distribution,

$$\mathcal{P}(n_i; \nu_i) = \frac{1}{n_i!} \nu_i^{n_i} e^{-\nu_i}, \tag{9.2.2}$$

ID	Channel	ID	Channel	ID	Channel
0	$D^-\pi^+$	19	$D^{*-}\pi^+\pi^+\pi^-\pi^0$	38	$\bar{D}^0D^{*0}K^+$
1	$D^-\pi^+\pi^0$	20	$D_s^{*+}D^-$	39	$\bar{D}^{*0}D^{*0}K^+$
2	$D^-\pi^+\pi^0\pi^0$	21	$D_s^+D^{*-}$	40	$D_s^+\bar{D}^0$
3	$D^-\pi^+\pi^+\pi^-$	22	$D_s^{*+}D^{*-}$	41	$\bar{D}^{*0}\pi^+$
4	$D^-\pi^+\pi^+\pi^-\pi^0$	23*	$J/\psi K_S^0$	42	$\bar{D}^{*0}\pi^+\pi^0$
5	$\bar{D}^0\pi^+\pi^-$	24	$J/\psi K^+\pi^-$	43	$\bar{D}^{*0}\pi^+\pi^0\pi^0$
6	$D^-D^0K^+$	25	$J/\psi K_S^0\pi^+\pi^-$	44	$\bar{D}^{*0}\pi^+\pi^+\pi^-$
7	$D^-D^{*0}K^+$	26	$\bar{D}^0\pi^+$	45	$\bar{D}^{*0}\pi^+\pi^+\pi^-\pi^0$
8	$D^{*-}D^0K^+$	27	$\bar{D}^0\pi^+\pi^0$	46	$D_s^{*+}\bar{D}^0$
9	$D^{*-}D^{*0}K^+$	28	$\bar{D}^0\pi^+\pi^0\pi^0$	47	$D_s^+\bar{D}^{*0}$
10*	$D^-D^+K_S^0$	29	$\bar{D}^0\pi^+\pi^+\pi^-$	48	$\bar{D}^0K^+$
11	$D^{*-}D^+K_S^0$	30	$\bar{D}^0\pi^+\pi^+\pi^-\pi^0$	49	$D^-\pi^+\pi^+$
12*	$D^-D^{*+}K_S^0$	31	$\bar{D}^0D^+$	50	$D^-\pi^+\pi^+\pi^0$
13*	$D^{*-}D^{*+}K_S^0$	32	$\bar{D}^0D^+K_S^0$	51*	$J/\psi K^+$
14	$D_s^+D^-$	33*	$\bar{D}^{*0}D^+K_S^0$	52	$J/\psi K^+\pi^+\pi^-$
15	$D^{*-}\pi^+$	34	$\bar{D}^0D^{*+}K_S^0$	53	$J/\psi K^+\pi^0$
16	$D^{*-}\pi^+\pi^0$	35*	$\bar{D}^{*0}D^{*+}K_S^0$	54	$J/\psi K_S^0\pi^+$
17	$D^{*-}\pi^+\pi^0\pi^0$	36	$\bar{D}^0D^0K^+$		
18	$D^{*-}\pi^+\pi^+\pi^-$	37	$\bar{D}^{*0}D^0K^+$		

Table 9.1: Mapping of *our* FEI decay channel ID and corresponding reconstruction mode. Note that the FEI uses a different numbering scheme (also referred to as decay channel ID) that does not distinguish between isospin conjugated decay modes. For an illustration of the frequencies of these decay channels, see Figure 6.2. Decay modes marked with an asterisk have $\mathcal{P}_{\text{FEI}} \equiv 0.$ and are excluded in the $B \to D^{(*)}\ell\nu_\ell$ dataset.

where n_i is the number of observed data events in bin i and ν_i is the corresponding expectation value. The expected yield ν_i is given by

$$\nu_i = \sum_{k=1}^{K} f_{ik}\eta_k, \tag{9.2.3}$$

where k runs over the different components of the fit, f_{ik} describes the shape of the component ($\sum_i f_{ik} = 1\ \forall k$) and η_k are the normalizations of the different components, i.e., the yields extracted by the fitter. The shape f_{ik} depends on nuisance parameters θ_{ik} that describe systematic uncertainties on the shape. The nuisance parameters are constrained by the Gaussian terms $\mathcal{G}_k$:

$$\mathcal{G}_k = \mathcal{G}_k(\vec{0}, \vec{\theta}_k, \Sigma_k) = \frac{1}{\sqrt{(2\pi)^N \det(\Sigma_k)}} \exp\left(-\frac{1}{2}(\vec{\theta}_k^{\top}\Sigma_k^{-1}\vec{\theta}_k)\right), \tag{9.2.4}$$

where $\vec{\theta}_k = (\theta_{ik})_i \in \mathbb{R}^N$ and θ_{ik} is the *nuisance parameter* for template k bin i; $\Sigma_k \in \mathbb{R}^{N\times N}$ is the covariance matrix of $\vec{\theta}_k$.

In the implementation, we minimize $-\ln\mathcal{L}$ rather than maximizing $\mathcal{L}$ because of performance benefits. We also approximate (9.2.2) via the Stirling approximation as

$$\begin{aligned} -\ln\mathcal{P}(n_i;\nu_i) &= -n_i\ln\nu_i + \nu_i + \ln(n_i!) = \\ &= \nu_i - n_i - n_i\ln\frac{\nu_i}{n_i} + \Theta(\ln n_i), \end{aligned} \tag{9.2.5}$$

and drop all constants in $-\ln \mathcal{G}_k$. Taking this together, the function to minimize is given by

$$f(\eta_k) = \sum_{i=1}^{N} \left(\nu_i - n_i - n_i \ln \frac{\nu_i}{n_i} \right) + \frac{1}{2} \sum_{k=1}^{K} \vec{\theta}_k^{\top} \Sigma_k^{-1} \vec{\theta}_k. \tag{9.2.6}$$

9.2.2 Goodness of fit

To check the quality of our data model and to exclude questionable calibration factors, we calculate a χ^2 figure from the minimized function f (Equation (9.2.6)). The Poisson distribution can be approximated with a Gaussian distribution:

$$\mathcal{P}(n_i; \nu_i) \approx \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{(n_i - \nu_i)^2}{2\sigma_i^2} \right), \tag{9.2.7}$$

Therefore, by comparison to Equation (9.2.5) (and $\sqrt{2\pi}\sigma_i$ negligible):

$$\nu_i - n_i - n_i \ln \frac{\nu_i}{n_i} \approx \frac{1}{2} \frac{(n_i - \nu_i)^2}{\sigma_i^2}. \tag{9.2.8}$$

Using this approximation in Equation (9.2.6), we obtain

$$2f(\eta_k) \approx \sum_{i=1}^{N} \frac{(n_i - \nu_i)^2}{\sigma_i^2} + \sum_{k=1}^{K} \vec{\theta}_k^{\top} \Sigma_k^{-1} \vec{\theta}_k. \tag{9.2.9}$$

Thus, we can approximately identify $2f$ with a χ^2_r distribution to obtain p-values. The number of degrees of freedom, r, is approximately given by the number of bins corrected by the number of floating components in the fit (that is, fit parameters η_k): $r = N - K$. The p-value distributions of our fit setup are discussed in Section 9.2.6.

9.2.3 Failed fits

A fit is considered to be successful if and only if the following criteria are met:

i) The minimization is successful, i.e., no is error reported by the `minuit` algorithm [344].

ii) The calibration factors ϵ_i^c of all components c (see Section 9.3) are within reasonable bounds: $\epsilon_i^c - \Delta\epsilon_i^c < 10$.

iii) The fitted yield $N_{c,i}^{\text{Data}}$ of all components is larger than zero (within 1σ): $N_{c,i}^{\text{Data}} + \Delta N_{c,i}^{\text{Data}} \geq 0$.

iv) $p \geq 10^{-4}$ for the p-value of the χ^2 test described in Section 9.2.2. This criterion is not applied for fits without systematic uncertainties used in the determination of the statistical part of the fitting uncertainties (Section 9.3.1).

A fit that does not meet these criteria is called a *failed fit*.

If a fit fails item i or ii, we repeat the fit procedure with different initial values for the fit parameters. The first attempt will be made with initial values from the MC expectation (i.e., corresponding to unit calibration factors). In subsequent attempts, we set the initial values of a subset of the fit parameters to half of their expectation and try again. This considerably reduces the number of categories for which no successful fit can be performed.

9.2.4 Fit components

We distinguish between the following events/components in the fits:

Signal The lepton is a true electron or muon and is the direct decay product of a true B meson.

Continuum No $\Upsilon(4S)$ meson is produced (see Section 7.2).

Fake lepton ("fake") The event is not a continuum event, but the particle reconstructed as an electron or muon is not a charged lepton, or the flavor or charge has been incorrectly reconstructed.

Secondary lepton ("secondary") The $\Upsilon(4S)$ decays to a pair of B mesons, and the charged electron or muon is correctly reconstructed. However, the lepton is not the *direct* product of a true B meson but is produced in the decay of an intermediate product. For example, events where the selected lepton is the decay product of a τ belong in this category.[1]

Every event falls into exactly one of these components. As already mentioned in Section 6.1 the reason for this distinction (and for having to fit at all) is that the signal component defined in this way is better understood than the background components. Isolating signal thus removes potential bias to the calibration factors that could be caused by an inaccurate description of the signal side decays in MC.

The three background components are generally difficult to distinguish, negatively impacting fit stability. For this reason, we usually either fit the signal component against all background components or against continuum and the combined fake and secondary lepton components.

Note that no information from the tag side is used in this signal definition: We currently do not distinguish between correctly and incorrectly reconstructed B_{tag} mesons. This follows all previous approaches to the tagging calibration and is also consistent with the approach of the $|V_{cb}|$ analysis: As the B_{tag} decay is independent of the B_{sig} decay, the reconstruction of the B_{tag} is also considered to decouple from the physics of the B_{sig} decay. For example, the $|\vec{p}^*_\ell|$ distribution is almost identical between samples with correctly reconstructed B_{tag} mesons and those where even the flavor of the B_{tag} is incorrectly reconstructed (see Figure 11.13). This means that it would generally be difficult for analyses to distinguish correctly and incorrectly reconstructed B_{tag} mesons (in addition to increased statistical uncertainties, which are already challenging in tagged analyses). For this reason, incorrectly reconstructed tags are to be calibrated (i.e., treated as signal in the calibration) as well.

However, the conclusion of Chapter 11 is that these previous practices are very problematic because the calibration factors required for correctly and incorrectly reconstructed B_{tag} mesons can differ significantly. This is discussed in detail in Chapter 11.

9.2.5 Choosing fit variables

This section motivates our choice of fit variables. The final fit setup is described in Section 9.2.6.

In multi-dimensional fits, we call the variable with more bins the *primary* variable and name it first, and the other variable (if any) the *secondary* variable.

The following variables have been considered as primary variables:

- $|\vec{p}^*_\ell|$: The lepton momentum in the center of mass frame of the B_{sig}.

[1] Technical detail: this definition does not exclude electrons exhibiting final state radiation, which is treated as electrons "decaying" into electrons and photons by the reconstruction framework. See Section 8.1.1 for the treatment of final-state radiation.

- m_X: Calculated as

$$m_X = \sqrt{E_{\text{ROE}}^2 - |\vec{p}_{\text{ROE}}|^2}, \tag{9.2.10}$$

where E_{ROE} and $\vec{p}_{\text{ROE}}$ are the energy and momentum of all particles in the ROE (all particles but the lepton and those included in the B_{tag} reconstruction that pass the selection criteria defined by the ROE mask from Section 8.1.2).

- $m_{\text{miss}}^2 = E_{\text{miss}}^2 - |\vec{p}_{\text{miss}}|^2$: The invariant mass squared of the missing momentum, calculated from the energies and momenta of the charged particles and photons (with applied ROE mask) in the center of mass system. Note that this quantity does not have the usual interpretation because the X particle is included in the ROE. In particular, the background components peak stronger at zero than the signal component.

- $E_{\text{miss}} - |\vec{p}_{\text{miss}}|$, where E_{miss} and $|\vec{p}_{\text{miss}}|$ are calculated as for m_{miss}^2.

$|\vec{p}_\ell^{\,*}|$ shows the clearest separation between signal and background. Furthermore, it has the most straightforward physical interpretation and is less model-dependent than the other variables. For this reason, it is chosen as fit variable in this calibration. $|\vec{p}_\ell^{\,*}|$ is also the variable that has been used to calibrate the FEI for Belle II data [307] and the FR algorithm for Belle data in similar studies of $\text{B} \to X\ell\nu_\ell$ [306].

In Figure 9.1 the $|\vec{p}_\ell^{\,*}|$ distributions of continuum events, fake lepton events and secondary events are almost identical. However, this is only the case on average: In particular, the $|\vec{p}_\ell^{\,*}|$ distributions between continuum and the joint fake and secondary lepton component are increasingly diverging for low values of $\mathcal{P}_{\text{FEI}}$ (Figure 9.3).

As the normalization of the continuum component is not well known and because the FEI might behave differently for continuum events than for $\text{B}\overline{\text{B}}$ events, we seek to fit continuum events as a separate component. However, the continuum events are challenging to fit against the other background components using only $|\vec{p}_\ell^{\,*}|$. Therefore, a second variable is used that is specifically sensitive to the continuum component.

The observables $\mathcal{L}_{\text{CS}}$, R_2^{tag}, and $\cos(\theta_{\text{TBTO}})$ (all described in Section 8.3) have been considered as secondary fit variables. Distributions of the three observables are shown in Figure 9.4. Based on its fit performances, $\mathcal{L}_{\text{CS}}$ is chosen for the final fits.

9.2.6 Fitting setups summarized

For the extraction of the calibration factors, an additional $|\vec{p}_\ell| > 1\,\text{GeV}$ cut is applied to exclude potential bias from fake and secondary leptons that are known to be insufficiently described at low momenta. Additionally, the $|\vec{p}_\ell^{\,*}|$ tail is cropped since a small excess compatible with an underestimated $u\ell\nu$ component is observed at high $|\vec{p}_\ell^{\,*}|$. While both effects only lead to sub-percent level changes in the calibration factors, the large total amount of data allows us to simply exclude both regions.

Two fit setups are used to extract the calibration factors:

Primary fit Two-dimensional fit with five equidistant bins in $|\vec{p}_\ell^{\,*}|$ between 1 GeV and 1.95 GeV and two bins in $\mathcal{L}_{\text{CS}}$ given by

$$\Big[0.5,\ \mu_{1/2}(\mathcal{L}_{\text{CS}})\Big] \quad \text{and} \quad \Big[\mu_{1/2}(\mathcal{L}_{\text{CS}}),\ 1.0\Big],$$

where $\mu_{1/2}(\mathcal{L}_{\text{CS}})$ is the median of $\mathcal{L}_{\text{CS}}$. Three fit components are distinguished: signal, fake/secondary, and continuum. As discussed in Section 9.2.5, the $|\vec{p}_\ell^{\,*}|$ distributions of the fake and secondary component is *almost* identical. To cover any remaining uncertainty

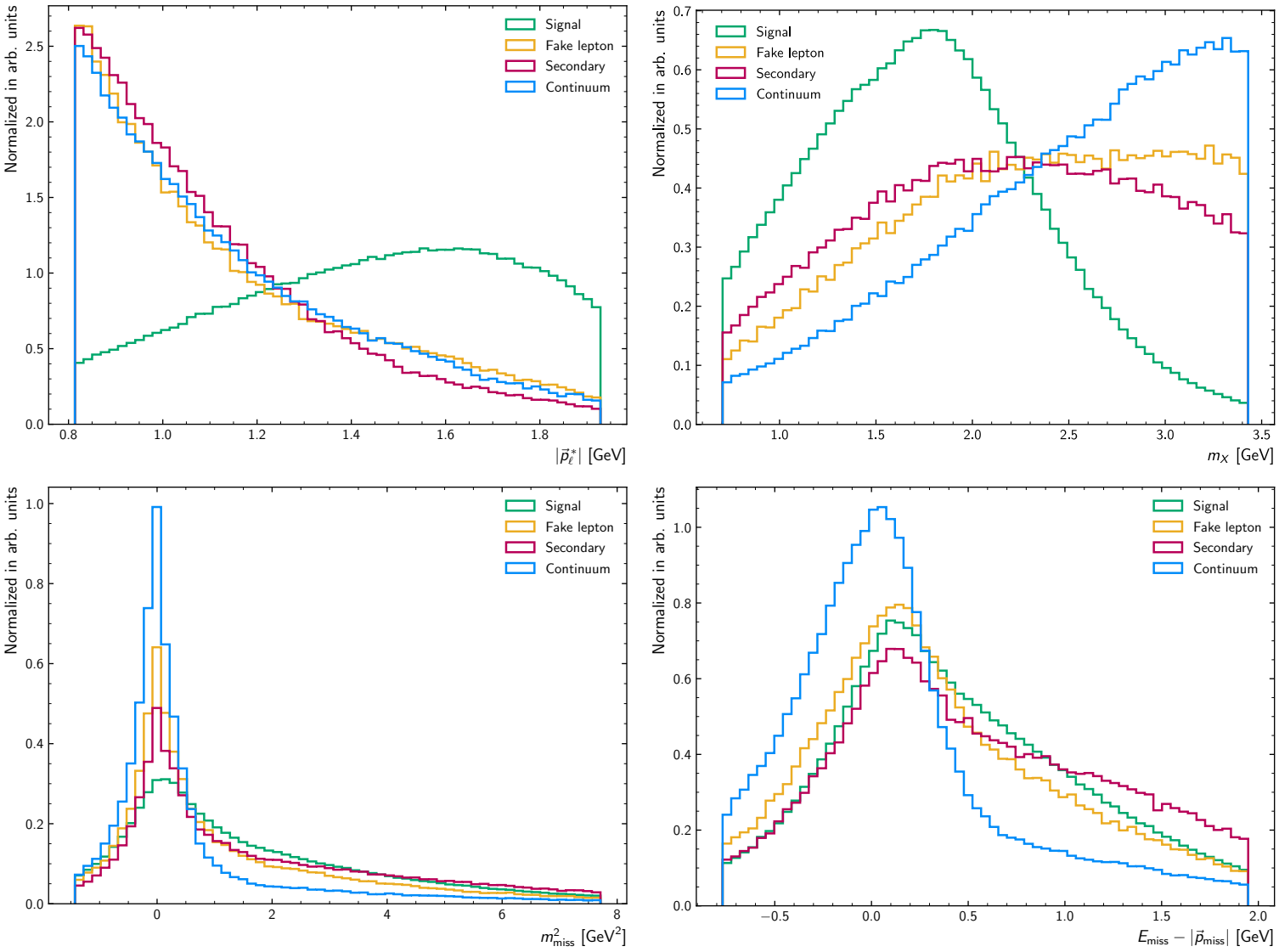


Figure 9.1: Probability density functions of possible primary fit variables that were considered.

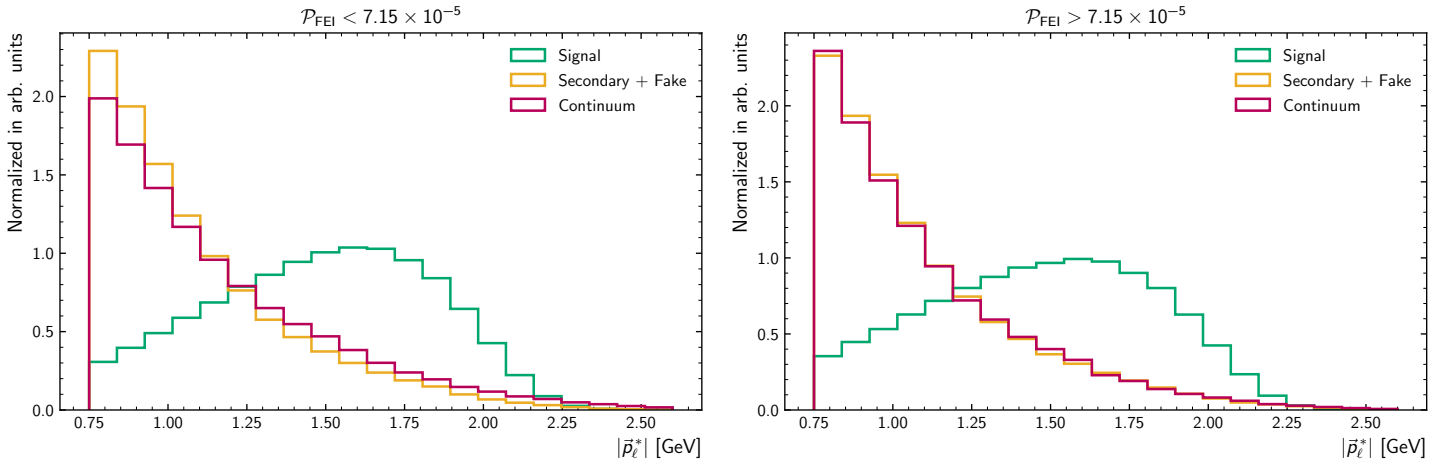


Figure 9.2: $|\vec{p}_\ell^{\,*}|$ distributions below and above the median of $\mathcal{P}_{\text{FEI}}$.

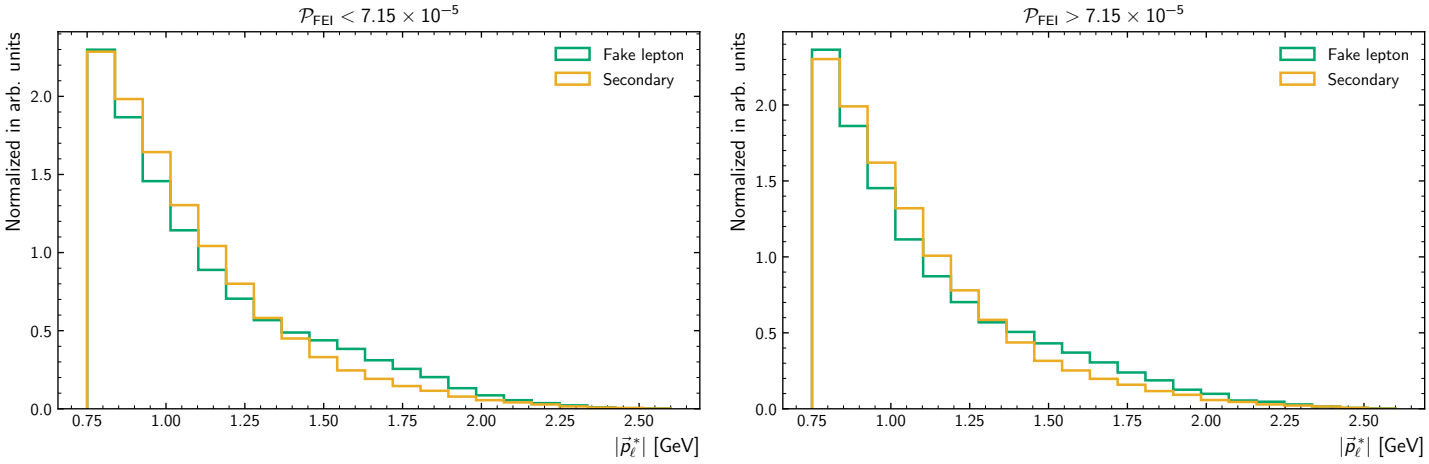


Figure 9.3: Fake and secondary $|\vec{p}_\ell^{\,*}|$ distributions below and above the median of $\mathcal{P}_{\text{FEI}}$.

on the probability density functions (PDF) that might arise from poorly normalized MC components, we assign a conservative 20 % uncertainty (*PDF composition uncertainty*) on the ratio of the normalizations of the fake lepton and secondary component.

Fallback fit One-dimensional fit to five equidistant bins in $|\vec{p}_\ell^{\,*}|$ between 1 GeV and 1.95 GeV. Two fit components are distinguished: signal and background, where background combines the fake lepton, secondary and continuum components. As discussed in Section 9.2.5, the $|\vec{p}_\ell^{\,*}|$ distributions of fake leptons, secondary lepton and continuum are relatively similar. To cover any remaining shape uncertainty that might arise from poorly normalized MC components, we assign 20 % uncertainties on each of these three normalizations.

To quantify the impact of the PDF composition uncertainties, we perform the fit once with and once without these uncertainties and calculate the uncertainty due to the PDF composition from the squared differences of uncertainties. The result[2] is shown in Figure 9.5. On average, the composition uncertainty is 0.3 % for the primary fit and 0.6 % for the secondary fit. The additional uncertainty only induces a very small change in the central values of the calibration factors ϵ_i^{sig} (Figure 9.6).

The distributions of p-values (calculated as described in Section 9.2.2) for both fit strategies (and both gap models) are shown in Figure 9.7. If only statistical uncertainties are considered, the fits in the different categories are independent, and the uniformity of the p-value distribution is a good indicator of the fit quality (and modeling of our kinematic distribution in MC). However, as we use systematic uncertainties that are approximately shared between the categories, this straightforward evaluation is no longer possible. If, for example, the real branching fraction of one of the decay channels differs from its assumed central value in MC, then the corresponding uncertainty is subject to a pull in every category, and every p-value is lowered simultaneously. Given this caveat, the p-value distributions appear convincing for both gap models and both fit setups.

[2] To be precise, the results were obtained by fitting once without any systematic uncertainties, and once with only the PDF composition uncertainty. An earlier, slightly different version of the categorization and dataset was being used.

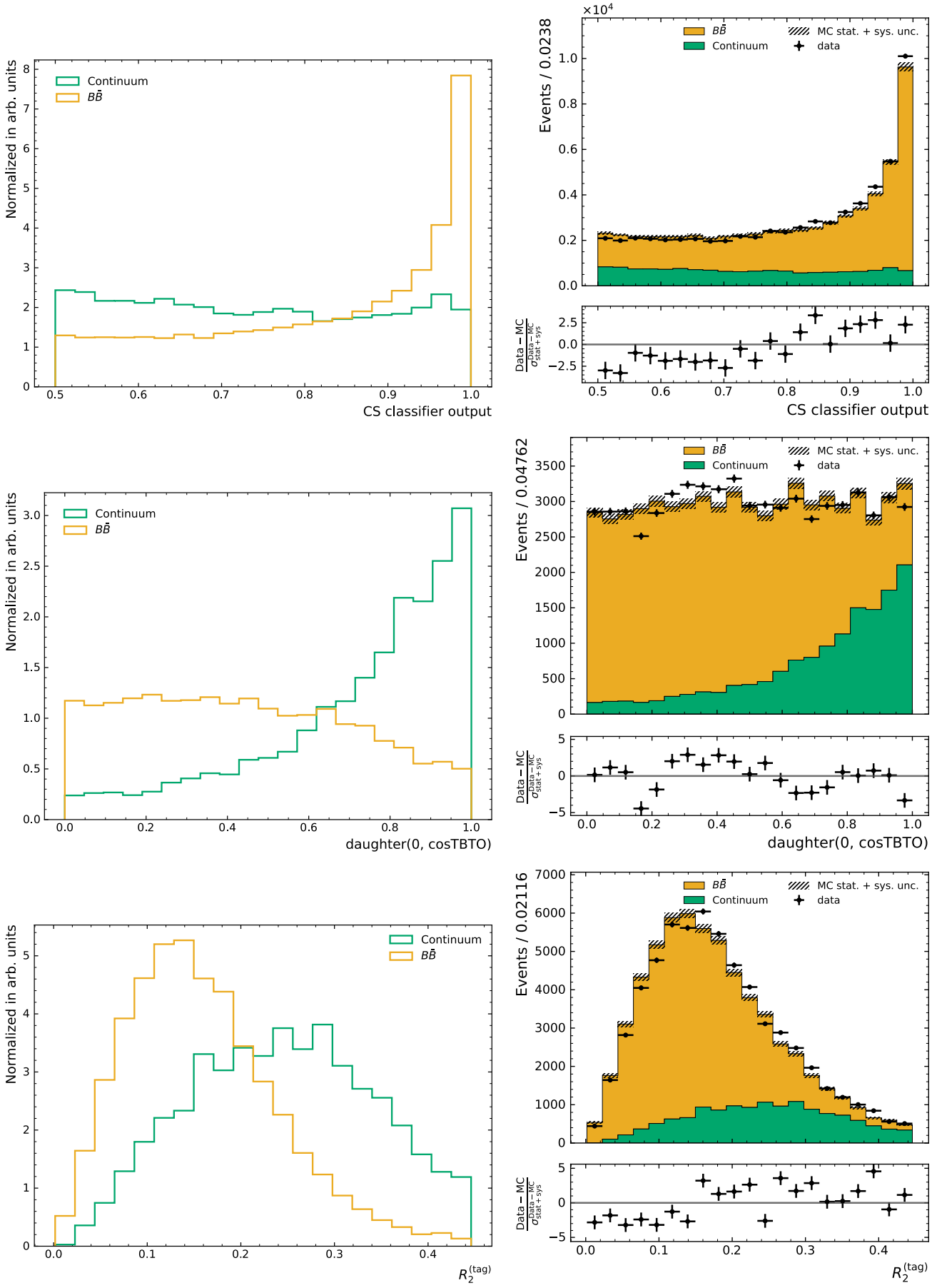
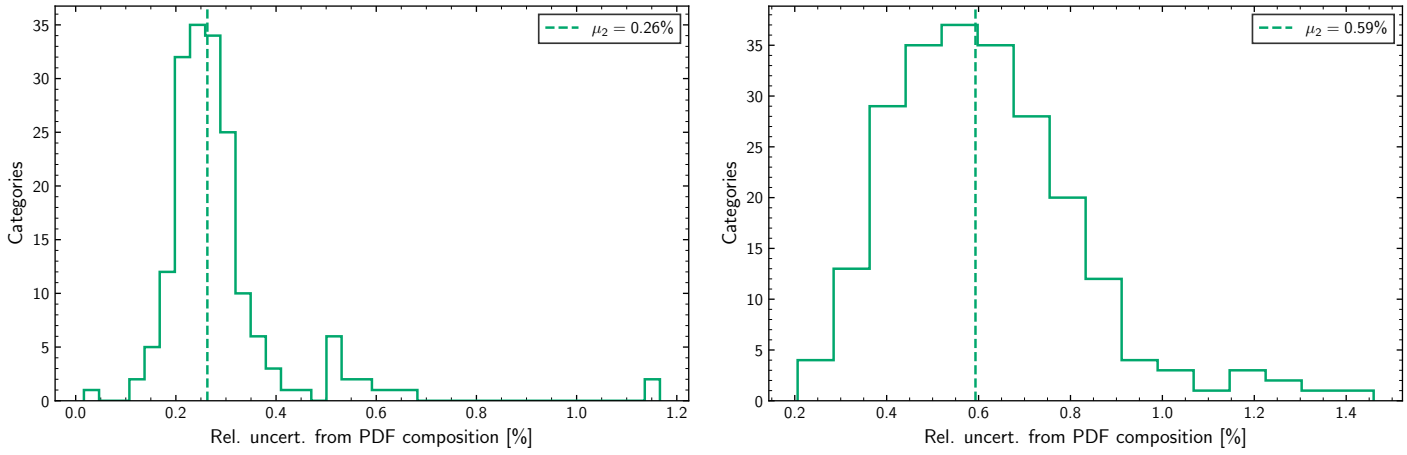
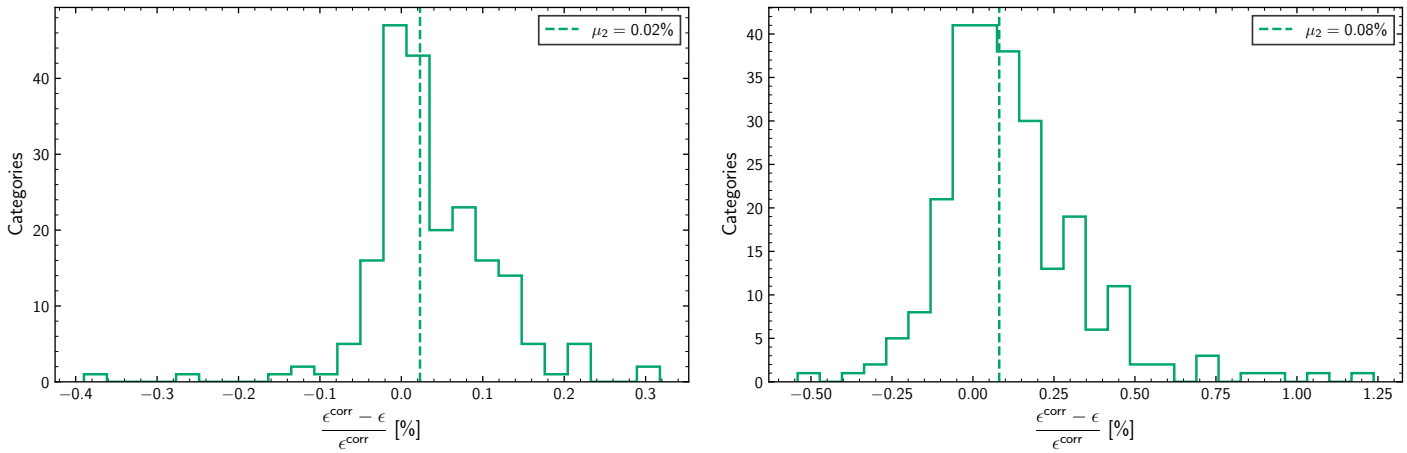


Figure 9.4: Continuum suppression variables $\mathcal{L}_{CS}$, R_2^{tag}, $\cos(\theta_{TBTO})$ after the continuum suppression cuts of Section 8.3 have been applied. The data vs. MC plots do not include calibration.



(a) Primary fit (b) Fallback fit

Figure 9.5: Uncertainty on $\epsilon_{\text{sig},i}$ from the PDF composition uncertainty. μ_2 denotes the median of the distribution.



(a) Primary fit (b) Fallback fit

Figure 9.6: Impact of the PDF composition uncertainty on the central value of ϵ_i^{sig}. ϵ^{corr} denotes the result with the additional uncertainty, ϵ without. μ_2 refers to the median of the distribution.

9.3 Calculating Calibration factors

For each fitted component c, and each category C_i, a calibration factor is calculated as

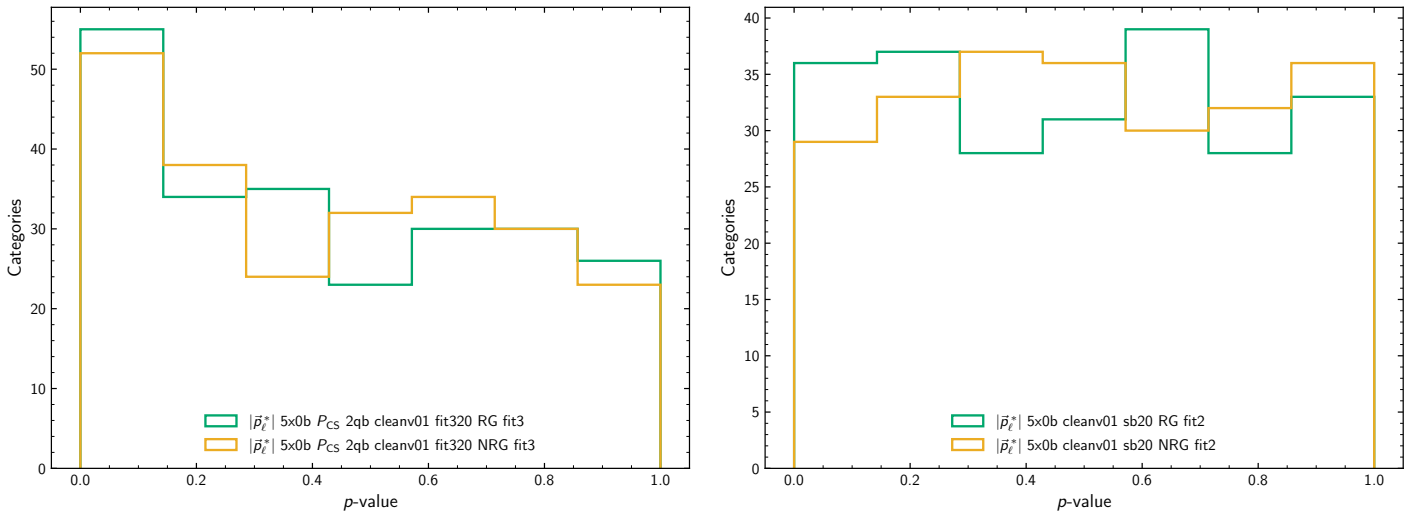
$$\epsilon_i^c = \frac{N_{c,i}^{\text{Data}}}{N_{c,i}^{\text{MC}}}, \tag{9.3.1}$$

where $N_{c,i}^{\text{Data}}$ is the fitted yield of component c for the data in C_i. The calibration factors that will be applied to the $|V_{\text{cb}}|$ dataset (as defined in Equation (6.2.1)) are those for the signal component, i.e., ϵ_i^{sig}.

9.3.1 Statistical and systematic uncertainty

The uncertainties on ϵ_i^c are calculated as

$$\frac{\tilde{\Delta}\epsilon_i^c}{\epsilon_i^c} := \frac{\Delta N_{c,i}^{\text{Data}}}{N_{c,i}^{\text{Data}}} \oplus \frac{\Delta N_{c,i}^{\text{MC}}}{N_{c,i}^{\text{MC}}}, \tag{9.3.2}$$



(a) Primary fit

(b) Fallback fit

Figure 9.7: p-value distributions for both fit strategies for the resonant gap model (RG, green) and non-resonant gap model (NRG, yellow).

where $\oplus$ denotes the usual addition of uncertainties[3]. We have placed a tilde on $\tilde{\Delta}\epsilon_i^c$ because additional uncertainties will be added later on (Section 9.5).

Equation (9.3.2) assumes that the fitted yield $N_{c,i}^{\mathrm{Data}}$ is uncorrelated to the MC yield $N_{c,i}^{\mathrm{MC}}$. This assumption is justified as follows: $N_{c,i}^{\mathrm{Data}}$ depends only on the shape of the MC templates, while the $N_{c,i}^{\mathrm{MC}}$ only depends on their normalizations. This means that systematic uncertainties mainly affecting the shape will be subleading uncertainties for $N_{c,i}^{\mathrm{MC}}$ and systematic uncertainties primarily affecting the normalization will be subleading uncertainties for $N_{c,i}^{\mathrm{Data}}$.

The uncertainty $\Delta N_{c,i}^{\mathrm{Data}}$ is directly calculated in the fit process. The systematic uncertainty on $\Delta N_{c,i}^{\mathrm{MC}}$ is calculated from

- event weights corresponding to up and down variations of independent parameters (e.g., for form factor parameters),
- event weights corresponding to parameter variations sampled according to a covariance matrix of multiple correlated parameters (e.g., for PID uncertainties).

The uncertainty on the fitted yield $\Delta N_{c,i}^{\mathrm{Data}}$ is split up into a statistical and systematic part by performing two separate fits of the same category:

- One fit with systematic uncertainties (with result $\tilde{N}_{c,i}^{\mathrm{Data}} \pm \Delta\tilde{N}_{c,i}^{\mathrm{Data}}$)
- One fit without systematic uncertainties ($\hat{N}_{c,i}^{\mathrm{Data}} \pm \Delta\hat{N}_{c,i}^{\mathrm{Data}}$).

In almost all cases $\tilde{N}_{c,i}^{\mathrm{Data}} \approx \hat{N}_{c,i}^{\mathrm{Data}}$. We then set

$$\begin{aligned} N_{c,i}^{\mathrm{Data}} &= \tilde{N}_{c,i}^{\mathrm{Data}} \\ \Delta N_{c,i}^{\mathrm{Data}} &= \max\{\Delta\tilde{N}_{c,i}^{\mathrm{Data}}, \Delta\hat{N}_{c,i}^{\mathrm{Data}}\} \\ \Delta^{\mathrm{stat}} N_{c,i}^{\mathrm{Data}} &= \Delta\hat{N}_{c,i}^{\mathrm{Data}} \\ \Delta^{\mathrm{syst}} N_{c,i}^{\mathrm{Data}} &= \Delta N_{c,i}^{\mathrm{Data}} \ominus \Delta^{\mathrm{stat}} N_{c,i}^{\mathrm{Data}}, \end{aligned} \tag{9.3.3}$$

[3] I.e., $a \oplus b := \sqrt{a^2 + b^2}$.

where $\ominus$ is defined analog to $\oplus$.[4] The max in Equation (9.3.3) ensures that the systematic uncertainty is well defined, but it is rarely necessary (this is separately highlighted in the final results, see Appendix E).

The statistical and systematic part of ϵ_i^c is then calculated using Equation (9.3.2). It should be noted that by doing so, we include the PDF modeling uncertainty (statistical uncertainty of the MC sample) in the statistical part. This is for simplicity because we will later only distinguish between uncertainties correlated or uncorrelated across categories, and the statistical uncertainty of the MC modeling belongs to the uncorrelated part. The uncertainties are finalized in Section 9.5.

It should be noted that the statistical and systematic uncertainty on ϵ_i^c due to $N_{c,i}^{\text{Data}}$ are of order 1 %, which is small in comparison to the overall uncertainty of about 4 %. Thus, the uncertainties are driven by the uncertainties on the MC yields.

9.3.1.1 Failing no-systematic fits

Note: This minor detail currently only affects the uncertainty calculation of two fit results for the resonant gap model.

In some cases, the fit with only statistical uncertainties fails, even though the fit with systematic uncertainties is successful. This can happen because, in some cases, the shape adjustment of the fit templates corresponding to the systematic uncertainties is necessary to achieve a sufficient fit quality (defined according to the definitions of Section 9.2.3).

For such cases, we assume that the relative systematic uncertainty is the same as the mean of the relative systematic uncertainties as calculated in 9.3.3.

To be precise, we consider the set

$$E_c := \left\{ \left(\frac{\Delta \tilde{N}_{c,i}^{\text{Data}}}{\tilde{N}_{c,i}^{\text{Data}}} \right)^2 - \left(\frac{\Delta \hat{N}_{c,i}^{\text{Data}}}{\hat{N}_{c,i}^{\text{Data}}} \right)^2 \,\middle|\, \text{both fits succeed} \right\}. \tag{9.3.4}$$

Under the assumption that the central values of both fits agree ($\tilde{N}_{c,u}^{\text{Data}} = \hat{N}_{c,i}^{\text{Data}}\ \forall i$), this is the set of all squared relative systematic uncertainties calculated as in (9.3.3). We then set the relative systematic uncertainty of the failed no-systematic fit case using the arithmetic mean $\mu(E_c)$ of E_c, i.e., define

$$\Delta \hat{N}_{c,i}^{\text{Data}} := \Delta \tilde{N}_{c,i}^{\text{Data}} \ominus \sqrt{\mu(E_c)} \tilde{N}_{c,i}^{\text{Data}}. \tag{9.3.5}$$

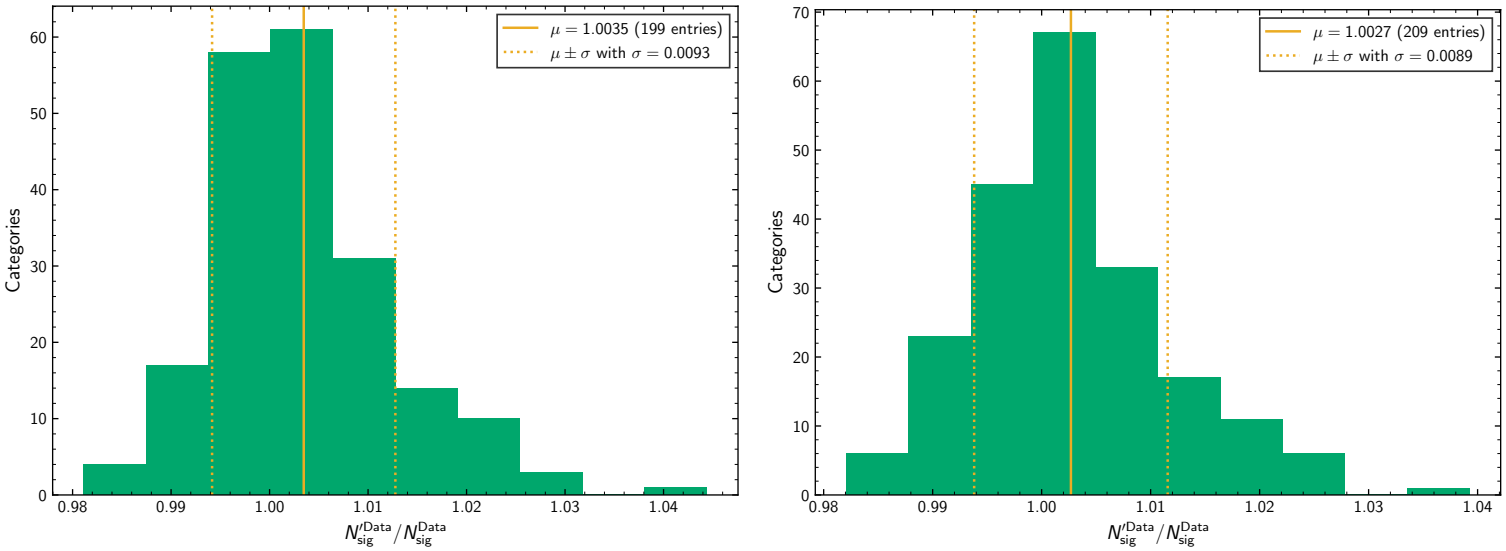
and apply all definitions from Equation (9.3.3) as before. The final results highlight the cases where this substitution is made.

9.3.2 Substituting results from simplified fits

In some categories, the default fit procedure (with systematic uncertainties) fails. This only happens for 32 (23) of 233 categories for the resonant gap model (non-resonant gap model). In this case, we use a simplified fit configuration (*fallback fit*, see Section 9.2.6). By comparing both fit configurations for categories where both fits pass, we can see that the results differ in some categories (Figure 9.9)

We use a correction factor to correct systematic shifts and assign a systematic uncertainty that should cover such differences: Let us assume that the simplified fit (with systematic uncertainties) returns the fit result $\tilde{N}_{c,i}^{\prime\text{Data}} \pm \Delta \tilde{N}_{c,i}^{\prime\text{Data}}$. The correction factor is calculated as the average of the

[4] I.e., $a \ominus b := \sqrt{a^2 - b^2}$



(a) Resonant gap

(b) Non-resonant gap

Figure 9.8: Ratios of yields fitted with the primary fit configuration and the fallback fit configuration.

ratio of fitted yields between both strategies (from all categories that can be successfully fit in both ways).

To be precise, we consider the set

$$D_c^{\text{fb}} := \left\{ \frac{\tilde{N}'^{\text{Data}}_{c,j}}{\tilde{N}^{\text{Data}}_{c,j}} \;\middle|\; j \text{ where both fits succeed on } C_j \right\}. \tag{9.3.6}$$

For any fallback fit case, we then define

$$\begin{aligned} \tilde{N}^{\text{Data}}_{c,i} &= \tilde{N}'^{\text{Data}} r_c^{\text{fb}} \\ \hat{N}^{\text{Data}}_{c,i} &= \hat{N}'^{\text{Data}} r_c^{\text{fb}} \quad \text{(if successful)} \end{aligned} \tag{9.3.7}$$

with a random variable

$$r_c^{\text{fb}} \sim \mathcal{N}\Big(\mu(D_c^{\text{fb}}), \sigma(D_c^{\text{fb}})\Big), \tag{9.3.8}$$

where $\mu(D_c^{\text{fb}})$ are the mean of D_c^{fb} and $\sigma(D_c^{\text{fb}})$ its standard deviation. Equation (9.3.3) is then applied as usual.

The sets $D_{\text{sig}}^{\text{fb}}$ are shown in Figure 9.8. We obtain

$$r_{\text{sig}}^{\text{fb}} = 1.003 \pm 0.009. \tag{9.3.9}$$

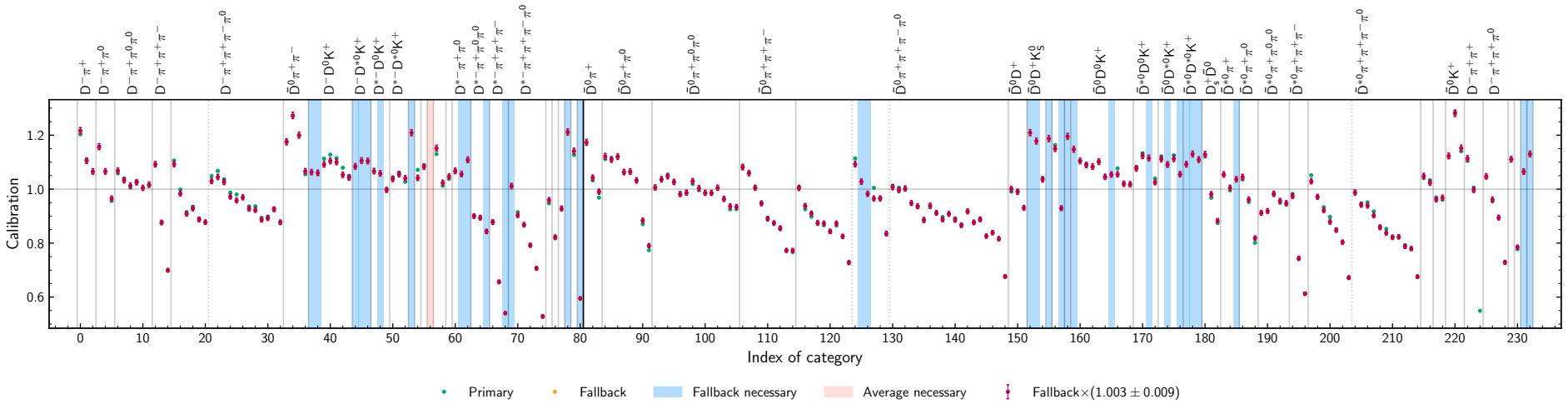
The uncertainty due to r_c^{fb} is assumed to be uncorrelated between different calibration factors.

9.3.3 Substituting average calibration factors

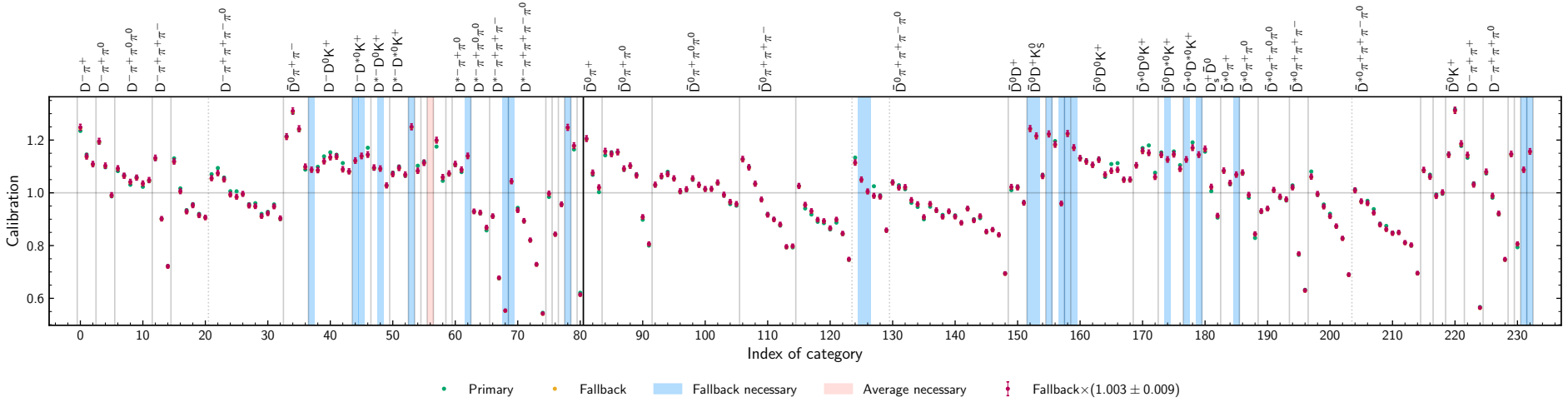
For one very small category ($\text{D}^{*-}\text{D}^{*+}\text{K}_{\text{S}}^0$, highlighted by the salmon-colored band in Figure 9.9), neither of the two fit strategies is successful (as defined in Section 9.2.3).

Here, we instead return an average calibration factor

$$\bar{\epsilon}_i := \frac{N_i^{\text{Data}}}{N_i^{\text{MC}}}, \tag{9.3.10}$$



(a) Resonant gap



(b) Non-resonant gap

Figure 9.9: Comparing calibration factors determined with the primary and fallback fit strategy. The vertical black line separates B^0 from B^+ decay modes; vertical gray lines separate different reconstruction modes; vertical dashed lines separate sub decay channels; the last sub decay channel usually is a collection of all sub decay channels not considered separately. The categories highlighted in blue are where the fallback fit is applied. Highlighted in salmon are categories where neither fit succeeds (no calibration factor is shown there). The data points obtained with the fallback fit result corrected by $r_{\text{sig}}^{\text{fb}}$ (red) hide the data points for the uncorrected results (yellow).

where N_i^{Data} and N_i^{MC} are the total numbers of events in data and MC for the category C_i.

This calibration factor is in fact the weighted average of the calibration factors defined in Equation (9.3.1):

$$\bar{\epsilon}_i = \frac{N_i^{\text{Data}}}{N_i^{\text{MC}}} = \sum_c \frac{N_{c,i}^{\text{Data}}}{N_i^{\text{MC}}} = \sum_c \epsilon_i^c \frac{N_{c,i}^{\text{MC}}}{N_i^{\text{MC}}} = \sum_c \epsilon_i^c f_c^{\text{MC}}, \tag{9.3.11}$$

where f^{MC} are the relative contributions of the different MC components ($\sum_c f_c^{\text{MC}} = 1$).

Similar to the procedure in Section 9.3.2, we use a correction factor to correct for systematic differences and to assign an uncertainty for the substitution. This correction factor is the average ratio of the calibration factor derived from the fit and the average calibration factor (for all categories where the fit is successful).

To be precise:

$$D_c^{\text{avg}} := \left\{ \frac{\epsilon_{c,j}}{\bar{\epsilon}_j} \;\middle|\; j \text{ where fit succeeds on } C_j \right\}. \tag{9.3.12}$$

We then define

$$\epsilon_{c,i} := \bar{\epsilon}_i \, r_c^{\text{avg}} \tag{9.3.13}$$

with a random variable

$$r_c^{\text{avg}} \sim \mathcal{N}\Big(\mu(D_c^{\text{avg}}), \sigma(D_c^{\text{avg}})\Big), \tag{9.3.14}$$

where $\mu(D_c^{\text{avg}})$ are the mean of D_c^{avg} and $\sigma(D_c^{\text{avg}})$ its standard deviation.

The uncertainty on $\bar{\epsilon}_i$ is calculated analog to Equation (9.3.2), but the calculation is simpler because only statistical uncertainties need to be considered for N_i^{Data}. The split-up of systematic and statistical uncertainty on $\bar{\epsilon}_i$ poses no problems, because no fit is involved.

9.4 The different gap models

As described in Section 7.3.4, two different gap models are considered (resonant and non-resonant). To arrive at one final result, we average the calibration factors that are separately calculated for the two models.

Denoting the dataset with the resonant (non-resonant) gap model as *RG* (*NRG*), two sets of calibration factors are calculated as described in the previous sections:

$$\epsilon_i^{c\,(\text{RG})} := \frac{N_{c,i}^{\text{Data (RG)}}}{N_{c,i}^{\text{MC (RG)}}}, \qquad \epsilon_i^{c\,(\text{NRG})} := \frac{N_{c,i}^{\text{Data (NRG)}}}{N_{c,i}^{\text{MC (NRG)}}}. \tag{9.4.1}$$

The two sets of calibration factors are shown in Figure 9.10a. A consistent shift of 2.8 % can be observed. This is mostly the result of shifts in MC yields (Figure 9.10d). However, the gap model also has a minor impact on the fitted yields (Figure 9.10c). This was already discussed to some extent in Section 7.3.4.

The final calibration factor is defined as

$$\epsilon_i^c := \frac{\epsilon_i^{c\,(\text{RG})} + \epsilon_i^{c\,(\text{NRG})}}{2}. \tag{9.4.2}$$

For the uncertainty, we define

$$\begin{aligned} \tilde{\Delta}^{\text{stat}}\epsilon_i^c &:= \max\left\{ \tilde{\Delta}^{\text{stat}}\epsilon_i^{c\,(\text{RG})}, \tilde{\Delta}^{\text{stat}}\epsilon_i^{c\,(\text{NRG})} \right\}, \\ \tilde{\Delta}^{\text{syst}}\epsilon_i^c &:= \max\left\{ \tilde{\Delta}^{\text{syst}}\epsilon_i^{c\,(\text{RG})}, \tilde{\Delta}^{\text{syst}}\epsilon_i^{c\,(\text{NRG})} \right\} \oplus \frac{\epsilon_i^{c\,(\text{RG})} - \epsilon_i^{c\,(\text{NRG})}}{2}, \end{aligned} \tag{9.4.3}$$

where $\oplus$ denotes the usual addition of uncertainties and uncertainties on $\epsilon_i^{c\,(\mathrm{RG/NRG})}$ are defined as described in Section 9.3.1.

9.5 Additional uncertainties

Rather than distinguishing between statistical and systematic uncertainties, we will slightly generalize this division and distinguish between uncertainties that are correlated across categories and those that are not. This is the only meaningful distinction for applying the uncertainties in the application of the calibration.

- All uncertainties so far summarized in the statistical part are uncorrelated between categories.
- We assume that all uncertainties that were summarized in the systematic part are fully correlated between categories: This assumption is warranted because all uncertainties added so far only depend mostly on signal side properties and thus should affect all categories similarly.

We now add some systematic uncertainties to arrive at the final uncertainties assigned to the calibration factors:

$$\Delta^{\mathrm{corr}}\epsilon_i^c := \tilde{\Delta}^{\mathrm{stat}}\epsilon_i^c \oplus e_{\mathrm{track}}^{\mathrm{corr}} \oplus e_{\mathrm{SM}}^{\mathrm{corr}} \oplus e_{\mathrm{cat}}^{\mathrm{corr}}\,, \tag{9.5.1}$$

$$\Delta^{\mathrm{uncorr}}\epsilon_i^c := \tilde{\Delta}^{\mathrm{syst}}\epsilon_i^c \oplus e_{\mathrm{SM}}^{\mathrm{uncorr}}, \tag{9.5.2}$$

where the additional terms are introduced in the following sections.

9.5.1 Tracking uncertainty

Not every track left by a charged particle is reconstructed, meaning that there is a non-unit *tracking efficiency* $\epsilon_{\mathrm{track}}$. Because we only reconstruct one track on the signal side, we have

$$\begin{aligned}\epsilon_i^c|_{\text{no track missed}} &= \frac{N_{c,i}^{\mathrm{Data}}|_{\text{no track missed}}}{N_{c,i}^{\mathrm{MC}}|_{\text{no track missed}}} = \frac{(\epsilon_{\mathrm{track}}^{\mathrm{Data}})^{-1}\cdot N_{c,i}^{\mathrm{Data}}}{(\epsilon_{\mathrm{track}}^{\mathrm{MC}})^{-1}\cdot N_{c,i}^{\mathrm{MC}}} = \\ &= \frac{\epsilon_{\mathrm{track}}^{\mathrm{MC}}}{\epsilon_{\mathrm{track}}^{\mathrm{Data}}}\cdot\epsilon_i^c =: r_{\mathrm{track}}^{-1}\cdot\epsilon_i^c\end{aligned} \tag{9.5.3}$$

[345] calculates r_{track} for tracks with $p_T > 200\,\mathrm{MeV}$ as

$$r_{\mathrm{track}} - 1 = (-0.13 \pm 0.30_{\mathrm{stat}} \pm 0.10_{\mathrm{syst}})\%. \tag{9.5.4}$$

Because the central value is minimal in comparison to the uncertainties, [345] recommends to apply this result as

$$r_{\mathrm{track}} = (100 \pm 0.35)\%, \tag{9.5.5}$$

i.e., to include the central value into the uncertainty rather than use it as a correction factor. Clearly, this uncertainty affects the calibration factors for all categories equally. Thus we prescribe

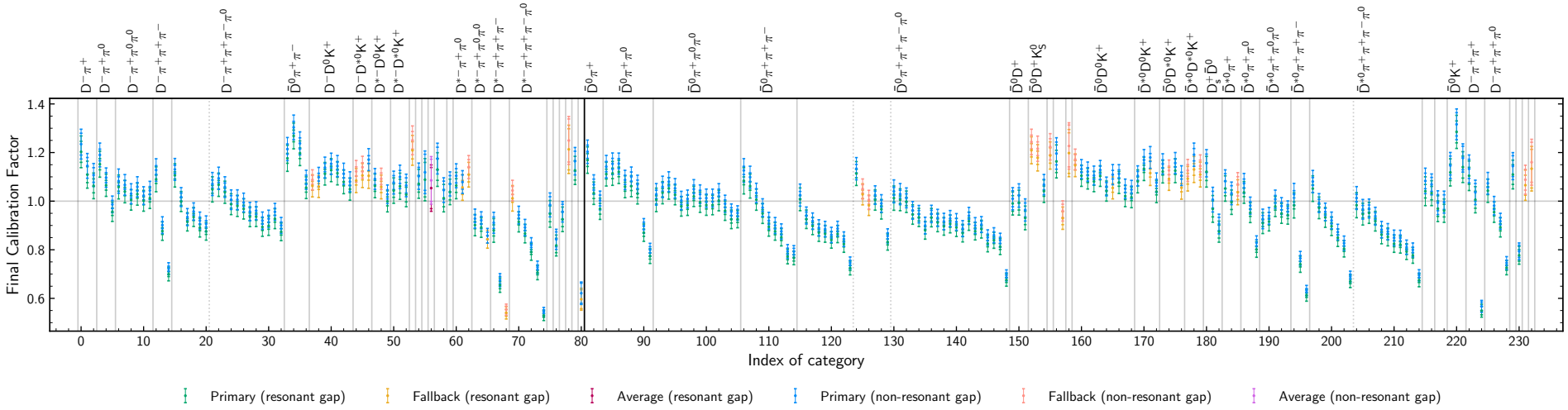
$$e_{\mathrm{track}}^{\mathrm{corr}} = 0.35\,\%.$$

9.5.2 Signal track multiplicity dependency

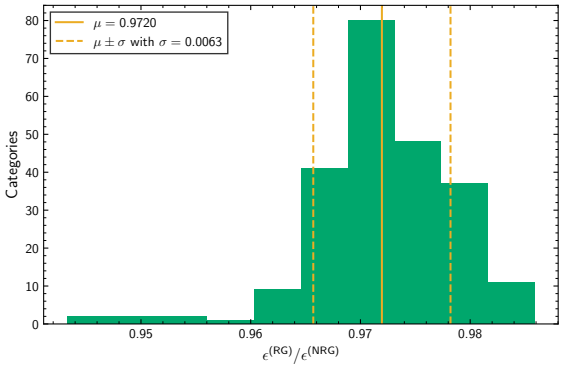
The side-study presented in Appendix B.2 assigns

$$e_{\mathrm{STM}}^{\mathrm{corr}} = 0.35\,\% \qquad \text{and} \qquad e_{\mathrm{STM}}^{\mathrm{uncorr}} = 0.4\,\%$$

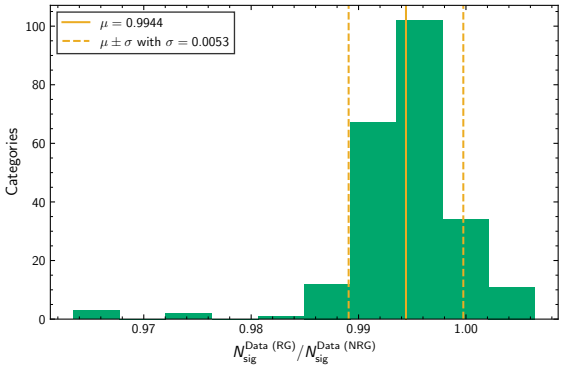
for possible influences of the signal side reconstruction on the calibration factor.



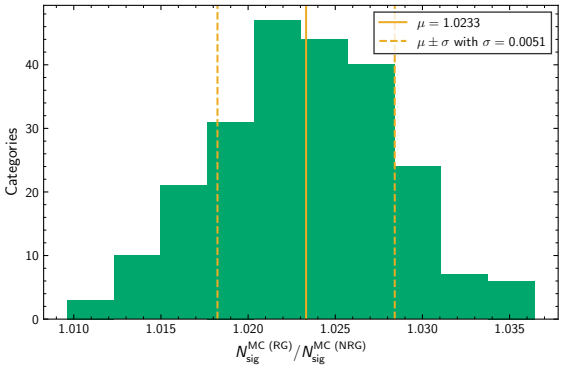
(a) Calibration factors vs. category before taking the average over both gap models. See Figure 9.9 for explanation of the general plot layout. Statistical and full uncertainty are shown separately as two sets of whiskers for each error bar.



(b) Ratios of calibration factors.



(c) Ratios of fitted yields.



(d) Ratios of MC normalizations.

Figure 9.10: Comparing the results for the resonant gap and non-resonant gap model.

9.5.3 Categorization

The side-study presented in Appendix B.1 assigns an uncertainty of

$$e_{\text{cat}}^{\text{corr}} = \begin{cases} 0.5\,\% & \text{B}^0 \\ 0.6\,\% & \text{B}^+. \end{cases}$$

for differences in the average calibration factors when choosing different categorizations.

Chapter 10

Summary and Results

The hadronically tagged $B \to X\ell\nu_\ell$ decay is reconstructed, and the events are projected into 232 categories according to the FEI reconstruction mode and classifier output (Section 9.1). Fits to $|\vec{p}_\ell^{\,*}|$ and $\mathcal{L}_{\text{CS}}$ are performed to subtract contributions from continuum events, and events with fake or secondary leptons (Section 9.2.6). In very few categories a simplified fit setup is used instead (Section 9.2.6, Section 9.3.2). As we use two different gap models (Section 7.3.4), two sets of calibration factors are calculated and subsequently averaged with an additional uncertainty covering their difference (Section 9.4). Finally, several side-studies probe underlying assumptions and assign additional uncertainties (Appendix B).

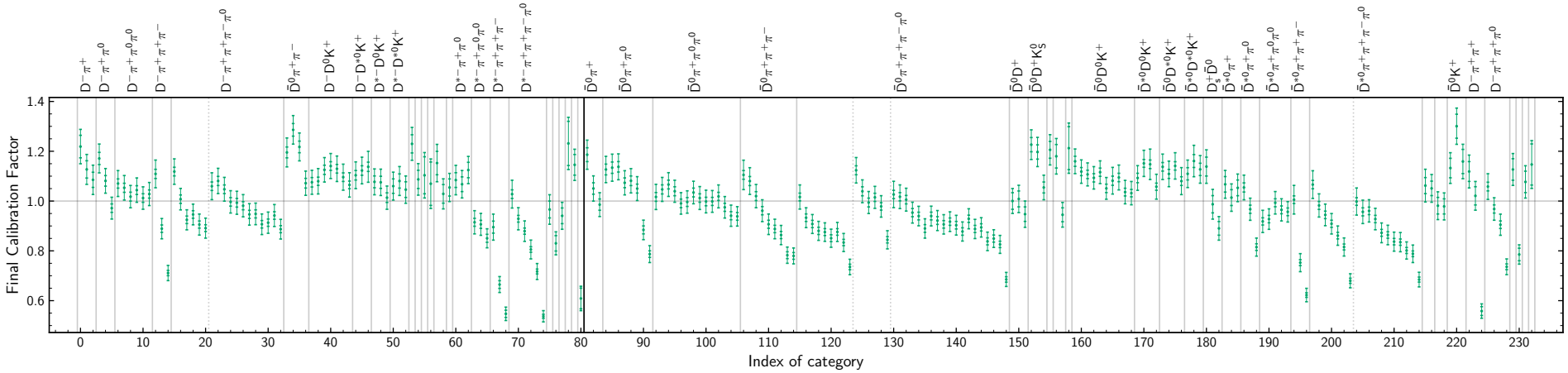
The final results are presented in Figure 10.1a. As shown in Figures 10.1b and 10.1c, the calibration factors depend significantly on both the reconstruction mode and the FEI classifier output $\mathcal{P}_{\text{FEI}}$. While the dependency on $\mathcal{P}_{\text{FEI}}$ is slightly different for each reconstruction mode, the calibration factors generally decrease with increasing $\mathcal{P}_{\text{FEI}}$.

A table of numerical results and annotations is found in Appendix E. Table 10.1 shows weighted averages of the calibration factors for different applications.

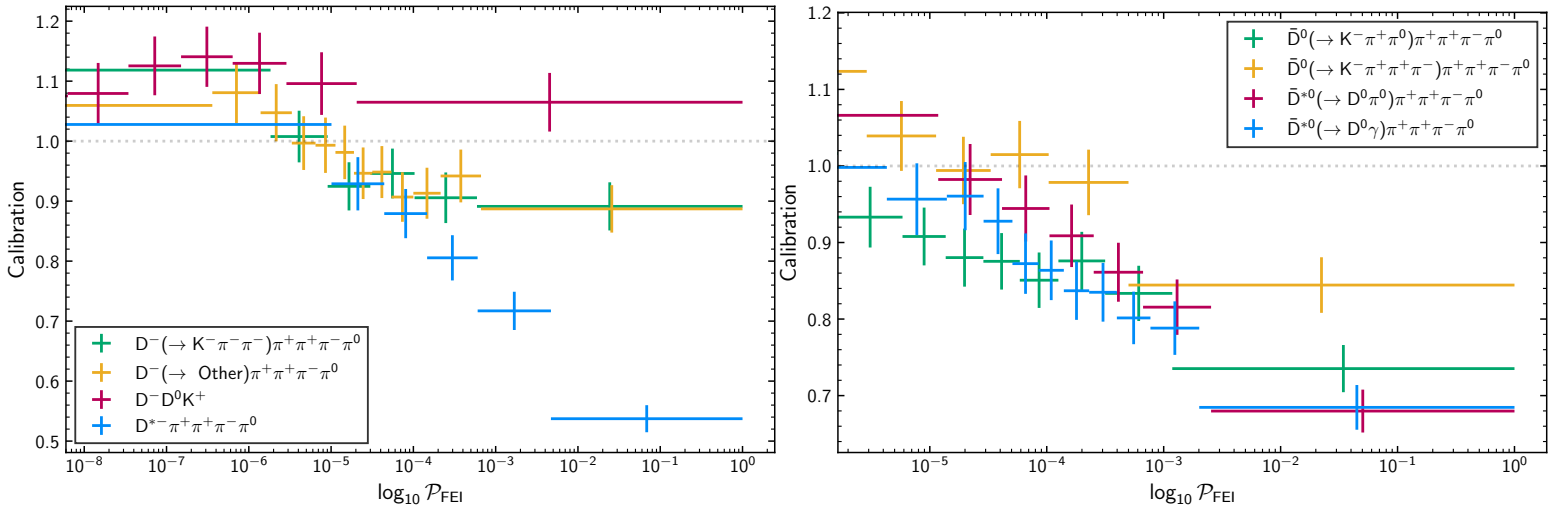
The final uncertainties are summarized in Table 10.3. The median of the total uncertainty is $(1.6_{\text{uncorr}} \oplus 4.3_{\text{corr}})\%$, where the first uncertainty is uncorrelated between calibration factors for different categories and the second is fully correlated. When applied to the $B \to D^*\ell\nu_\ell$ dataset for the $|V_{\text{cb}}|$ measurement, the uncorrelated uncertainties become negligible, and the final uncertainty due to the tagging calibration in the $|V_{\text{cb}}|$ analysis is 4.3 %. The most significant contribution to the uncertainties is the 3.1 % uncertainty on the branching fractions of each gap model, in particular of the resonant gap model (Table 10.3b). In particular, the results are strongly dominated by systematic uncertainties on the MC normalization, while the uncertainty on the fitted yield is only $(1.4_{\text{stat}} \oplus 1.5_{\text{syst}})\%$.

Several closure tests confirm the validity of the background subtraction and the general calibration procedure. They are presented in Appendix C and include studies with toys, the comparison of separate calibrations with $B \to Xe^-\nu_e$ and $B \to X\mu^-\nu_\mu$, studies of signal side influence on the categories, and the application of the calibration factors to $B \to X\ell\nu_\ell$.

Furthermore, we compare the calibration factors with calibration results derived from the $B \to D^{(*)}\ell\nu_\ell$ dataset. For example, the weighted average from Table 10.1 can be compared with the result of a calibration fit on $B \to D\ell\nu_\ell$. This is shown in Table 10.2 demonstrating the transferability of the calibration factors to the exclusive datasets. Performing a similar comparison in bins of $\mathcal{P}_{\text{FEI}}$ in Figure 10.2, we can also confirm that the $\mathcal{P}_{\text{FEI}}$ dependency of the calibration factors does not depend on the dataset. Studies of this kind are presented in detail in Appendix C.1.



(a) Final calibration factors. The uncorrelated part of the uncertainty and the full uncertainty are shown separately as two sets of whiskers. The black vertical line divides B^0 channels and B^+ channels; vertical gray lines divide different FEI reconstruction modes; dashed vertical lines divide sub-reconstruction modes; the last sub decay channel usually is a collection of all sub decay channels not considered separately. A full table of all results is available in Appendix E.



(b) Mixed modes

(c) Charged modes

Figure 10.1: Final calibration factors vs. $\mathcal{P}_{\mathrm{FEI}}$ for different decay channels.

Weights	B^0	B^+	Both
Exclusive $B \to D\ell\nu_\ell$	0.94 ± 0.04	0.93 ± 0.04	0.93 ± 0.04
Exclusive $B \to D\ell\nu_\ell$ signal ℓ	0.93 ± 0.04	0.92 ± 0.04	0.92 ± 0.04
Exclusive $B \to D\ell\nu_\ell$ signal	0.91 ± 0.04	0.91 ± 0.04	0.91 ± 0.04
Exclusive $B \to D^*\ell\nu_\ell$	0.92 ± 0.04	0.91 ± 0.04	0.91 ± 0.04
Exclusive $B \to D^*\ell\nu_\ell$ signal ℓ	0.92 ± 0.04	0.90 ± 0.04	0.91 ± 0.04
Exclusive $B \to D^*\ell\nu_\ell$ signal	0.90 ± 0.04	0.90 ± 0.04	0.90 ± 0.04
Inclusive $X\ell\nu_\ell$ (RG)	0.98 ± 0.04	0.95 ± 0.04	0.96 ± 0.04
Inclusive $X\ell\nu_\ell$ signal ℓ (RG)	0.98 ± 0.04	0.95 ± 0.04	0.96 ± 0.04

Table 10.1: Averages of calibration factors weighted by the number of events per category for different datasets. Signal ℓ means that we use the signal definition from Section 9.2.4 (correctly reconstructed primary lepton from a $B\overline{B}$ event), *signal* means correctly reconstructed $B \to D^{(*)}\ell\nu_\ell$ decays (including the $D^{(*)}$ reconstruction). The averages based on the inclusive dataset use the event counts of $B \to X\ell\nu_\ell$ with the resonant gap model (RG). The event counts are visualized in Figure D.9.

	B^0	B^+	Both
$D\ell\nu$ calibration fit	0.90 ± 0.03	0.92 ± 0.03	0.91 ± 0.03
Incl. calib. rwgt. to $D\ell\nu$ signal ℓ	0.93 ± 0.04	0.92 ± 0.04	0.92 ± 0.04
Tension	-0.53	-0.08	-0.21

Table 10.2: Comparing weighted averages of the calibration factors (second line) to results obtained by applying the calibration fits to the $B \to D\ell\nu_\ell$ dataset (first line). For the uncertainty of the latter, 3 % of systematic uncertainty for the branching ratio of $B \to D\ell\nu_\ell$ have been added to the statistical uncertainty. The *tension* is the difference between both results expressed in units of its uncertainty. The B flavor is defined via the tag side. This comparison is explained in detail in Appendix C.1.1.

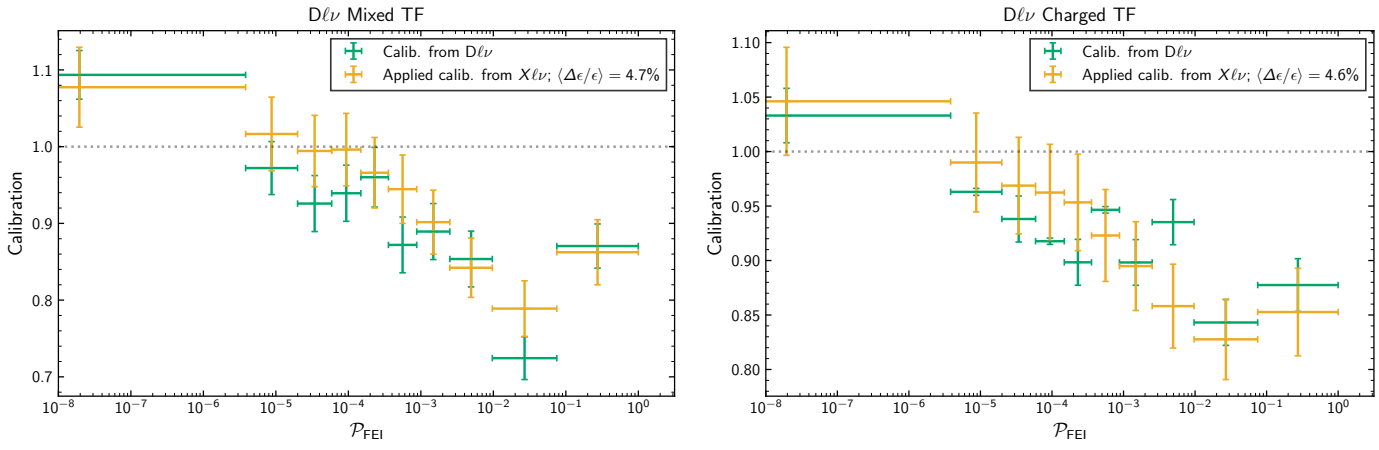


Figure 10.2: Comparing the weighted averages of the calibration factors with results obtained by applying the calibration fits to the $B \to D\ell\nu_\ell$ dataset. Only statistical uncertainties are shown for the latter. The B flavor is defined via the tag side. Additional plots of similar comparisons in the light of recent findings are shown in Figure 11.1.

%	Source
0.04	FF $B \to D$
0.07	FF $B \to D^*$
0.07	FF $B \to D_2$
0.08	$N_{\text{track}}^{\text{tag}}$ dependency (uncorr.)
0.10	FF $B \to D'^1$
0.11	FF $B \to D^{*0}$
0.13	FF $B \to D_1$
0.20	$N_{\text{track}}^{\text{tag}}$ dependency (corr.)
0.26	BR $B \to D_2^* \ell\nu_\ell$
0.27	BR $B \to D_1 \ell\nu_\ell$
0.35	$N_{\text{track}}^{\text{sig}}$ dependency (corr.)
0.35	Lepton track (corr.)
0.40	$N_{\text{track}}^{\text{sig}}$ dependency (uncorr.)
0.51	Muon efficiency
0.55	Categorization
0.57	BR $B \to D\ell\nu$
0.58	BR $B \to D_0^* \ell\nu_\ell$
0.72	MC norm. stat. uncert.
0.77	BR $B \to D_1' \ell\nu_\ell$
0.77	Electron efficiency
1.36	Different gap models
1.36	BR $B \to D^* \ell\nu$
1.43	Fitter (stat)
1.50	Fitter (syst)
3.09	Gap model BRs
4.59	Total

(a) Final uncertainties.

%	Source
0.2	BR Gap $B \to D_{0s}(\to D\pi\pi)\ell\nu_\ell$
0.3	BR Gap $B \to D_1'(\to D\pi\pi)\ell\nu_\ell$
0.7	BR Gap $B \to D_{0s}(\to D^*\pi\pi)\ell\nu_\ell$
0.8	BR Gap $B \to D_1'(\to D^*\pi\pi)\ell\nu_\ell$
1.0	BR $B \to D_1(\to D\pi\pi)\ell\nu_\ell$
1.9	BR Gap $B \to D_{0s}(\to D\eta)\ell\nu_\ell$
2.0	BR Gap $B \to D_1'(\to D^*\eta)\ell\nu_\ell$
3.1	Total

(b) BR uncertainties in the resonant gap model.

%	Source
0.3	BR Gap $B \to D\pi\pi\ell\nu_\ell$
0.3	BR Gap $B \to D^*\pi\pi\ell\nu_\ell$
1.0	BR $B \to D_1(\to D\pi\pi)\ell\nu_\ell$
1.2	BR Gap $B \to D^{*0}\eta\ell\nu_\ell$
1.5	BR Gap $B \to D\eta\ell\nu_\ell$
2.2	Total

(c) BR uncertainties in the non-resonant gap model.

Table 10.3: Final uncertainties on the calibration factors (medians). For each category, the maximum of each source of uncertainty between the two gap models has been taken. Because of the different components involved, the uncertainty on the gap branching fractions is summarized in one component in Table 10.3a and then separately split up for the two gap models in Tables 10.3b

Chapter 11

Recent findings

While Table 10.2 and Figure 10.2 show closure with the exclusive datasets, there is one subtlety: The tests define the reconstructed flavor of the B meson pair based on the tag side reconstructed by the FEI. This is the natural and only possible definition for the tagging calibration since the B meson on signal side is not fully reconstructed.

However, this definition matters! In Figure 11.1, we compare the calibration factors calculated by applying the calibration fits directly to $B \to D^{(*)}\ell\nu_\ell$ with the appropriate weighted average of the calibration factors from $B \to X\ell\nu_\ell$. On the one hand, the results agree well when the flavor is defined based on the tag side. On the other hand, the two sets of calibration factors are consistently and significantly different when the flavor is defined via the signal side. In this case, the calibration derived from $B \to X\ell\nu_\ell$ appears to be too high for mixed B mesons.

The fact that the calibration factors are different between the two definitions of the event flavor is to be expected: Neither the flavor reported by the FEI for the B_{tag} nor the reconstructed flavor of the B_{sig} is always correct. Consequently, the two reconstructed flavors do not always agree. Thus, the averages of the calibration factors for a specified signal side or tag side flavor correspond to different subsets of events and do not need to agree.

This fact alone cannot cause tension with the calibration factors from $B \to X\ell\nu_\ell$ when considering the reconstructed signal side flavor. If we assume that the signal side efficiencies are well understood, the most likely explanation is that the required calibration depends on the correct reconstruction of the B_{tag} flavor. For example, the probability of reconstructing a tag with the correct (incorrect) flavor might be lower (higher) on data than on MC since the FEI algorithm was trained using MC samples. This bias would cause the calibration factors for correctly and incorrectly reconstructed tag flavors to diverge.

There are three different strategies to account for this effect:

1. Can we consider a subset of the dataset where the effect of incorrectly reconstructed tags is less prominent? For example, if we find selection criteria based on on $\mathcal{P}_{\text{FEI}}$ or C_{FEI}, such that the incorrect tag flavor (ITF) fraction is similar between the inclusive and the exclusive dataset, the previously measured calibration factors are applicable without correction. We discuss this option in Section 11.1 and show that the strategy does not provide a sufficient solution.

2. Can we use information from $B \to D\ell\nu_\ell$ to correct the calibration factors derived from $B \to X\ell\nu_\ell$? While the $B \to D\ell\nu_\ell$ sample is too small to derive calibration factors of appropriate granularity, the additional information obtained from reconstructing the signal side can be exploited to investigate and correct the $B \to X\ell\nu_\ell$ calibration factors. Section 11.2 finds that the calibration factors for correctly and incorrectly reconstructed tag

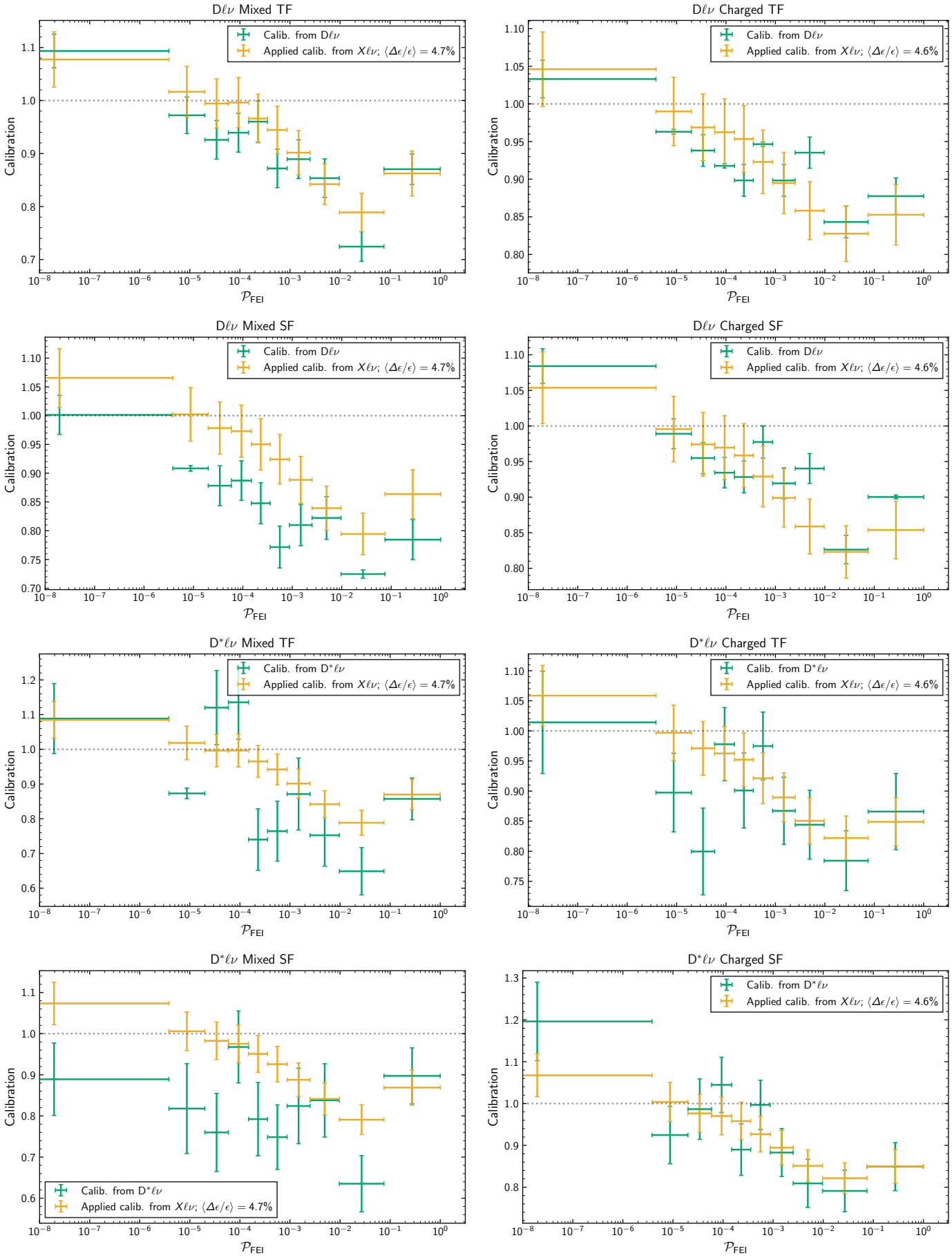


Figure 11.1: Calibration factors fitted on $B \to D\ell\nu_\ell$ vs. the applied calibration transferred from $B \to X\ell\nu_\ell$ for different definitions of B meson flavor. *TF* (*SF*) denotes that the flavor is defined via the tag (signal) side. $\langle \Delta\epsilon/\epsilon \rangle$ refers to the average relative uncertainty on the calibration.

flavors differ by a constant factor. After applying this factor to the $B \to X\ell\nu_\ell$ calibration factors, they can be used to calibrate $B \to D^*\ell\nu_\ell$.

3. Can we extend the calibration setup to calibrate correctly and incorrectly reconstructed tag flavors separately? Unfortunately, both cases cannot be easily distinguished using the previously discussed signal side variables, such as $|\vec{p}^*_\ell|$. There is also not a wide selection of available tag side variables: Most variables that possess enough separation power are already included in the decision of the FEI and cannot be used for the calibration. A possible avenue is the variable $m_{\text{bc}}^{\text{tag}}$, which we explore in Section 11.3. However, only limited investigations can be performed due to the current analysis selection requirement of $m_{\text{bc}}^{\text{tag}} > 5.27\,\text{GeV}$. A dataset extending to lower values of $m_{\text{bc}}^{\text{tag}}$ is in preparation but not yet available.

As a general note, all results that require fits to the exclusive datasets shown in this chapter (like the ones shown in Figure 11.1) were obtained with the same fit setup as used for extracting calibration factors on $B \to X\ell\nu_\ell$. In particular, we only fit for signal leptons, and $B \to D\ell\nu_\ell$ and $B \to D^*\ell\nu_\ell$ crossfeed is not separated. Therefore, the $B \to D\ell\nu_\ell$ and $B \to D^*\ell\nu_\ell$ datasets have significant event overlap. Furthermore, only statistical uncertainties are considered in the fits to the exclusive dataset. Possible caveats of this are discussed in Section 11.3.3.

We have also assigned 10^{-10} as lowest value of $\mathcal{P}_{\text{FEI}}$ for presentational purposes (i.e., for all $\mathcal{P}_{\text{FEI}} < 10^{-10}$ we replaced $\mathcal{P}_{\text{FEI}}$ with 10^{-10}).

11.1 Counting flavor configurations

Up to three different sources of flavor information are available in any event: reconstructed tag flavor, reconstructed signal flavor (available for exclusive data only), and true flavor (available on MC only). We can therefore distinguish between eight different flavor configurations.

In the following, we use the abbreviations:

TF (*tag flavor*): reconstructed B flavor defined via the tag side,

SF (*signal flavor*): reconstructed B flavor defined via the signal side,

IF (*inconsistent flavor*; excl. only): TF $\neq$ SF,

CF (*consistent flavor*; excl. only): TF = SF,

ITF (*incorrect tag flavor*; MC only): TF does not match true B flavor,

CTF (*correct tag flavor*; MC only): TF matches true B flavor,

ISF (*incorrect signal flavor*; excl. MC only): SF does not match true B flavor,

CSF (*correct signal flavor*; excl. MC only): SF matches true B flavor.

For example, the notation *IF Mixed (TF)* means that we reconstructed a B^0 on the tag side and a B^+ on the signal side.

To estimate the magnitude of possible effects, we compile a series of plots to determine the number of occurrences of each case as well as the dependency on $\mathcal{P}_{\text{FEI}}$.

11.1.1 Inconsistent flavor

The fraction of events with inconsistent tag and signal flavors, f_{IF}, is only available for the exclusive dataset. Figures 11.2a and 11.2b show f_{IF} for data and MC. A value of $f_{\text{ITF}} = 50\,\%$ indicates that signal and tag flavor coincide purely by chance, thus we expect probabilities below this value. We recognize clear differences between the B flavors and different datasets. The worst performance is observed for $\text{B} \to \text{D}\ell\nu_\ell$ Mixed TF, where even at high $\mathcal{P}_{\text{FEI}}$, 40 % of events have inconsistently reconstructed flavor.

Figures 11.2c to 11.2f show the ratios $f_{\text{IF}}^{\text{Data}}/f_{\text{IF}}^{\text{MC}}$ and $f_{\text{CF}}^{\text{Data}}/f_{\text{CF}}^{\text{MC}} = (1 - f_{\text{IF}}^{\text{Data}})/(1 - f_{\text{IF}}^{\text{MC}})$. Note that in this fraction, $f_{\text{IF}}^{\text{MC}}$ and $f_{\text{CF}}^{\text{MC}}$ are calculated without matching signal leptons to allow for a simple cut-and-count comparison with data. Since f_{IF} and f_{CF} are already relative to data, the usual calibration factors (based on tag flavor) cancel. Thus, the fact that these double ratios are different from unity is already a clear sign of a calibration effect not covered by the calibration factors derived so far. However, in order to obtain a thorough comparison and to avoid bias from insufficiently modeled background leptons, we need to fit the signal lepton component in data . This is done in Section 11.2.1.

11.1.2 Incorrect flavor

In contrast to flavor consistency, the necessary information to investigate the truth-matched flavors is only available on MC. The fraction f_{ITF} (Figure 11.3) differs significantly between the inclusive and exclusive datasets with the ratio $f_{\text{ITF}}^{\text{excl}}/f_{\text{ITF}}^{\text{incl}}$ generally decreasing with increasing $\mathcal{P}_{\text{FEI}}$ (Figure 11.4). However, since both $f_{\text{ITF}}^{\text{excl}}$ and $f_{\text{ITF}}^{\text{incl}}$ converge to zero for high $\mathcal{P}_{\text{FEI}}$, the ratio of $f_{\text{CTF}}^{\text{excl}}/f_{\text{CTF}}^{\text{incl}} = (1 - f_{\text{ITF}}^{\text{excl}})/(1 - f_{\text{ITF}}^{\text{incl}})$ still converges to unity for very high $\mathcal{P}_{\text{FEI}}$ (Figure 11.4b).

Generally, the fraction of incorrect signal flavors (f_{ISF}, Figure 11.5) is much lower, starting at 25 % even for the lowest values of $\mathcal{P}_{\text{FEI}}$. An exception is observed for $\text{B} \to \text{D}\ell\nu_\ell$ Mixed TF, which remains consistently at 40 % (Figure 11.5a). This observation is linked to the very high values obtained for f_{IF} (Figures 11.2a and 11.2b). This quantity depends strongly on the definition of the flavor: f_{ISF} for $\text{B} \to \text{D}\ell\nu_\ell$ Mixed SF is still very high but converges to zero for high $\mathcal{P}_{\text{FEI}}$. The fact that a dependency on the tag side quantity $\mathcal{P}_{\text{FEI}}$ can exist is due to the connection of tag and signal side reconstruction by the completeness constraints applied to the exclusive dataset.

exclusive best candidate selection, which combines tag side B and signal side B mesons and selects the most promising combination (though it does not explicitly enforce consistent flavors).

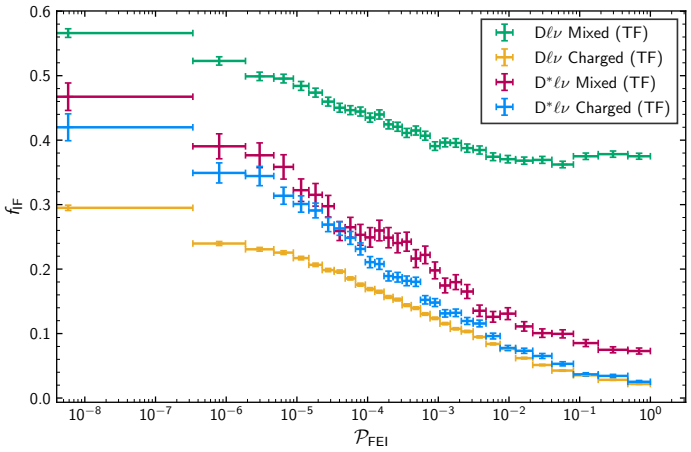
11.1.3 Observing the tension using only exclusive data

Figure 11.1 compares the calibration factors derived from fits to the exclusive dataset with the calibration factors derived from the $\text{B} \to X\ell\nu_\ell$ dataset. The observed tension indicates an additional dependency of the calibration factors (such as ITF/CTF). If this is the case, similar tensions should be observable using only the exclusive dataset. If the calibration factors do *not* depend on ITF/CTF/IF/CF, we expect that the calibration factors $\epsilon^{\text{SF}x}$ agree with the weighted averages[1]

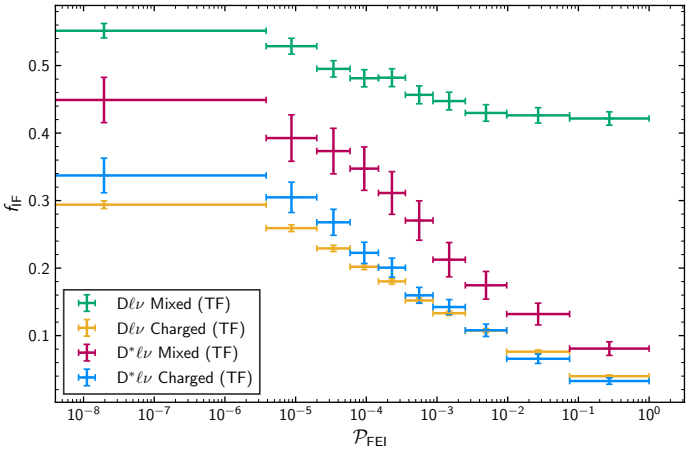
$$\tilde{\epsilon}^{\text{SF}x} = f_{\text{TF0}}^{\text{SF}x}\epsilon^{\text{TF0}} + f_{\text{TF+}}^{\text{SF}x}\epsilon^{\text{TF+}}, \quad \text{with} \quad f_{\text{TF}y}^{\text{SF}x} := \frac{\#\{\text{SF} = x \wedge \text{TF} = y\}}{\#\{\text{SF} = x\}}, \tag{11.1.1}$$

where $x = 0, +$ and $y = 0, +$ denote the mixed/charged flavor of the respective reconstructed B meson. The result of this comparison is shown in Figure 11.6. The resulting tension appears very similar to that of Figure 11.1 and particularly affects $\text{B} \to \text{D}\ell\nu_\ell$ Mixed SF. Therefore, it is

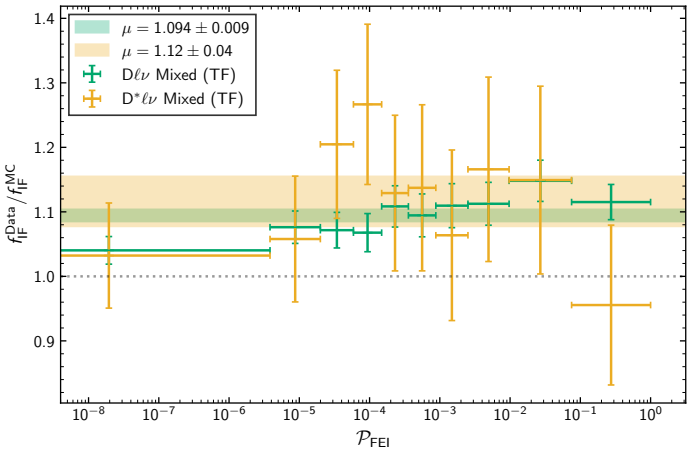
[1] This also neglects the effects of possible changes in the admixture of decay channels. However, these effects should be smaller than the deviation that is found. Note that this limitation does not apply to the comparison in Figure 11.1.



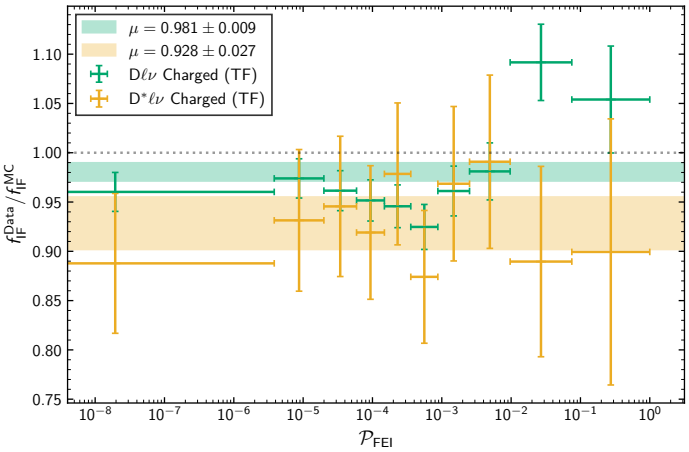
(a) Fraction of inconsistent flavors (signal ℓ MC).



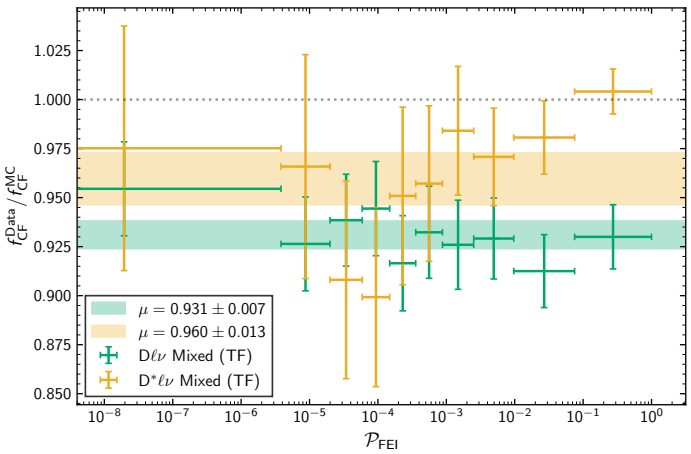
(b) Fraction of inconsistent flavors (Data).



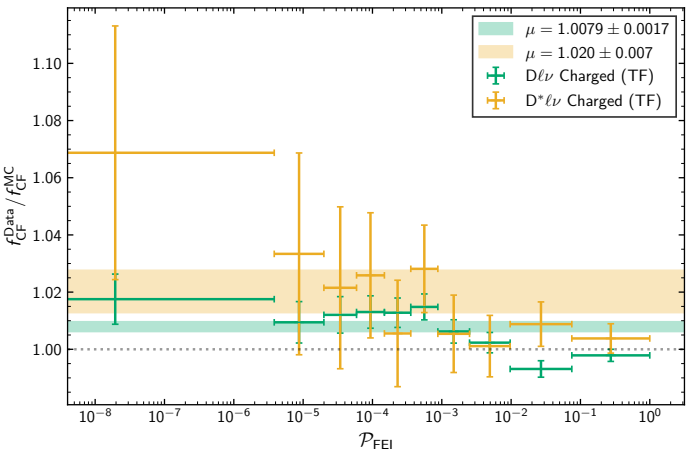
(c) Data/MC ratio of IF Mixed TF.



(d) Data/MC ratio of IF Charged TF.



(e) Data/MC ratio of CF Mixed TF.



(f) Data/MC ratio of CF Mixed TF.

Figure 11.2: Fraction of events with inconsistent and consistent flavor. In Figure 11.2a, the signal lepton component of MC is used. All other plots use the full MC dataset to allow for comparisons with data.

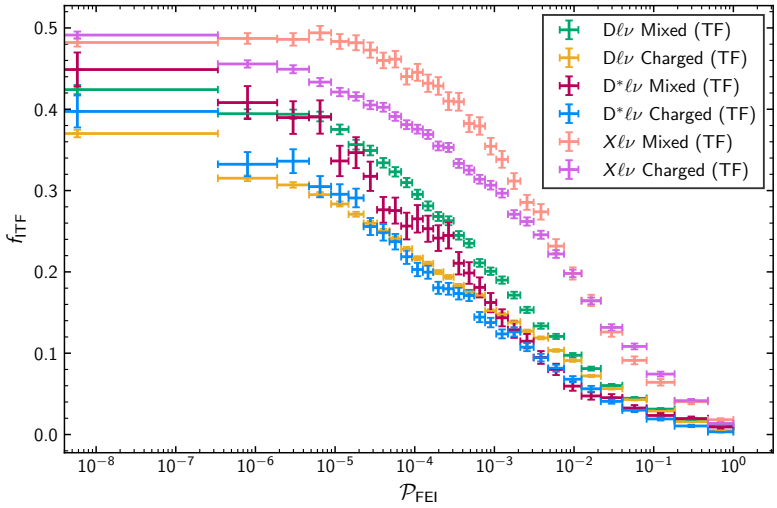
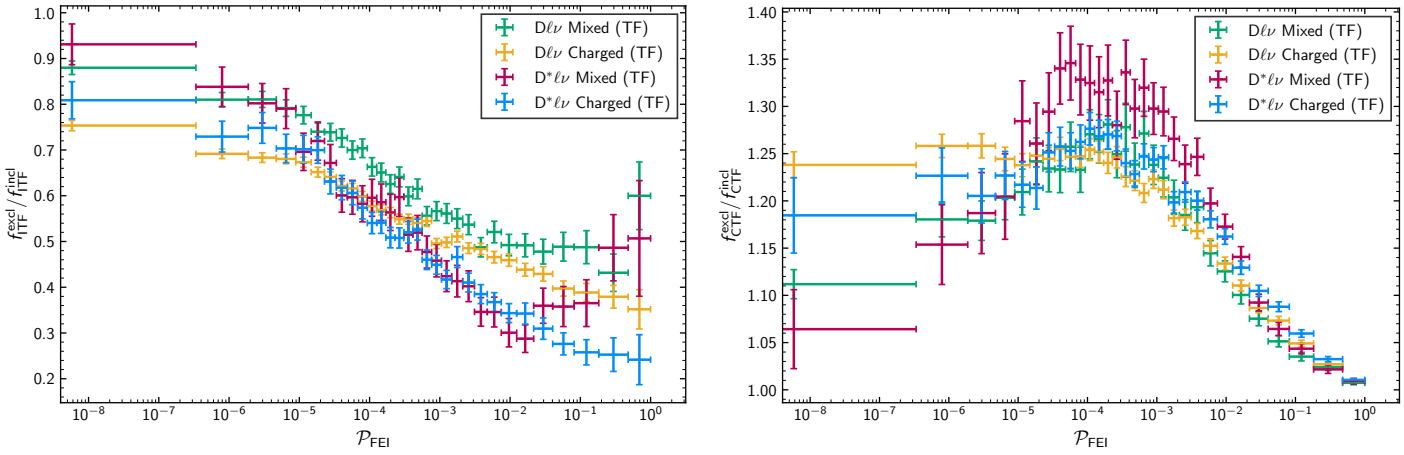


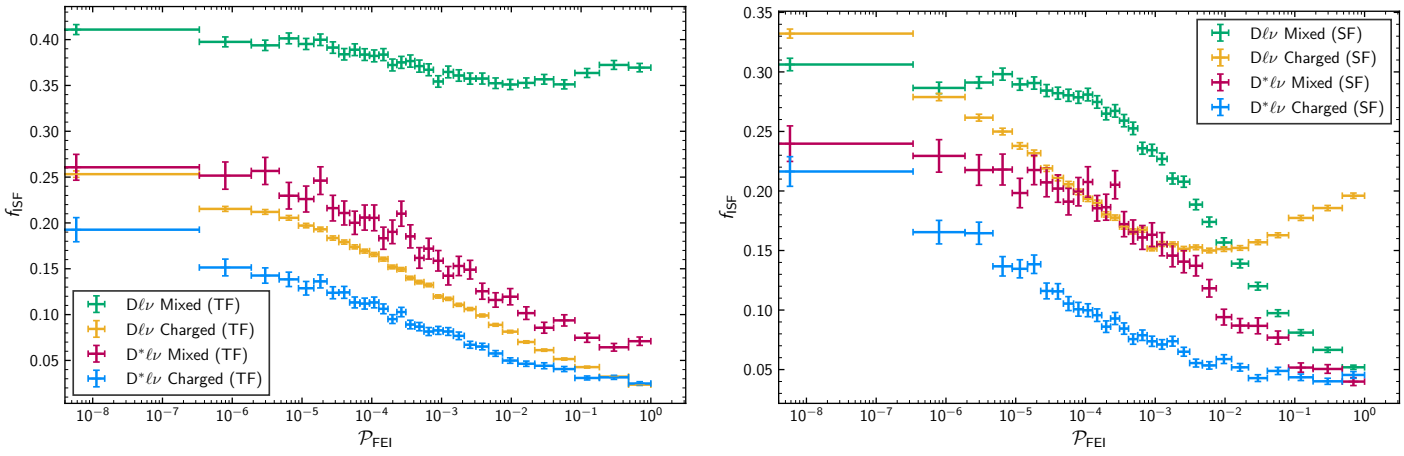
Figure 11.3: Fraction of incorrect tag flavors (signal lepton component of MC). Ratios of inclusive and exclusive curves are shown in Figure 11.4.



(a) Ratios of f_{ITF}.

(b) Ratios of f_{CTF}.

Figure 11.4: Comparing f_{ITF} and f_{CTF} between the inclusive and exclusive datasets (signal lepton component of MC).



(a) By tag flavor

(b) By signal flavor

Figure 11.5: Fraction of incorrect signal flavors (signal lepton component of MC).

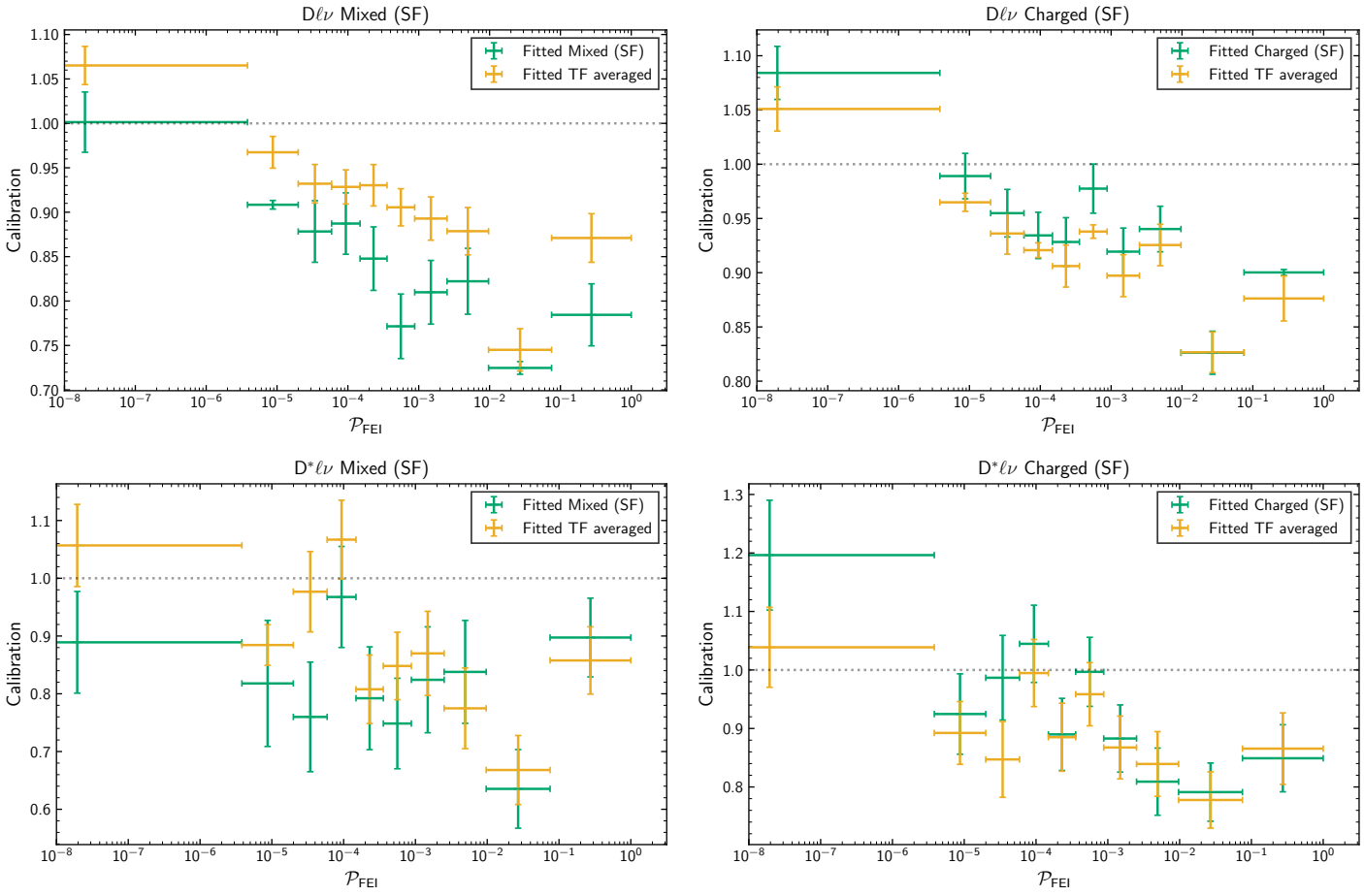


Figure 11.6: Comparing calibration factors from fits to $\mathrm{B} \to \mathrm{D}^{(*)}\ell\nu_\ell$ using specified signal flavors (green) with reweighted calibration factors from fits to $\mathrm{B} \to \mathrm{D}^{(*)}\ell\nu_\ell$ (yellow).

likely that this tension is caused by a general effect and not by an issue related to differences between the inclusive and exclusive datasets.

11.1.4 Adjusting the selection

One possibility to still apply the calibration factors presented in Chapter 10 is to define a subset where the tension is less prominent and can be covered by an additional systematic uncertainty. For example, if the effect is indeed caused by diverging calibration factors for ITF/CTF, finding a selection where f_{ITF} is similar between the inclusive and exclusive datasets can reduce the observed tension.

The most obvious choice for such a selection would be a lower bound on $\mathcal{P}_{\text{FEI}}$. However, the tension does not disappear even for the highest $\mathcal{P}_{\text{FEI}}$ bins (see Figure 11.6 or Figure 11.1). Moreover, we see that the difference[2] in f_{ITF} remains substantial for high $\mathcal{P}_{\text{FEI}}$ (Figure 11.7).

Another possibility is a selection based on the FEI reconstruction modes. As mentioned in Section 4.1.1, the substantial efficiency gains of the FEI over the FR in the low purity region are connected to the addition of more complicated reconstruction modes. Removing reconstruction modes of large purity differences between the inclusive and exclusive datasets could potentially reduce the observed tension. Unfortunately, Figure 11.8 shows large purity differences for almost all reconstruction modes, making this strategy not feasible.

[2] While the ratio $f_{\text{ITF}}^{\text{incl}}/f_{\text{ITF}}^{\text{excl}}$ is more natural to understand the general picture, the difference $f_{\text{ITF}}^{\text{incl}} - f_{\text{ITF}}^{\text{excl}}$ is better suited to quantify the impact on the observed tension, see Equation (11.3.2).

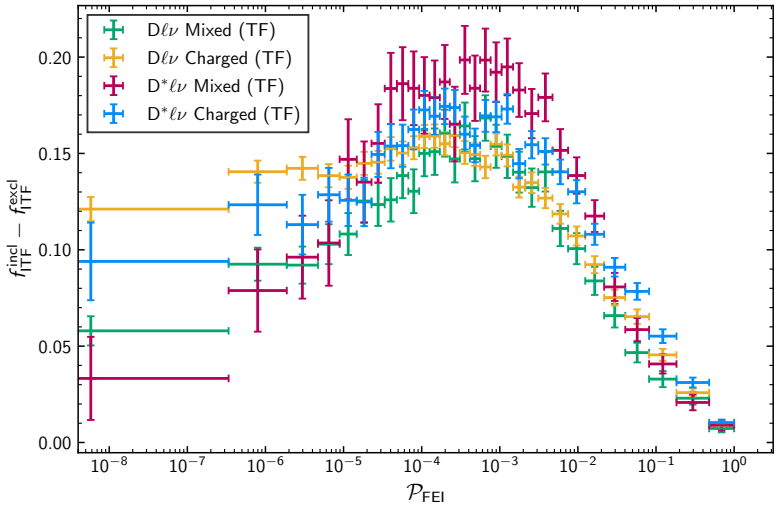


Figure 11.7: Difference between incorrectly reconstructed tag fractions of the inclusive and exclusive datasets. For the same information as a ratio, see Figure 11.4a.

11.2 Using the signal side in exclusive data

This section investigates the effect by performing additional fits using the signal side flavor variables in the exclusive datasets. This allows separately measuring ϵ_{IF} and ϵ_{CF}, and to relate these quantities to values for ϵ_{ITF} and ϵ_{CTF}. While this is not possible in $\mathrm{B} \to X\ell\nu_\ell$, we expect to improve our understanding and find a simple way to correct the granular calibration factors from $\mathrm{B} \to X\ell\nu_\ell$.

11.2.1 Fitting inconsistent flavor calibration factors

Since it is possible to determine the flavor consistency on both data and MC, we use separate fits to determine the IF and CF calibration factors. This is shown in Figure 11.9. As shown in Figure 11.10, using these two sets of calibration factors can resolve the tension. The resulting closure is not entirely trivial, since a total of 8 combinations of truth, TF and SF flavors can be distinguished, and we only use 4 different calibration factors (Mixed TF IF, Charged TF IF, Mixed TF CF, Mixed TF CF).

Figures 11.9e and 11.9f show the ratios $\epsilon_{\mathrm{IF}}/\epsilon_{\mathrm{CF}}$. While the ratio for $\mathrm{B} \to \mathrm{D}\ell\nu_\ell$ Mixed TF is flat ($p = 90\,\%$), checks for flatness for the other cases result in p-values between 2 % and 3 % after excluding any outliers more than 3σ away from the average. The ratios between $\mathrm{B} \to \mathrm{D}\ell\nu_\ell$ and $\mathrm{B} \to \mathrm{D}^*\ell\nu_\ell$ are only compatible at a $p = 6\,\%$ level after excluding outliers with the same definition. However, it should be noted that this is a very preliminary treatment of uncertainties: Not only do the $\mathrm{B} \to \mathrm{D}\ell\nu_\ell$ and $\mathrm{B} \to \mathrm{D}^*\ell\nu_\ell$ samples have a significant overlap (we only match for signal ℓ), but we also do not consider any systematic uncertainties on signal side.

11.2.2 Translation to (in)correct tag flavor calibration factors

We cannot directly *fit* ITF/CTF calibration factors since we do not have the underlying truth flavor on data. Instead, we can determine them using a system of linear equations, assuming that any difference between ϵ_{IF} and ϵ_{CF} is caused by the different values in f_{ITF}.

The system of equations is given by:[3]

[3] It appear more natural to directly consider ϵ^{TF0}, $\epsilon^{\mathrm{TF+}}$, ϵ^{SF0}, and $\epsilon^{\mathrm{SF+}}$ that motivated the investigation. However, this leads to a singular mixing matrix: Because Mixed/Charged TF and Mixed/Charged SF both encompass the full dataset, each calibration factor is a linear combination of the three others (with coefficients from MC).

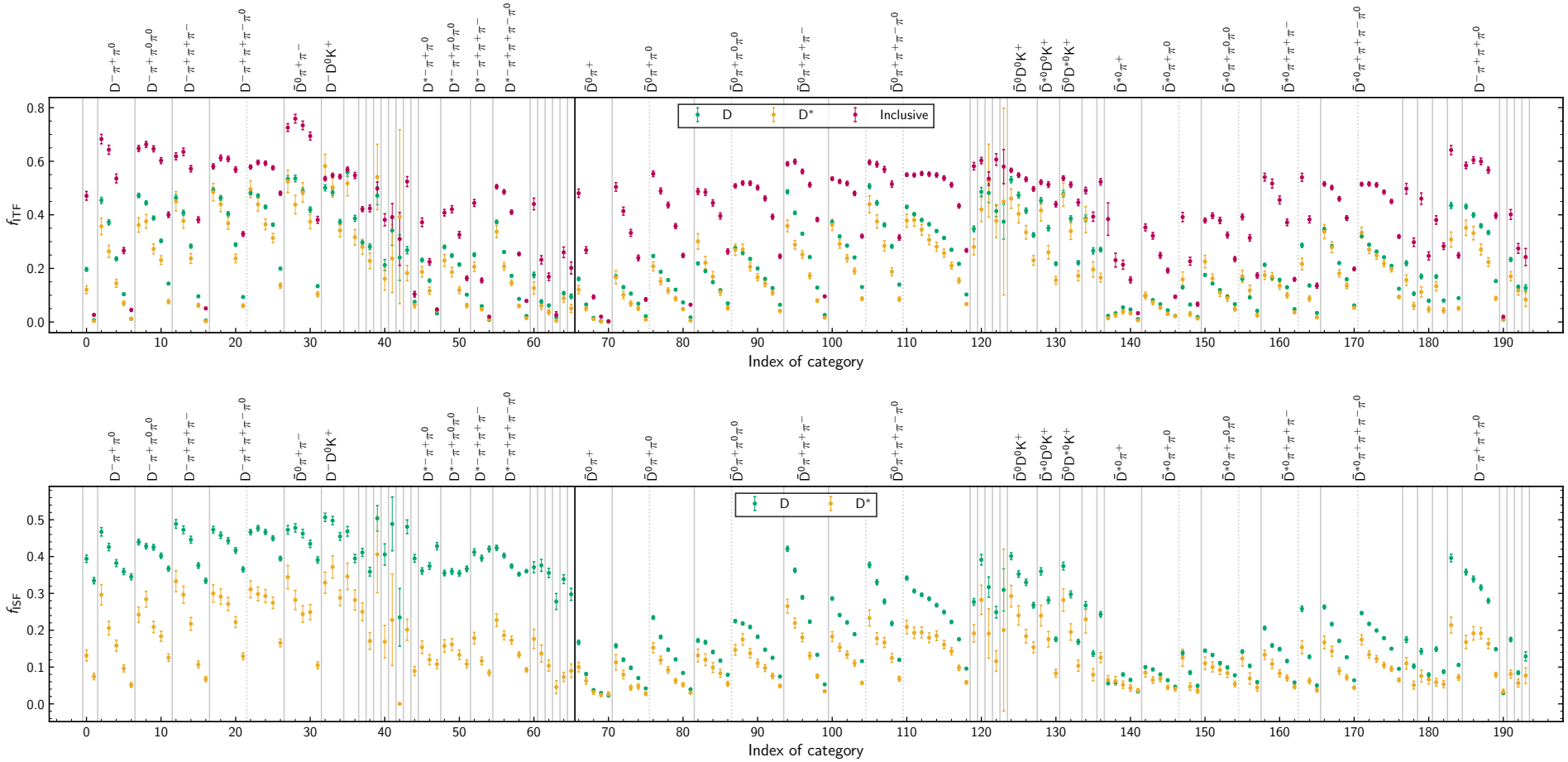


Figure 11.8: Comparing the fractions f_{ITF} and f_{ISF} for all categories.

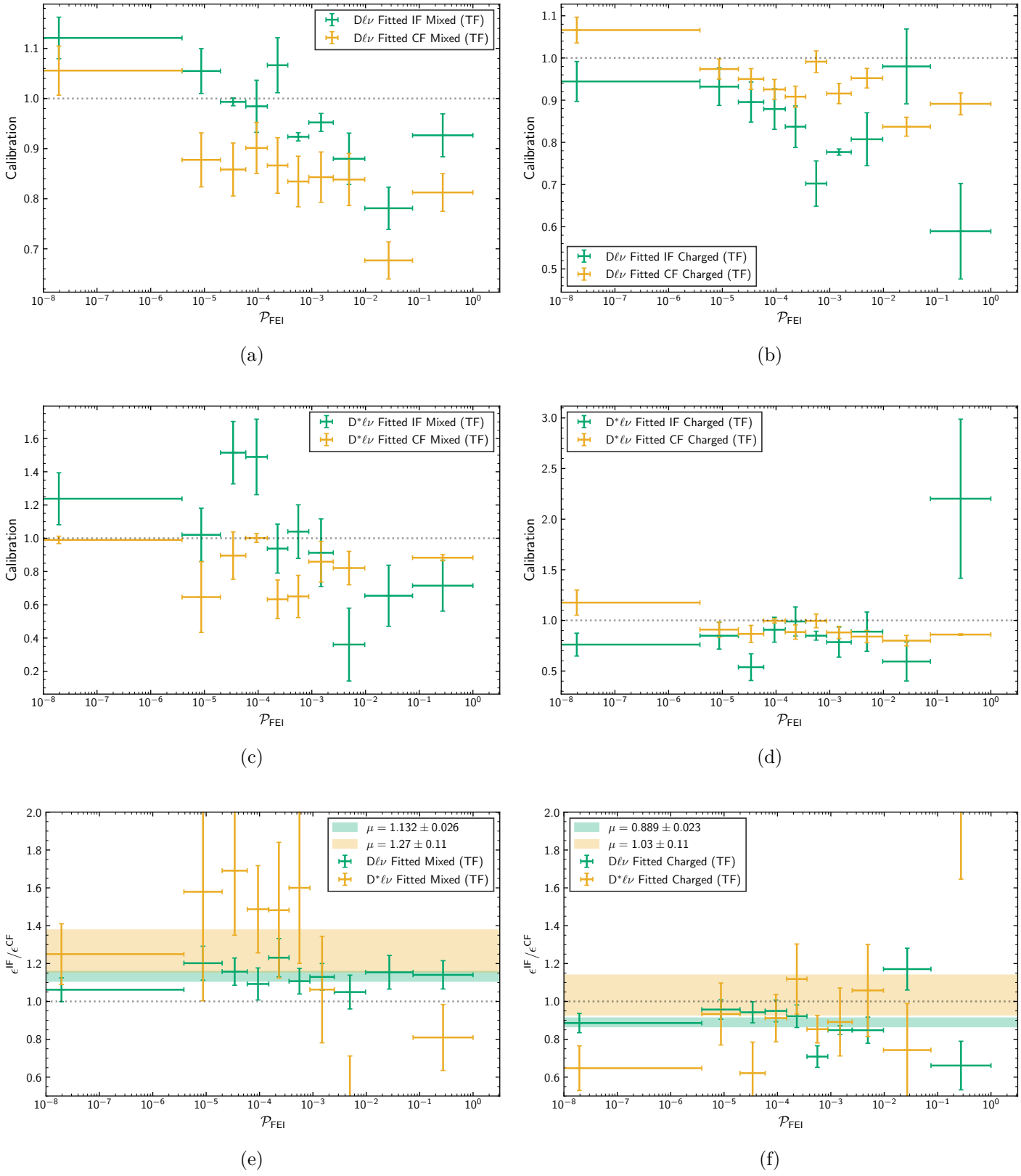


Figure 11.9: Separately calibrating consistent and inconsistent flavor on the exclusive dataset.

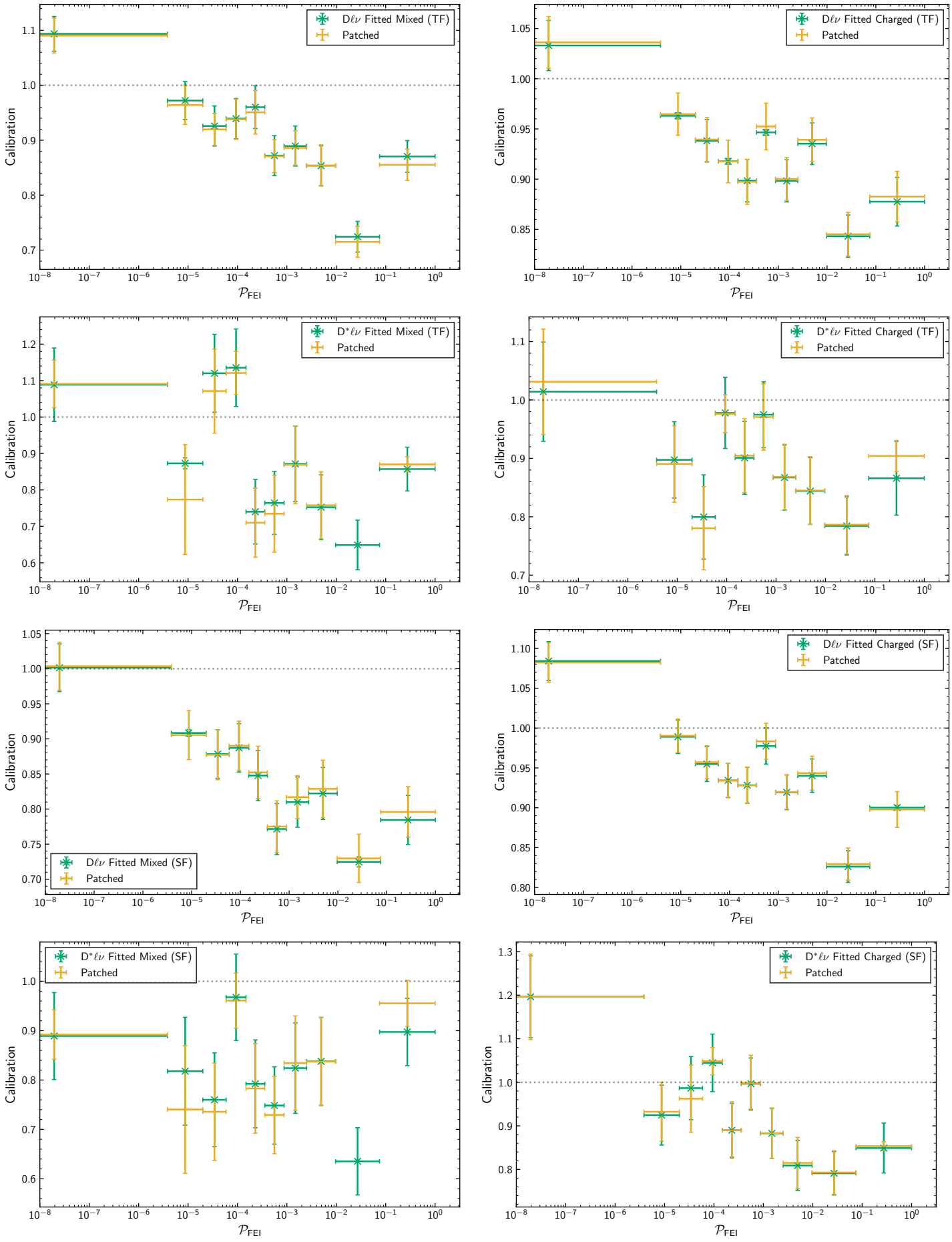


Figure 11.10: Comparing linear combinations of ϵ_{IF} and ϵ_{CF} with the calibration factors fitted for the two different definitions of flavor.

	$\langle\epsilon_{\mathrm{ITF}}/\epsilon_{\mathrm{CF}}\rangle$	χ^2/ndf	p	ndf
$\mathrm{B}\to\mathrm{D}\ell\nu_\ell$ Mixed TF	1.89 ± 0.24	0.64	0.72	7
$\mathrm{B}\to\mathrm{D}^*\ell\nu_\ell$ Mixed TF	2.1 ± 0.4	0.66	0.68	6
$\mathrm{B}\to\mathrm{D}\ell\nu_\ell$ Charged TF	0.85 ± 0.04	3.84 (1.17)	2×10^{-4} (0.32)	8 (6)
$\mathrm{B}\to\mathrm{D}^*\ell\nu_\ell$ Charged TF	0.79 ± 0.10	1.12	0.35	8

Table 11.1: Tests for uniformity of $\epsilon_{\mathrm{ITF}}/\epsilon_{\mathrm{CF}}$. Only values between 0.0 and 3.0 are considered (resulting in different degrees of freedom, *ndf*). The uncertainties on the average $\langle\epsilon_{\mathrm{ITF}}/\epsilon_{\mathrm{CF}}\rangle$ are not considered in the χ^2 test; the number of degrees of freedom is set to the number of bins reduced by one. For $\mathrm{B}\to\mathrm{D}\ell\nu_\ell$ Charged TF, the second set of values in parentheses shows the results after excluding outliers of $>3\sigma$.

$$\begin{pmatrix}\epsilon_{\mathrm{IF}}^{\mathrm{TF0}}\\ \epsilon_{\mathrm{CF}}^{\mathrm{TF0}}\\ \epsilon_{\mathrm{IF}}^{\mathrm{TF+}}\\ \epsilon_{\mathrm{CF}}^{\mathrm{TF+}}\end{pmatrix}=\begin{pmatrix}f_{\mathrm{ITF}}^{\mathrm{IF\,TF0}} & 1-f_{\mathrm{ITF}}^{\mathrm{IF\,TF0}} & 0 & 0\\ f_{\mathrm{ITF}}^{\mathrm{CF\,TF0}} & 1-f_{\mathrm{ITF}}^{\mathrm{CF\,TF0}} & 0 & 0\\ 0 & 0 & f_{\mathrm{ITF}}^{\mathrm{IF\,TF+}} & 1-f_{\mathrm{ITF}}^{\mathrm{IF\,TF+}}\\ 0 & 0 & f_{\mathrm{ITF}}^{\mathrm{CF\,TF+}} & 1-f_{\mathrm{ITF}}^{\mathrm{CF\,TF+}}\end{pmatrix}\begin{pmatrix}\epsilon_{\mathrm{ITF}}^{\mathrm{TF0}}\\ \epsilon_{\mathrm{CTF}}^{\mathrm{TF0}}\\ \epsilon_{\mathrm{ITF}}^{\mathrm{TF+}}\\ \epsilon_{\mathrm{CTF}}^{\mathrm{TF+}}\end{pmatrix}, \tag{11.2.1}$$

where

$$f_{\mathrm{ITF}}^{X\,\mathrm{TF}y}=\frac{\#\{\mathrm{TF}=y\wedge X\wedge\mathrm{ITF}\}}{\#\{\mathrm{TF}=y\wedge X\}},\quad(X=\mathrm{IF},\mathrm{CF}\text{ and }y=0,+). \tag{11.2.2}$$

Equation (11.2.1) is solved for every bin in $\mathcal{P}_{\mathrm{FEI}}$. The result is shown in Equation (11.2.1). The ratio $\epsilon_{\mathrm{ITF}}/\epsilon_{\mathrm{CF}}$ generally assumes large positive values for Mixed TF, but remains below unity for the Charged TF.

The individual ratios are compatible between the $\mathrm{B}\to\mathrm{D}\ell\nu_\ell$ and $\mathrm{B}\to\mathrm{D}^*\ell\nu_\ell$ Mixed TF (Charged TF) datasets with $p=0.94$ ($p=0.10$) when considering values of $0<\epsilon_{\mathrm{ITF}}/\epsilon_{\mathrm{CF}}<3$ (which excludes unreasonable results possibly produced by near-singular mixing matrices). However, this comparison does not account for the event overlap between the two samples and thus requires further investigation.

Table 11.1 shows that the distributions of the ratios are compatible with a uniform distribution given by the average ratio. Furthermore, these averages are well-compatible between $\mathrm{B}\to\mathrm{D}\ell\nu_\ell$ and $\mathrm{B}\to\mathrm{D}^*\ell\nu_\ell$.

11.2.3 Using ratios to correct the $X\ell\nu_\ell$ calibration factors

We can use the just derived ratios to correct the calibration factors derived from the $X\ell\nu_\ell$ calibration. For this, we use the values derived from $\mathrm{B}\to\mathrm{D}\ell\nu_\ell$ (see Table 11.1),

$$r^0=1.89\pm0.24\quad\text{and}\quad r^+=0.85\pm0.04, \tag{11.2.3}$$

where $r=\epsilon_{\mathrm{ITF}}/\epsilon_{\mathrm{CF}}$, and the superscript denotes the tag flavor.

Using the differential $X\ell\nu_\ell$ calibration factors averaged over the corresponding bin in the exclusive data, denoted $\epsilon^{\mathrm{TF}x}$, we can calculate separate CTF and ITF calibration factors as the linear combinations

$$\epsilon_{\mathrm{CTF}}^{\mathrm{TF}x}=\frac{\epsilon^{\mathrm{TF}x}}{f_{\mathrm{ITF}}^{\mathrm{TF}x}r^x+(1-f_{\mathrm{ITF}}^{\mathrm{TF}x})}\quad\text{and}\quad\epsilon_{\mathrm{ITF}}^{\mathrm{TF}x}=r^x\epsilon_{\mathrm{CTF}}^{\mathrm{TF}x}\quad(x=0,+), \tag{11.2.4}$$

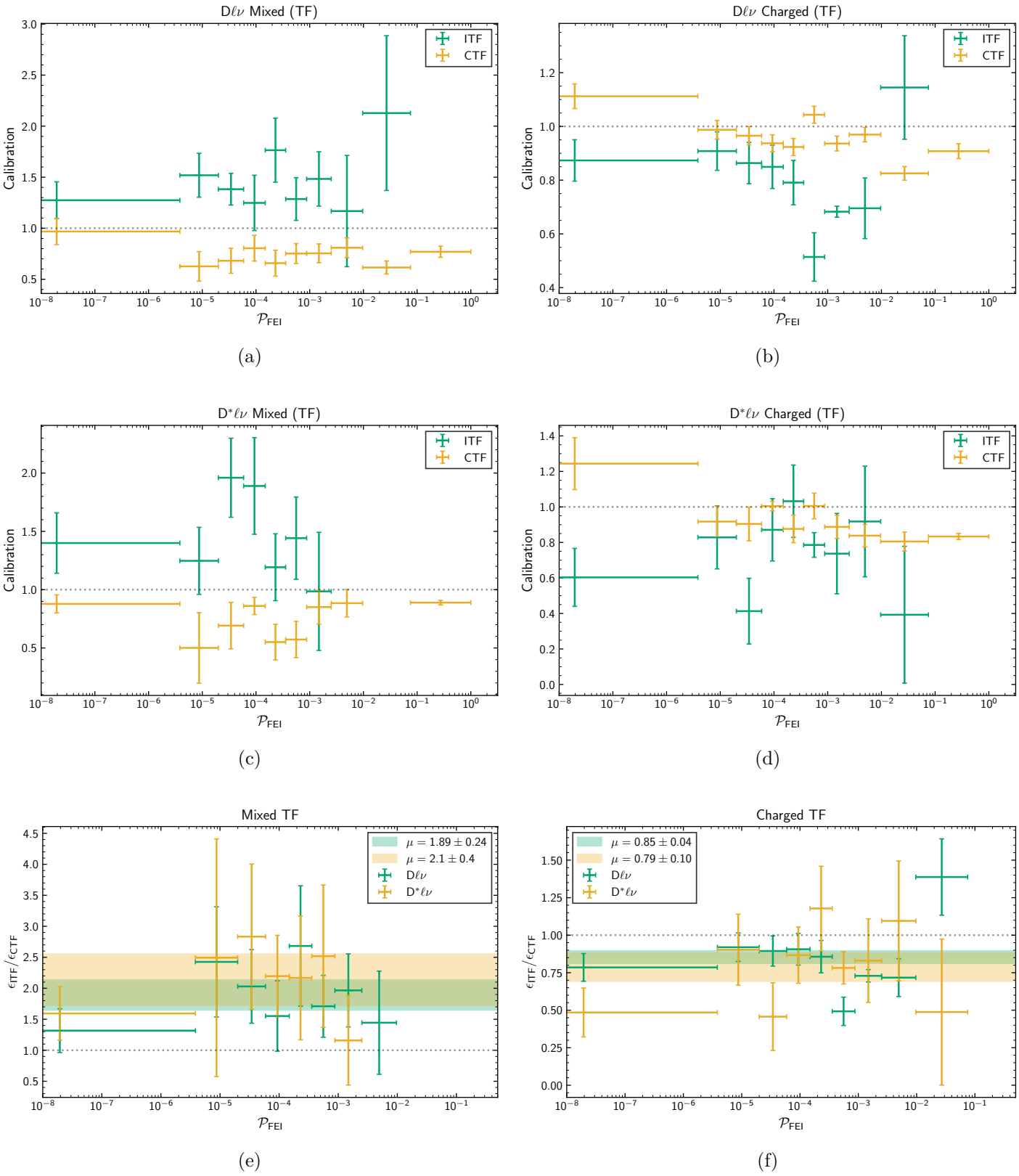


Figure 11.11: Solving for the calibration factors of correctly and incorrectly reconstructed tag flavors. Values not between 0.0 and 3.0 are not shown and excluded from the calculation of the mean values.

where the fractions f are defined as in Equation (11.2.2) (but evaluated on the $X\ell\nu_\ell$ dataset!).[4]

Using these calibration factors, we can then calculate corrected calibration factors for any subset Y of the exclusive dataset as

$$\epsilon^Y = \sum_{x=0,+} \sum_{Z=\text{ITF,CTF}} f^Y_{\text{TF}x\,Z}\, \epsilon_Z^{\text{TF}x}, \tag{11.2.5}$$

where

$$f^Y_{\text{TF}x\,Z} = \frac{\#\{Y \wedge \text{TF} = x \wedge Z\}}{\#\{Y\}}, \quad (x = 0, + \text{ and } Z = \text{ITF}, \text{CTF}). \tag{11.2.6}$$

is evaluated on the exclusive dataset. With this, we can perform a similar comparison to Figure 11.1 that motivated our investigation. The result is presented in Figure 11.12 and shows an overall good agreement. Because the two quantities r^0 and r^+ are the only additional inputs from the $\text{B} \to \text{D}\ell\nu_\ell$ dataset, this confirms the efficacy of the $\text{B} \to X\ell\nu_\ell$ calibration approach.

Using the same strategy with the fit model presented in Section 5.1.2 allows subtracting $\text{B} \to \text{D}^*\ell\nu_\ell$ crossfeed in the $\text{B} \to \text{D}\ell\nu_\ell$ dataset and applying the corrected calibration factors on $\text{B} \to \text{D}^*\ell\nu_\ell$. This result is currently in preparation.

11.3 Measuring (in)correct tag flavor calibration factors using $m_{\text{bc}}^{\text{tag}}$

The last section has shown that information from the $\text{B} \to \text{D}\ell\nu_\ell$ dataset can be used to correct the calibration factors measured on $\text{B} \to X\ell\nu_\ell$. A different avenue is to use a variable that can discriminate between correctly and incorrectly reconstructed tag flavors and directly measure separate calibration factors on $\text{B} \to X\ell\nu_\ell$. The respective shapes of our previous fit variables, $|\vec{p}^{\,*}_\ell|$ and $\mathcal{L}_{\text{CS}}$ are not suitable for this (Figure 11.13). As already mentioned, one option is to utilize the $m_{\text{bc}}^{\text{tag}}$ observable (Figure 11.14).

However, the currently available dataset imposes a selection requirement of $m_{\text{bc}}^{\text{tag}} > 5.27\,\text{GeV}$. This means that the available discriminatory power is relatively limited, leading to strong correlations between calibration factors for ITF and CTF and large uncertainties.

We consider four different methods to extract the calibration factors:

2D fit We perform a two-dimensional fit using $|\vec{p}^{\,*}_\ell|$ (5 bins) and $m_{\text{bc}}^{\text{tag}}$ (9 bins) with three components: ITF signal, CTF signal, and background. Here, *signal* is defined as usual for the tagging calibration.

3D fit We utilize a three-dimensional fit using $|\vec{p}^{\,*}_\ell|$ (5 bins), $m_{\text{bc}}^{\text{tag}}$ (5 bins), $\mathcal{L}_{\text{CS}}$ (2 bins) with the same fit components.

Mixing matrix strategy (MMS) We perform a $|\vec{p}^{\,*}_\ell|$ fit (9 bins, signal vs background) for the low and high $m_{\text{bc}}^{\text{tag}}$ regions. The calibration factors for ITF and CTF are then inferred by solving a two-dimensional system of linear equations (similar to Section 11.2.2). This strategy is explained in detail in Section 11.3.1.

Extended mixing matrix (EMMS) This method is similar to the previous strategy, but we perform the fits for *several* bins in $m_{\text{bc}}^{\text{tag}}$. The calibration factors for ITF and CTF are then given by the optimal solutions for an overdetermined system of linear equations. This strategy is explained in detail in Section 11.3.2.

All four strategies require that the calibration factors do not depend on $m_{\text{bc}}^{\text{tag}}$ other than by correlation with f_{ITF}. This is further investigated in Section 11.3.2.

[4] This is a simplified approach: The more precise way would be to use the ϵ_i for all categories, apply Equation (11.2.4) to each of them, and only then take the average over the bin in the exclusive data.

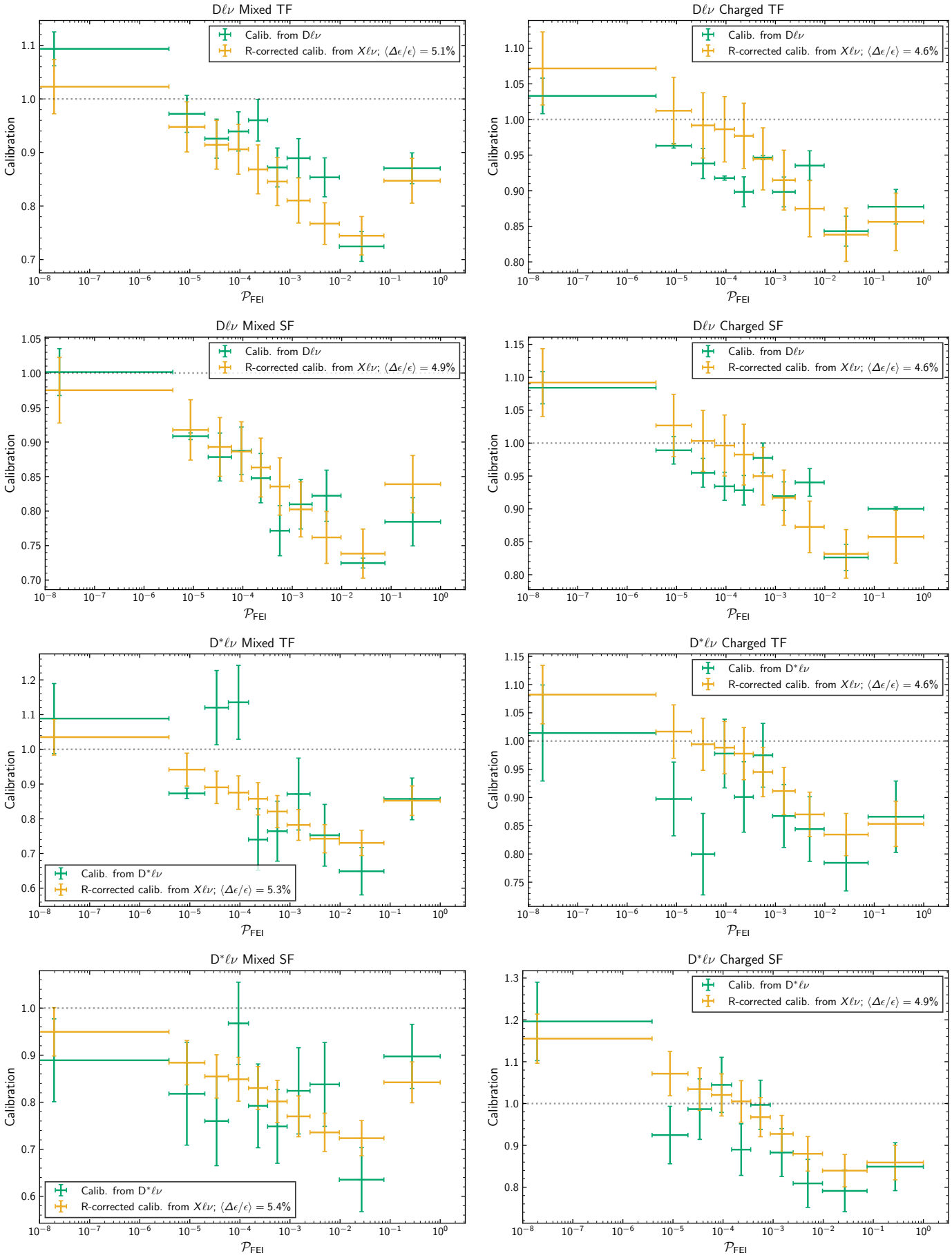


Figure 11.12: Calibration factors fitted on $B \to D^{(*)}\ell\nu_\ell$ vs. the weighted averages of the calibration factors from $B \to X\ell\nu_\ell$ that have been corrected with r^0, r^+. $\langle\Delta\epsilon/\epsilon\rangle$ refers to the average relative calibration uncertainty. The equivalent plot without the correction is Figure 11.1.

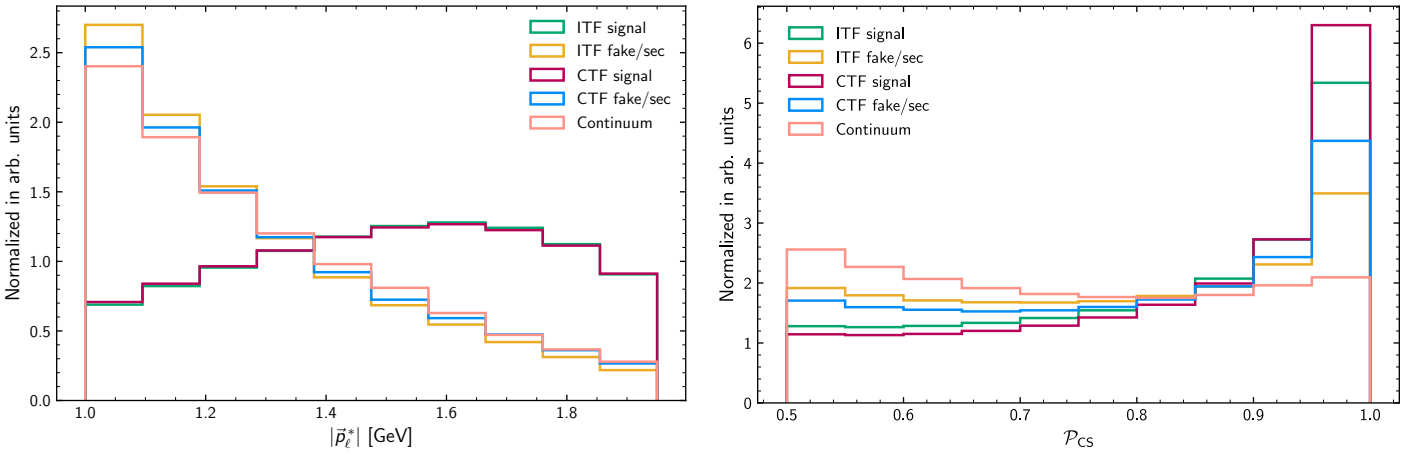


Figure 11.13: Fit variable distributions for different flavor configurations.

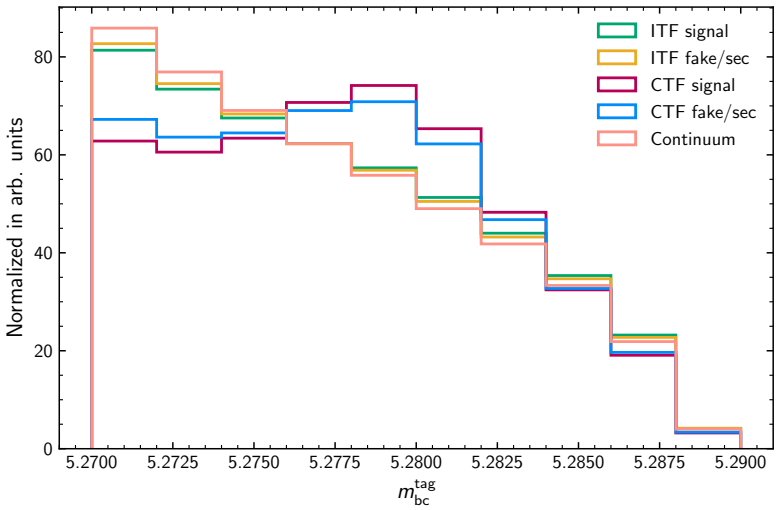


Figure 11.14: The $m_{\text{bc}}^{\text{tag}}$ distribution for different flavor configurations.

11.3.1 Mixing matrix strategy

This strategy divides data and MC into two subsets with distinct incorrect flavor tag fractions $f_{\text{ITF}} \neq f'_{\text{ITF}}$ and calculates calibration factors ϵ and ϵ' for both subsets. Assuming that the difference in calibration factors can be attributed to the difference between f_{ITF} and f'_{ITF}, we observe

$$\begin{pmatrix} \epsilon \\ \epsilon' \end{pmatrix} = \begin{pmatrix} f_{\text{ITF}} & 1 - f_{\text{ITF}} \\ f'_{\text{ITF}} & 1 - f'_{\text{ITF}} \end{pmatrix} \begin{pmatrix} \epsilon_{\text{ITF}} \\ \epsilon_{\text{CTF}} \end{pmatrix}. \tag{11.3.1}$$

By inverting the matrix[5], we can thus calculate ϵ_{ITF} and ϵ_{CTF} from ϵ and ϵ'. A simple conclusion is

$$\epsilon_{\text{ITF}} - \epsilon_{\text{CTF}} = \frac{\epsilon - \epsilon'}{f_{\text{ITF}} - f'_{\text{ITF}}}, \tag{11.3.2}$$

which shows that this method becomes numerically unstable for low values of $|f_{\text{ITF}} - f'_{\text{ITF}}|$.

To apply this strategy, we divide the dataset into $m_{\text{bc}}^{\text{tag}} < s$ and $m_{\text{bc}}^{\text{tag}} > s$. The threshold s should be chosen to maximize $f_{\text{ITF}} - f'_{\text{ITF}}$ while ensuring that both datasets have a sufficient number of events. We choose $s = 5.275\,\text{GeV}$ (see Figure 11.15). Currently, the uncertainty on f_{ITF} is not

[5]The determinant is $f_{\text{ITF}} - f'_{\text{ITF}}$, so this possible for all $f_{\text{ITF}} \neq f'_{\text{ITF}}$.

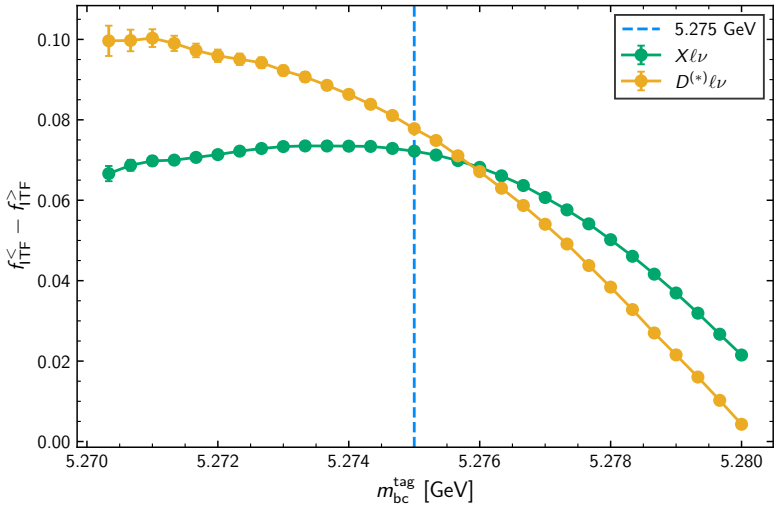


Figure 11.15: Optimizing the $m_{\mathrm{bc}}^{\mathrm{tag}}$ cut to divide the dataset. Shown is the difference between incorrect tag flavor fractions of the subsets satisfying $m_{\mathrm{bc}}^{\mathrm{tag}} < s$ and $m_{\mathrm{bc}}^{\mathrm{tag}} > s$ vs. the value of s.

included in the uncertainty for the results.

11.3.2 Extended mixing matrix strategy

A variant of the previous method is to consider multiple subsets of our dataset, leading to an overdetermination of the linear system. This means that we assume

$$\epsilon_{\mathrm{predicted}}^{(i)} = f_{\mathrm{ITF}}^{(i)}\epsilon_{\mathrm{ITF}} + (1 - f_{\mathrm{ITF}}^{(i)})\epsilon_{\mathrm{CTF}}, \tag{11.3.3}$$

where the index i enumerates the subset of our dataset and ϵ_{ITF}, ϵ_{CTF} are constants. We can then optimize the correspondence between $\epsilon_{\mathrm{predicted}}^{(i)}$ and $\epsilon_{\mathrm{measured}}^{(i)}$ (e.g., with a χ^2 test) with respect to ϵ_{ITF}, ϵ_{CTF}. An advantage of this strategy is that we automatically test the underlying assumption that all differences in $\epsilon^{(i)}$ can be attributed to $f_{\mathrm{ITF}}^{(i)}$.

The result is shown in Figure 11.16. For $\mathcal{P}_{\mathrm{FEI}} < 10^{-4}$ (approximately the lower 50 % quantile), this condition is well fulfilled, with p-values of 0.69 (Mixed TF) and 0.37 (Charged TF). For $\mathcal{P}_{\mathrm{FEI}} > 10^{-4}$, $\epsilon_{\mathrm{predicted}}^{(i)}$ and $\epsilon_{\mathrm{measured}}^{(i)}$ have statistically significant differences, in particular at high values of $m_{\mathrm{bc}}^{\mathrm{tag}}$. An explanation is that $m_{\mathrm{bc}}^{\mathrm{tag}}$ and $\mathcal{P}_{\mathrm{FEI}}$ are not entirely uncorrelated and that two bins in $\mathcal{P}_{\mathrm{FEI}}$ are insufficient to describe this dependency. However, the differences in all bins are lower than 2 %, which we deem acceptable for this proof-of-concept study, considering that the differences between the ITF and CTF calibration factors are found to be very large. Once the low-$m_{\mathrm{bc}}^{\mathrm{tag}}$ dataset is available, these discrepancies can be investigated more thoroughly.

11.3.3 Results

Figure 11.17 compares different results for the global ITF and CTF calibration factors measured on the inclusive dataset. The results of the various strategies do not agree within statistical uncertainties. This tension does not necessarily imply that some methods are inadequate or that additional effects are involved. Instead, it can be assumed that the discriminatory power of the $m_{\mathrm{bc}}^{\mathrm{tag}}$ distribution is currently too limited and leads to numerical instabilities, or that additional calibration dependencies, such as $\mathcal{P}_{\mathrm{FEI}}$ or C_{FEI}, need to be considered. Nonetheless, there is a very clear and consistent separation between ITF and CTF calibration factors, with the ϵ_{ITF} calibration exceeding ϵ_{CTF} up to twofold.

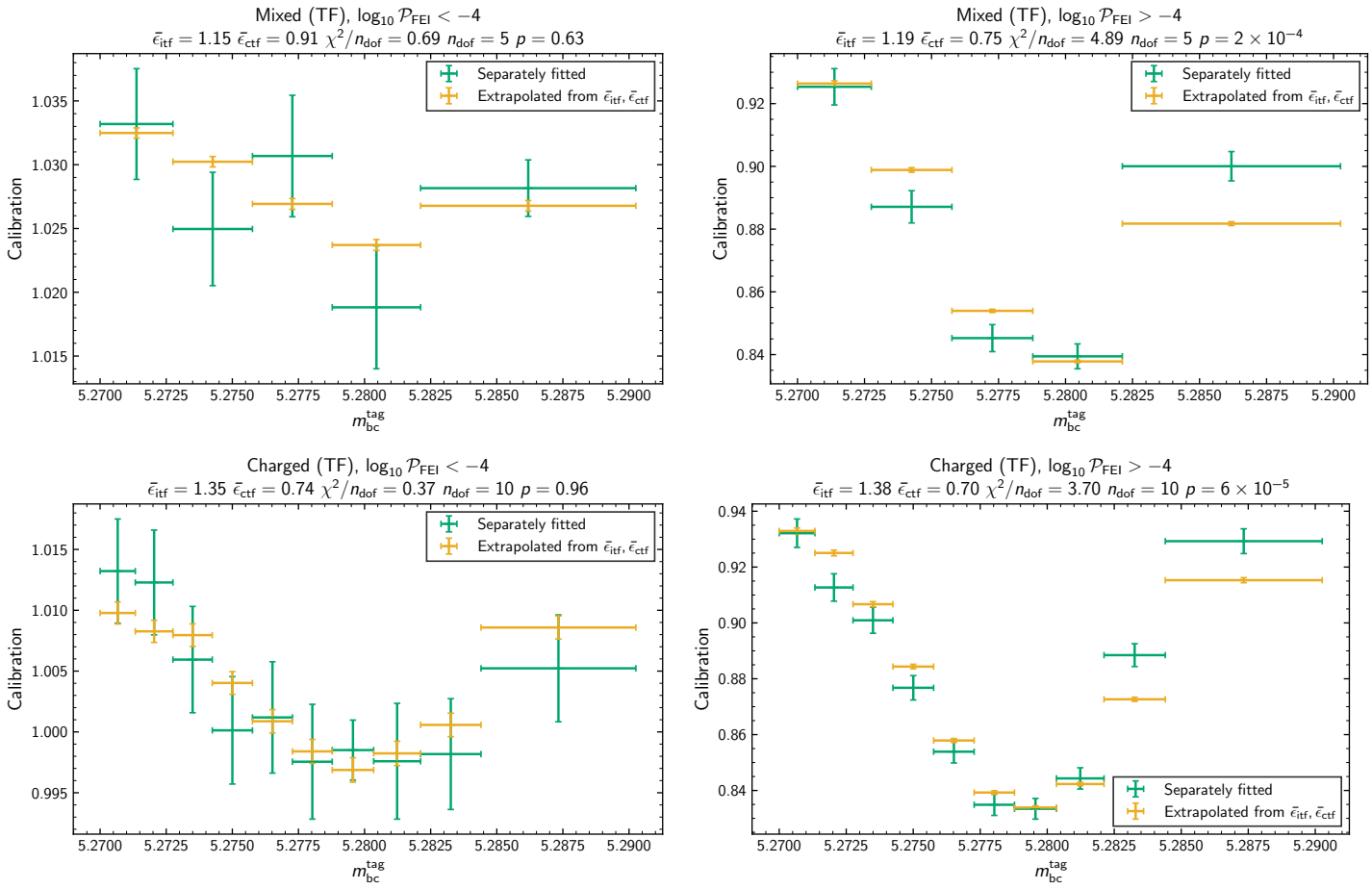


Figure 11.16: Testing whether the dependency of the calibration factors on $m_{\text{bc}}^{\text{tag}}$ can be attributed to the different values of f_{ITF}. The yellow data points show the expectation based on Equation (11.3.3). Only the statistical uncertainties on f_{ITF} are included.

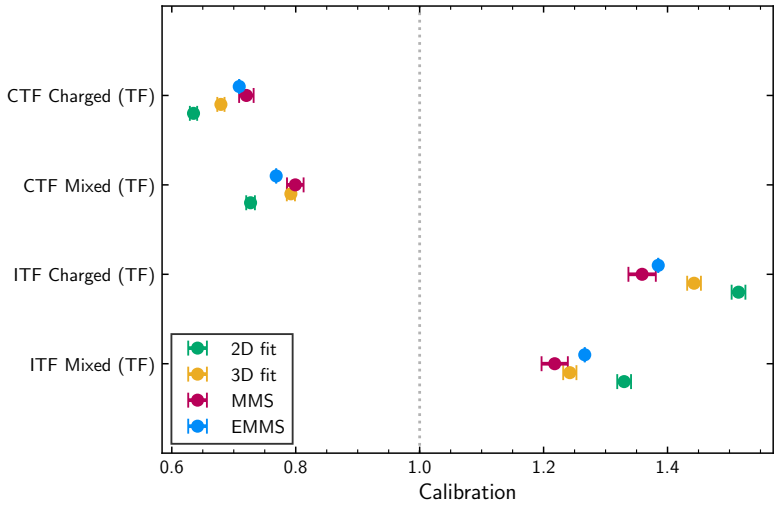


Figure 11.17: Comparing global ITF and CTF calibration factors extracted with different strategies on the inclusive dataset. Only statistical uncertainties are shown. No uncertainties are calculated for EMMS.

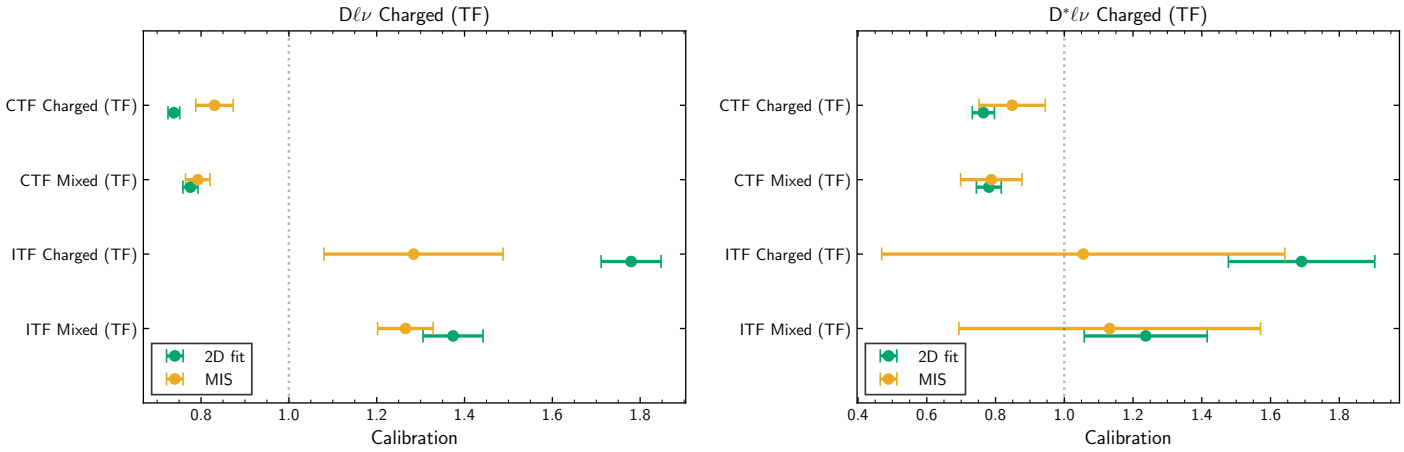


Figure 11.18: Comparing global ITF and CTF calibration factors extracted with different strategies on the exclusive datasets.

Applying the same method to the exclusive dataset produces similar results (Figure 11.18).[6] Due to the smaller data samples, the uncertainties are significantly increased and we do not use the three-dimensional fit or the EMMS strategy. The results partially contradict the observations of Section 11.2.2: While similar ratios of $\epsilon_{\text{ITF}}/\epsilon_{\text{CTF}} \sim 2$ are obtained for Mixed TF, the ratios for Charged TF disagree.

As already mentioned, the calculation of Section 11.2.2 relies on the assumption that the difference between calibration factors for IF and CF can be reduced to different values of f_{ITF}. Using the $m_{\text{bc}}^{\text{tag}}$ strategies, we can test this assumption by separately extracting ITF and CTF calibration factors for IF and CF. If the assumption holds true, the respective results should be compatible between IF and CF. The large uncertainties make this comparison difficult, but the results presented in Figure 11.19, particularly for $\text{B} \to \text{D}\ell\nu_\ell$ Mixed TF, cast doubt on the assumption. However, the same limitations mentioned in the previous paragraphs apply to the comparison.

Another caveat is that the fits from Section 11.2.2 do not include systematic uncertainties. If only tag side quantities are considered, these uncertainties generally only describe a uniform shift in the calibration factors, which is of lesser interest to the discussion of this chapter. However, if the signal flavor is examined, the $\text{B} \to \text{D}^{(*)}\ell\nu_\ell$ crossfeed and the respective branching ratio uncertainties can become relevant.

Finally, Figure 11.20 shows plots of the ITF and CTF calibration factors vs. $\mathcal{P}_{\text{FEI}}$ on the inclusive dataset. The results of MMS for the last bins in $\mathcal{P}_{\text{FEI}}$ are questionable since f_{ITF} becomes very small (and thus also $f_{\text{ITF}} - f'_{\text{ITF}}$, which leads to numerical instability, see Equation (11.3.2)). Excluding these results, the results are relatively consistent. The dependency of the calibration factors on $\mathcal{P}_{\text{FEI}}$ seems to be generally of opposite direction between ITF and CTF and for the tag flavors. In particular, the ratio $\epsilon_{\text{ITF}}/\epsilon_{\text{CTF}}$ is not constant.

11.4 Summary

The calibration factors derived from $\text{B} \to X\ell\nu_\ell$ appear reasonable when applied to the $\text{B} \to \text{D}\ell\nu_\ell$ dataset if only the flavor reconstructed on the tag side is considered. However, discrepancies are observed if the generator level flavor or the flavor reconstructed on the signal side is examined.

Prompted by this very recent observation, we investigate possible discrepancies between the calibration factors for correctly and incorrectly reconstructed tag flavors, ϵ_{CTF}, and ϵ_{ITF}.

[6] Because of the different $\mathcal{P}_{\text{FEI}}$ distributions of the inclusive and exclusive datasets, this is not guaranteed. Therefore, this observation is not a closure check.

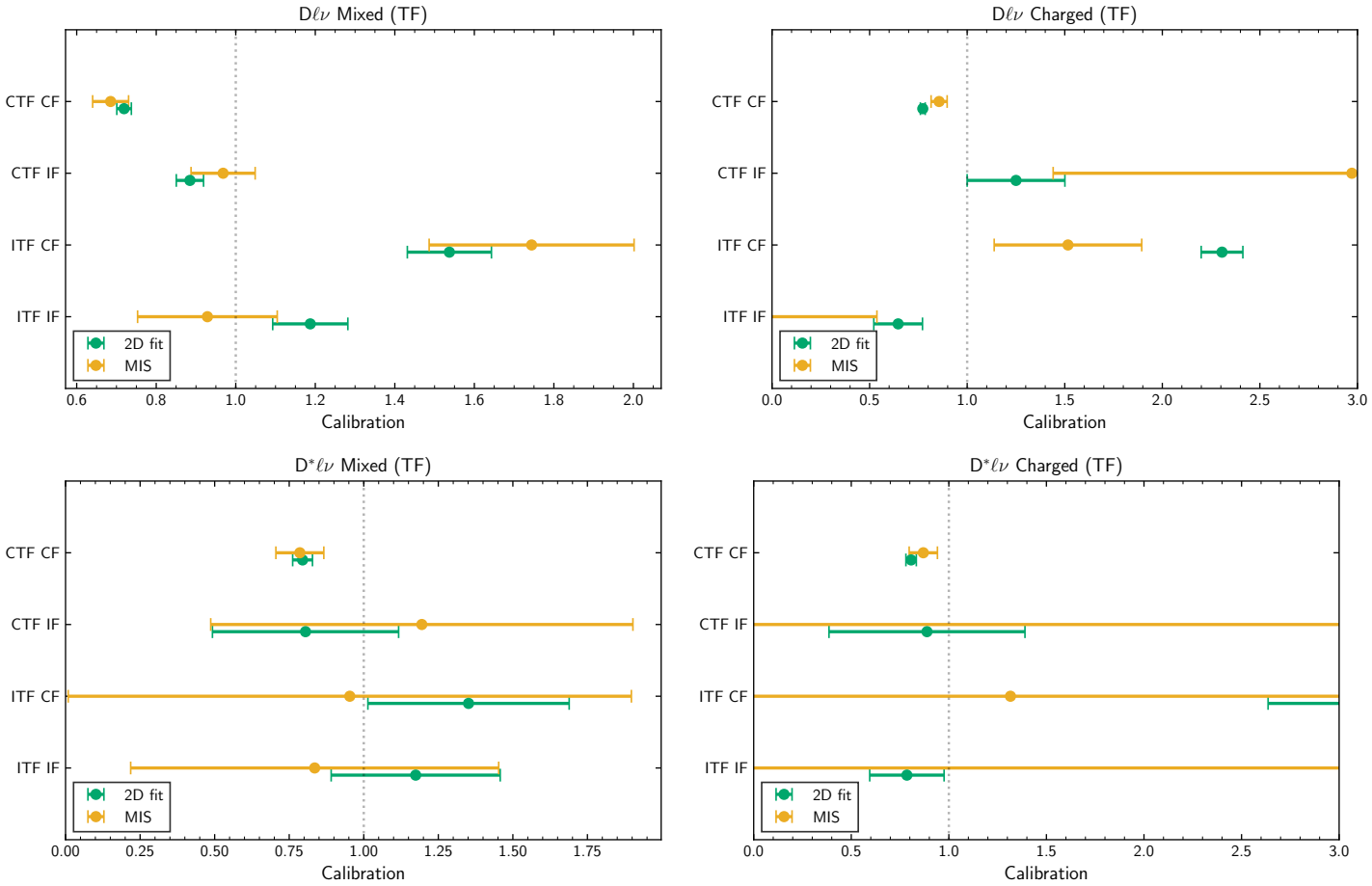


Figure 11.19: Comparing global ITF and CTF calibration factors for IF and CF.

First, we divide the exclusive dataset by the compatibility of reconstructed signal and tag side flavor and extract corresponding calibration factors ϵ_{IF} and ϵ_{CF}. Using a linear system of equations, values for ϵ_{CTF} and ϵ_{ITF} can be calculated. The ratios $\epsilon_{\text{ITF}}/\epsilon_{\text{CTF}}$ appear independent of $\mathcal{P}_{\text{FEI}}$ and can be used to correct the $\text{B} \to X\ell\nu_\ell$ calibration factors. Using only input of $\text{B} \to \text{D}\ell\nu_\ell$ to determine the ratios allows to use the corrected calibration factors for the $\text{B} \to \text{D}^*\ell\nu_\ell$ measurement.

A different avenue is to directly measure ϵ_{CTF} and ϵ_{ITF} using the $m_{\text{bc}}^{\text{tag}}$ observable. Four different strategies are compared. Since the currently available dataset includes a high lower bound on $m_{\text{bc}}^{\text{tag}}$, the available discriminatory power is limited and results in large uncertainties and correlations. Nonetheless, the clear conclusion is that ϵ_{ITF} and ϵ_{CTF} differ immensely, with the ratio $\epsilon_{\text{ITF}}/\epsilon_{\text{CTF}}$ reaching values of up to 2.

While the effects discovered in this chapter might have affected the FR to some extent, its purity at maximum efficiency is generally higher than that of the FEI (see Section 4.1.1.1) such that the effects would have been less prominent. As already mentioned in Section 4.1.1.1, the frequently presented comparison of purities and efficiencies on data also does not distinguish between correctly and incorrectly reconstructed B_{tag} flavors. Therefore, this difference between the two algorithms might be even more significant than typically assumed.

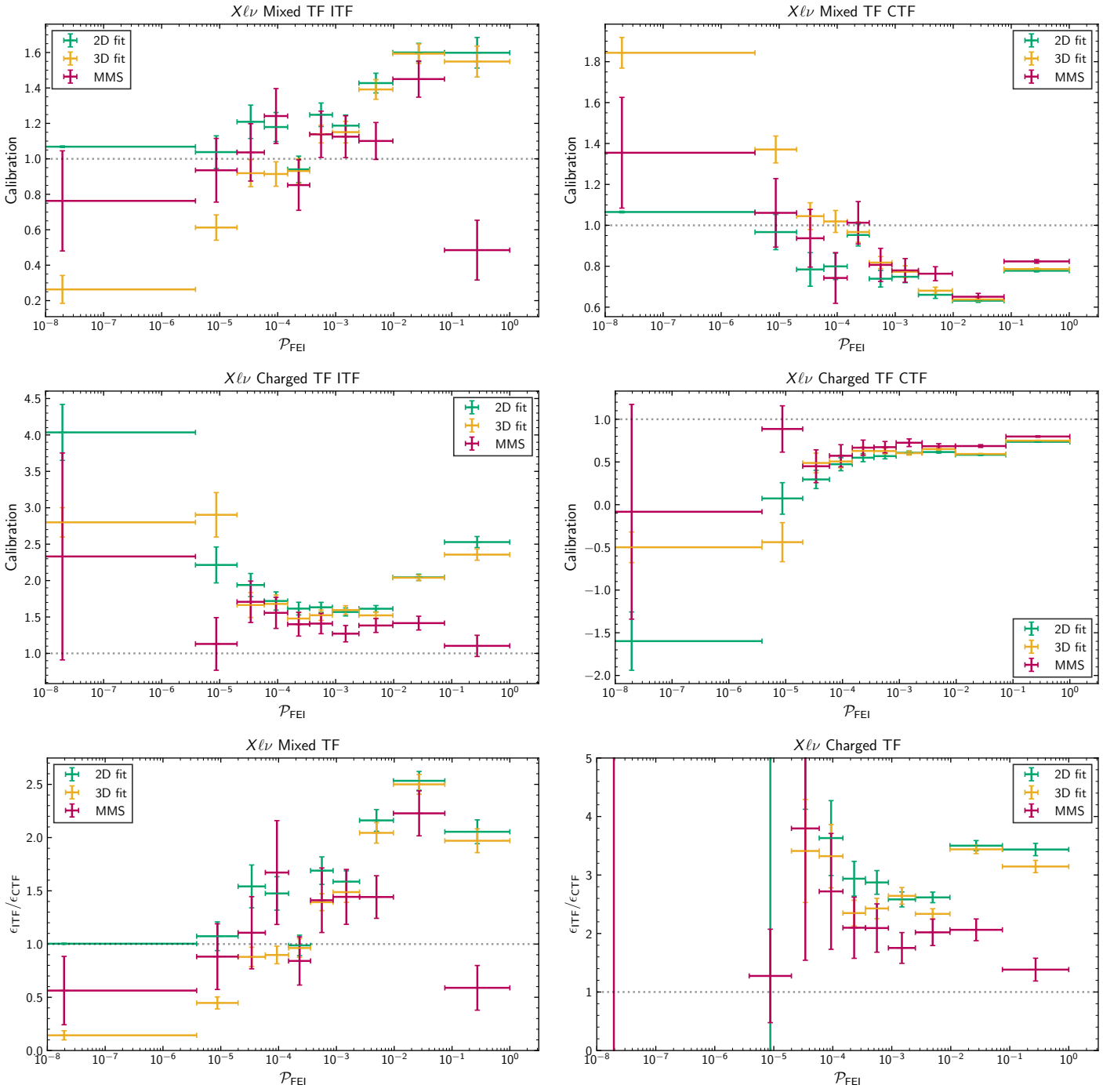


Figure 11.20: Different strategies to calculate separate ITF/CTF calibration factors using the $m_{\mathrm{bc}}^{\mathrm{tag}}$ distribution.

Chapter 12

Summary

The anomalies observed in measurements of $\mathcal{R}(\mathrm{D}^{(*)})$ motivate sensitivity and bias studies regarding the influence of possible new physics (NP) models contributing to the $\mathrm{B} \to \mathrm{D}^{*}\tau\nu_{\tau}$ decay. Clustering the parameter space of such contributions according to the resulting kinematic distributions allows identifying a limited number of benchmark points that represent the distinguishable NP scenarios. The number of required benchmark points is representative of the sensitivity of the considered observables to different NP models. This information can be used to optimize observables either for maximum model independence or maximum discriminatory power. In addition, benchmark points allow to simplify and compare the results of experimental and theoretical analyses.

In order to perform such clustering studies for the first time for the flavor anomalies, I present the open-source package `ClusterKinG`. Together with my collaborators, I implement relevant observables and analyze $\mathrm{B} \to \mathrm{D}^{*}\tau\nu_{\tau}$ decays. Our study characterizes the sensitivity and complementarity of different observables and partially explains the model dependency of experimental results for $\mathcal{R}(\mathrm{D}^{(*)})$.

After highlighting personal contributions to the coordination of software education activities at Belle II and at the HEP Software Foundation, I show preparations for a study of hadronically tagged $\mathrm{B} \to \mathrm{D}^{*}\ell\nu_{\ell}$ decays with the Belle dataset. This analysis improves on previous studies by using the new software framework developed for the Belle II collaboration and considering both B flavors in the reconstruction.

The $\mathrm{B} \to \mathrm{D}^{*}\ell\nu_{\ell}$ decays are used as the normalization mode for $\mathcal{R}(\mathrm{D}^{(*)})$ measurements. Therefore, ensuring a detailed understanding of this decay is essential for the continuing investigation of the flavor anomalies. Moreover, the decay allows measuring $|V_{\mathrm{cb}}|$ and hadronic form factors, which also serve as ingredients for improving theoretical predictions of $\mathcal{R}(\mathrm{D}^{(*)})$ and reducing uncertainties in other experimental analyses.

As demonstrated by fits to Asimov data and dedicated sideband studies, all elements of the analysis are ready for the unblinding of the data. The last building block yet to be approved is related to the Full Event Interpretation (FEI), a machine learning algorithm for the reconstruction of the tag side decays. Its usage is responsible for a substantial part of the increased sensitivity of the $\mathrm{B} \to \mathrm{D}^{*}\ell\nu_{\ell}$ analysis. However, as its performance differs between data and MC simulation, a calibration is required.

This calibration also depends on the FEI classifier output, which induces correlations to observables commonly used for background subtraction. Applying the calibration factors significantly improves the goodness of fit for m^{2}_{miss} fits performed in the $\mathrm{B} \to \mathrm{D}^{*}\ell\nu_{\ell}$ analysis. Similar benefits are expected for an upcoming $\mathcal{R}(\mathrm{D}^{(*)})$ analysis with the Belle dataset.

For the measurement of calibration factors, I analyze hadronically tagged $X\ell\nu_{\ell}$ decays and

compare the signal yield between data and MC. Assuming that the signal side reconstruction efficiencies are well understood, all efficiency differences can be attributed to the tag side. This allows calculating calibration factors that can be applied to analyses that utilize the FEI.

The calibration is challenging for multiple reasons: The MC dataset initially produced for the Belle experiment requires numerous corrections to match today's understanding. The updates particularly affect D^{**} decays and the modeling of the gap, a mixture of incompletely measured decays that are assumed to contribute to $X\ell\nu_\ell$. The loose selection criteria of $X\ell\nu_\ell$ result in a very large dataset requiring state-of-the-art big data analysis tools but allowing a very granular calibration. This high granularity is shown to be important: The required calibration depends significantly on both reconstruction modes and the classifier output.

The investigation presented in this thesis is more detailed than any previous calibration study of the FEI or its predecessor. Besides studying the $\mathcal{P}_{\mathrm{FEI}}$ dependency, the reconstruction modes are investigated beyond the immediate decay products, and significant differences are observed. Additionally, the fundamental assumptions of the calibration procedure are probed by several side studies. Various checks are performed to test the validity of the results. In particular, applying the calibration framework to a dataset of $\mathrm{B} \to \mathrm{D}\ell\nu_\ell$ decays confirms that the calibration factors measured on $\mathrm{B} \to X\ell\nu_\ell$ are reasonable.

However, the last chapter shows that the calibration differs significantly between $\mathrm{B}_{\mathrm{tag}}$ mesons of correctly and incorrectly reconstructed flavors. Since the application of completeness constraints in exclusive measurements changes the $\mathrm{B}_{\mathrm{tag}}$ purity, an additional correction is necessary. I present several strategies to perform this correction. In particular, I show that two ratios obtained from the $\mathrm{B} \to \mathrm{D}\ell\nu_\ell$ dataset are sufficient to correct the calibration factors described above. I also present preliminary results for a separate calibration of correctly and incorrectly reconstructed flavor tags using the $m_{\mathrm{bc}}^{\mathrm{tag}}$ observable.

The insights about the calibration have important consequences for many analyses at Belle and Belle II that make use of the FEI. With the steadily increasing amount of available data at Belle II, the calibration uncertainty will become dominant in many analyses, calling for detailed calibration studies to consider correct and incorrect flavor tags separately. In the light of these projections, future developments of the FEI should also consider prioritizing the reductions of bias over additional efficiency gains.

List of Publications

Papers

J. Aebischer, T. Kuhr, and K. Lieret. "Clustering of $\overline{B} \to D^{(*)}\tau^{-}\overline{\nu}_{\tau}$ kinematic distributions with ClusterKinG". In: *Journal of High Energy Physics* Apr. 2020. ISSN: 1029-8479. DOI: 10.1007/JHEP04(2020)007.

Conference papers

J. Kahn, E. Dorigatti, K. Lieret, A. Lindner, and T. Kuhr. "Selective background Monte Carlo simulation at Belle II". In: *EPJ Web of Conferences* Nov. 2020. DOI: 10.1051/epjconf/202024502028

S. Malik, S. Meehan, K. Lieret et al. "Software Training in HEP". In: "Computing and Software for Big Science" Dec. 2021. ISSN: 2510-2036, 2510-2044. DOI: 10.1007/s41781-021-00069-9.

In review

K. Lieret et al. "A new Software Training Model at Belle II". Submitted to *Journal Of Physics: Conference Series*.

Appendix A

Supplementary material for the clustering studies

A.1 ClusterKinG example

The following code example shows a full `ClusterKinG` workflow: Kinematic distributions for $\mathrm{dBR}(\mathrm{B}^+ \to \mathrm{D}^0 \tau^- \bar{\nu}_\tau)/\,\mathrm{d}q^2$ are generated for a selection of sample points in Wilson space. Uncertainties are added to the data, and the distributions are clustered. Finally, BPs are selected, and several plots are generated. Similar (and more systematic) examples can be found in the `examples` folder of the `ClusterKinG` repository [173].

```python
import flavio
import numpy as np
import clusterking as ck
from clusterking.maths.metric import chi2_metric


# Define kinematic function using the flavio package
def dBrdq2(w, q):
    return flavio.np_prediction("dBR/dq2(B+->Dtaunu)", w, q)


# Set up and configure Scanner
s = ck.scan.WilsonScanner(scale=5, eft="WET", basis="flavio")
# Set kinematic function
s.set_dfunction(
    dBrdq2,
    binning=np.linspace(3.2, 11.6, 10),
    normalize=True
)
# Set sampling points in Wilson space
s.set_spoints_equidist({
    "CVL_bctaunutau": (-0.5, 0.5, 10),
    "CSL_bctaunutau": (-0.5, 0.5, 10),
    "CT_bctaunutau": (-0.1, 0.1, 10)
})

# Run scanner and add errors
d = ck.DataWithErrors()     # Create data object to write results to
r = s.run(d)                # Run scanner
r.write()                   # Write results back to data object
d.add_err_poisson(1000)     # statistical uncertainties
d.add_rel_err_uncorr(0.1)   # 10% relative system uncertainties, uncorrelated

# Clustering
c = ck.cluster.HierarchyCluster()  # Initialize worker class
```

```
c.set_metric(chi2_metric)
c.set_max_d(1)    # "Cut off" value for hierarchy
r = c.run(d)      # Run clustering on d
r.write()         # Write results back to data object

# Benchmarking
b = ck.Benchmark()  # Initialize worker class
b.set_metric(chi2_metric)
r = b.run(d)        # Run benchmarking
r.write()           # Write results back to data object

# Optional: Save data (kinematic distributions, clusters, BPs, ...)
d.write("btaunu_q2.sql")

# Find closest benchmark point to new parameter point
# Similar function for spoints: find_closest_spoints
d.find_closest_bpoints(
        {
                "CVL_bctaunutau": 0.1,
                "CSL_bctaunutau": 0.1,
                "CT_bctaunutau": 0.1,
        },
        n=1
)

# Generate plots
d.plot_clusters_scatter(["CVL_bctaunutau", "CSL_bctaunutau"])
d.plot_dist_box()
d.plot_clusters_scatter(["CT_bctaunutau", "CVL_bctaunutau", "CSL_bctaunutau"])
d.plot_dist()
d.plot_bpoint_distance_matrix()
```

A.2 Differential decay rates

The differential decay rates for the decays $\bar{\mathrm{B}} \to \mathrm{D}^{0(*)}\tau^-\bar{\nu}_\tau$ are given by:

$$\frac{\mathrm{d}^2\Gamma(\mathrm{B}\to \mathrm{D}^{0(*)}\tau\nu)}{\mathrm{d}q^2\,\mathrm{d}(\cos\theta_\ell)} = \frac{3}{8}(I_1^c + 2I_1^s) + \cos(\theta_\ell)\frac{3}{8}(I_6^c + 2I_6^s) + \cos(2\theta_\ell)\frac{3}{8}(I_2^c + 2I_2^s)\,, \tag{A.2.1}$$

$$\frac{\mathrm{d}^2\Gamma(\mathrm{B}\to \mathrm{D}^{*0}\tau\nu)}{\mathrm{d}q^2\,\mathrm{d}(\cos\theta_V)} = -\cos^2(\theta_V)\frac{3}{8}(-3I_1^c + I_2^c) - \sin^2(\theta_V)\frac{3}{8}(-3I_1^s + I_2^s)\,, \tag{A.2.2}$$

$$\frac{\mathrm{d}\Gamma(\mathrm{B}\to \mathrm{D}^{*0}\tau\nu)}{\mathrm{d}q^2} = \frac{3}{4}(2I_1^s + I_1^c) - \frac{1}{4}(2I_2^s + I_2^c)\,, \tag{A.2.3}$$

where q^2 is the invariant mass of the lepton pair and I_i^j denote the angular coefficients, which depend on q^2 and the Wilson coefficients. The angle θ_ℓ is defined as the angle between the direction of the tau in the dilepton rest frame and the direction of the dilepton in the B rest frame, whereas θ_V is defined as the angle between the direction of the D^{*0} meson in the dilepton rest frame and the direction of the D^{*0} in the B rest frame. A general discussion of the kinematics of semileptonic meson decays (the context of $\mathrm{B}\to\mathrm{K}^*\bar{\ell}\ell$) together with the angular coefficients can be found in [346]. For $\bar{\mathrm{B}} \to \mathrm{D}^{0(*)}\tau^-\bar{\nu}_\tau$ decays the corresponding distributions are discussed for example in [185, 347]. Our implementation of the above decay rates has been made available as open-source software at [174].

A.3 Validation of statistical treatment

The statistical treatment of the examples shown in Section 2.4 is validated with toy experiments: For each point in parameter space, we consider the corresponding histogram and its covariance

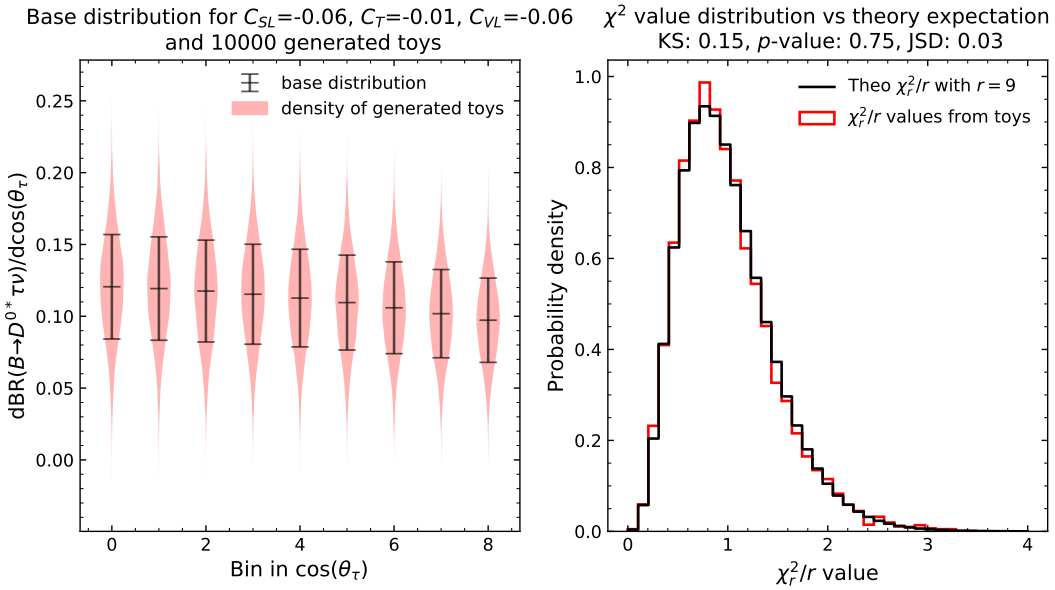


Figure A.1: Toy experiments to validate the implementation of the χ^2 metric. The right sided figure also reports several values quantifying the similarity of the toy distribution to the theoretical expectation: the Kolmogorov-Smirnov test statistic (KS), its corresponding p-value and the Jensen-Shannon Divergence (JSD).

matrix. Toy histograms are generated by drawing random values from the multivariate normal distribution with matching means and covariance matrix. We then calculate the test statistic χ^2 from (2.2.1) between each toy histogram and the original histogram. The distribution of all $\chi^2/(N-1)$ values is binned and compared to the calculated expected distribution $\chi^2_{N-1}/(N-1)$ using the Jenson-Shannon Divergence (JSD).

An example for one particular point in parameter space is shown in Figure A.1. Both histograms agree nicely, resulting in a low JSD value. The result of repeating the same procedure across all points is shown in Figure A.2, showing satisfactorily low divergence values.

Additional code to reproduce the figures shown here and to validate the statistical treatment has been added to the `ClusterKinG` repository [173].

A.4 Individual contributions

Because the paper and the results presented throughout Chapter 2 have been the result of a collaboration, this section gives a summary of the individual contributions.

The initial idea for the project was suggested by Alejandro Celis, who worked on several early proofs of concepts, but then left the project early on to pursue a career outside of academia.

All of the remaining development of the `ClusterKinG` package and its physics package has been performed by K. L.[1] The work on all software components is completely traceable via the commit history of the repositories [173, 174].

The physical examples shown in the paper have been developed by K. L. and Jason Aebischer (J. A.), who has in particular guided the selection of the considered observables and their appropriate ranges.

Likewise, the text of the paper has been written by both K. L. and J. A. J. A. has particularly

[1] This excludes 11 commits by Jason Aebischer that only affected the readme file.

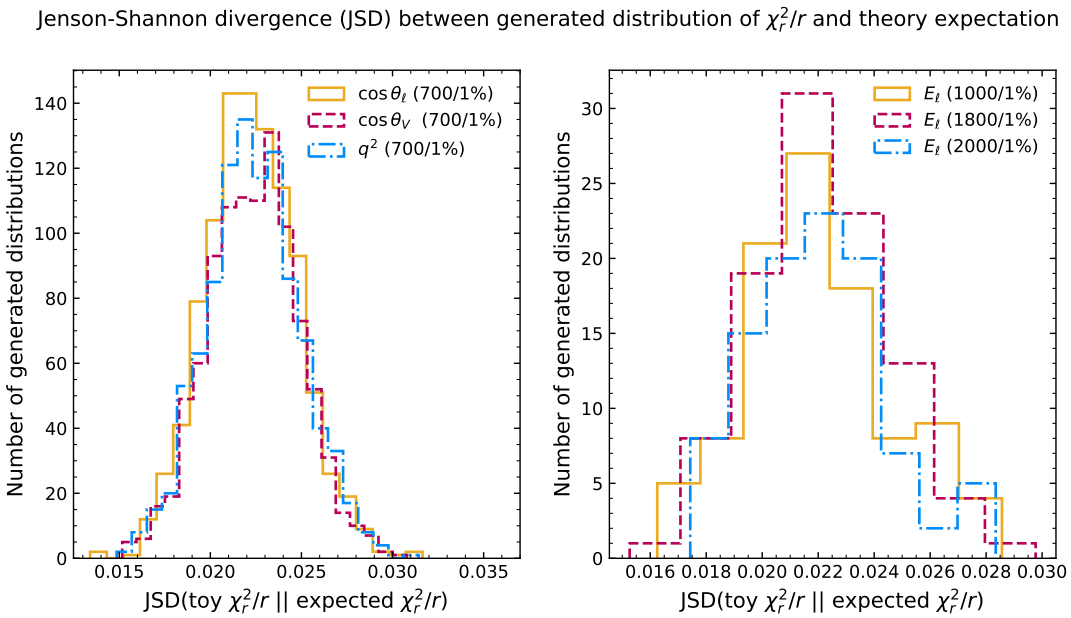


Figure A.2: Validating the shape of the χ^2 distribution for all points in the parameter space. The numbers in parentheses denote the assumed total yield corresponding to the Poisson uncertainties and the uncorrelated systematic uncertainty.

contributed to the theoretical and phenomenological background and K. L. to the technical details, the description of the algorithm, the statistical treatment, and the interpretation of the results.

Finally, Thomas Kuhr has supported this work with general guidance and many helpful comments.

Appendix B

Evaluation of systematic uncertainties for the calibration

This chapter describes two side-studies that lead to systematic uncertainties:

- *Categorization side study* (Appendix B.1): Is our categorization granular enough to avoid any bias from the different $\mathcal{P}_{\text{FEI}}$ or C_{FEI} distributions of $B \to D^* \ell \nu_\ell$? We investigate this question by comparing weighted averages of the calibration factors when considering more sub-decay channels or more bins in $\mathcal{P}_{\text{FEI}}$. We assign systematic uncertainties on the sub-percent level.
- *Signal track multiplicity study* (Appendix B.2): Do different signal side track counts impact the calibration factor? If the calibration factors are independent of the signal side (other than by correlations with $\mathcal{P}_{\text{FEI}}$ and C_{FEI}), no effect should be observed. This study therefore tests one of the important assumptions of the calibration procedure. We assign systematic uncertainties on the sub-percent level.

Additional side-studies that do not conclude with a systematic uncertainty are shown in Appendix C.

B.1 Categorization

As already mentioned in Section 9.1, choosing too coarse categories can lead to biased calibration factors. Because we already consider every decay channel separately, this can only be an issue for our consideration of sub-decay channels or binning in $\mathcal{P}_{\text{FEI}}$.

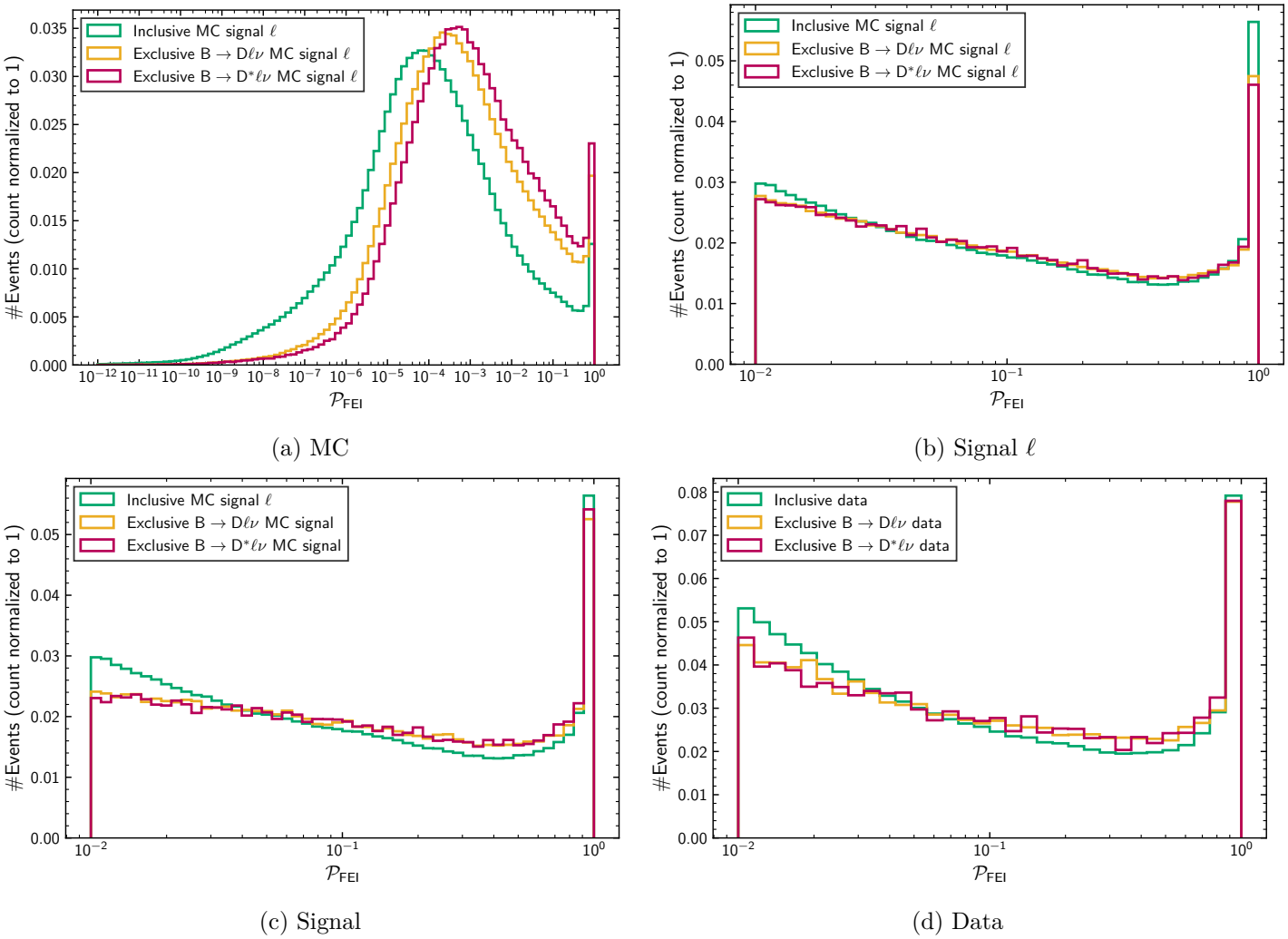
B.1.1 Some qualitative observations

For simplicity, the following arguments and conditions will be formulated about $\mathcal{P}_{\text{FEI}}$, but analog statements can be made about sub-decay channels (but we will see that they have less impact on the result). Considering $\mathcal{P}_{\text{FEI}}$, we have a problem if, for a category,

i. the real calibration factor depends very strongly on the $\mathcal{P}_{\text{FEI}}$ distribution in the region given by the $\mathcal{P}_{\text{FEI}}$ cut applied for the category (significant differences between the calibration factors of neighboring categories within the same decay channel are a symptom of this)

ii. the distribution of $\mathcal{P}_{\text{FEI}}$ in this region is very different between the inclusive and exclusive dataset, particularly in the signal/signal lepton component.

We know from Figure 6.1 that the $\mathcal{P}_{\text{FEI}}$ distributions of the inclusive and exclusive dataset are indeed very different: Because of their cleaner reconstruction that also affects the tag side



(a) MC

(b) Signal ℓ

(c) Signal

(d) Data

Figure B.1: $\mathcal{P}_{\text{FEI}}$ distributions. For the exclusive datasets, *signal* ℓ refers to the signal component as defined in Section 9.2.4, whereas *signal* refers to matching of the full decay $B \to D^{(*)}\ell\nu_\ell$.

reconstruction and selection, the $\mathcal{P}_{\text{FEI}}$ distributions of the exclusive datasets are shifted to the right. This effect is even more substantial when considering the MC distribution in the signal lepton component (i.e., with the signal component as defined in Section 9.2.4). This is shown in Figure B.1a and Figure D.9.

However, inspecting the high $\mathcal{P}_{\text{FEI}}$ region, it turns out that the shapes become very similar: Figure B.1b shows this for the MC signal component, Figure B.1c even more specifically for correctly reconstructed exclusive decays and Figure B.1d for data.

Thus condition i is mostly fulfilled for high $\mathcal{P}_{\text{FEI}}$, but condition ii is not fulfilled in these regions.

While this makes it plausible that large channels like $\overline{D}^0\pi^+\pi^+\pi^-\pi^0$ are categorized appropriately (despite the drop at the end), there is still the issue of the numerous small channels that have few to none bins in $\mathcal{P}_{\text{FEI}}$.

B.1.2 Comparing averaged results with different categorizations

In order to quantify any bias that we pick up, we calculate calibration factors for different categorizations and compare their weighted average using the relative weights of the categories in $B \to D^*\ell\nu_\ell$ signal MC. For simplicity, we only use the resonant gap model and the fallback fits. We also do not consider any systematic uncertainties (since they are assumed to be correlated

across categories).

Table B.1 summarizes the different categorizations (see Section 9.1 for the definition of the categories) and the results. Rather than specifying c_1, we list $\lceil N_1/c_1 \rceil$, the maximal number of $\mathcal{P}_{\text{FEI}}$ bins used for channels with few events ($n \leq N_1$).

B.1.2.1 Qualitative description

In the first group of results (A–D), both N_{sdc} (and the other parameters) are varied. Categorization A is what has been used so far in this note.

The binning in the sub decay channels is at its finest in results D and its coarsest in results A. The relative difference $((\bar{\epsilon}_{\text{A}} - \bar{\epsilon}_{\text{D}})/\bar{\epsilon}_{\text{A}})$ between both results is 0.1 % (B^0) and −0.4 % (B^+). For B^0, this is within the statistical uncertainties. Note that the difference cannot simply be attributed to the sub-decay channel granularity because (with the current setup), the division in sub-decay channels always changes the granularity in $\mathcal{P}_{\text{FEI}}$ (though this could be investigated with a fixed $\mathcal{P}_{\text{FEI}}$ binning side-study).

Looking now at the dependency on $\mathcal{P}_{\text{FEI}}$, we compare A and E–J. All of them have the same N_{sdc}, that is, equal granularity in sub-decay channels. Except for H (which uses a very fine and uniform division in $\mathcal{P}_{\text{FEI}}$), N_1 is also shared. Ordering by increasing granularity in $\mathcal{P}_{\text{FEI}}$, we can compare A, E, F, G, I, H. No clear dependency is apparent: For B^0, the average calibration factor mostly seems to decrease with increasing granularity. However, I is certainly more granular than G but shows an increase in the calibration factor. For B^+, the behavior is even less clear.

B.1.2.2 Systematic uncertainty

For B^0, categorization A leads to the maximal value, G to the minimal value. We thus assign a relative uncertainty of $(\bar{\epsilon}_{\text{A}} - \bar{\epsilon}_{\text{G}})/\bar{\epsilon}_{\text{A}} \sim 0.5\,\%$.

For B^+, the result for B is most different from A, so we assign a relative uncertainty of $(\bar{\epsilon}_{\text{A}} - \bar{\epsilon}_{\text{B}})/\bar{\epsilon}_{\text{A}} \sim 0.6\,\%$.

To summarize, we assign

$$e^{\text{corr}}_{\text{cat}} = \begin{cases} 0.5\,\% & \text{B}^0 \\ 0.6\,\% & \text{B}^+. \end{cases} \tag{B.1.1}$$

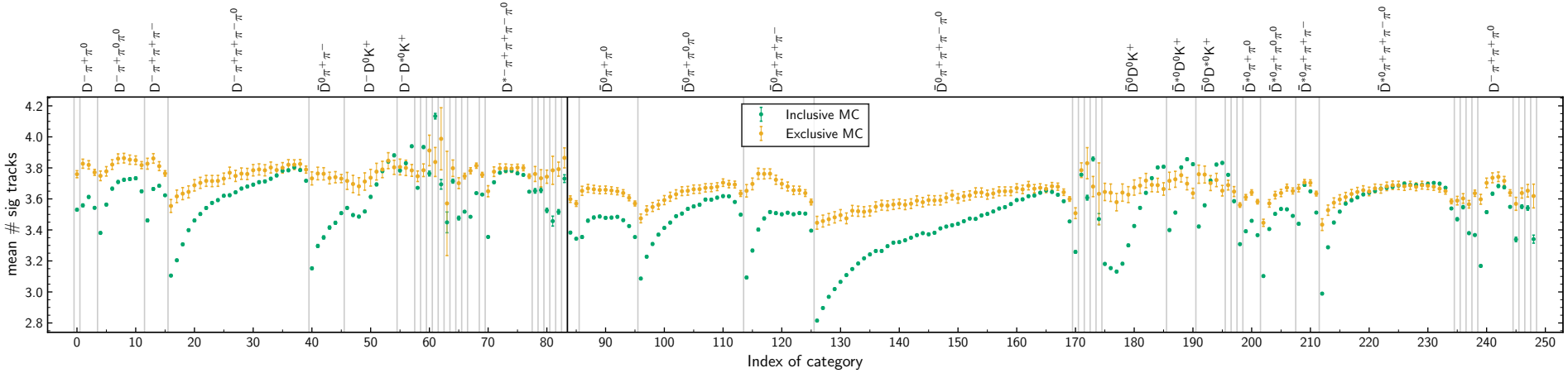
B.2 Calibration factors and signal side tracks

The key question that is also investigated in Appendix C.3 is: Does the calibration factor depend on the reconstruction on the signal side? This side-study investigates whether the number of tracks on the signal side ($n^{\text{sig}}_{\text{track}}$) influences the calibration factor. For example, many tracks on the signal side might give the FEI more incorrect track candidates to use in reconstruction and change the reconstruction efficiency in MC and data slightly differently. This is of importance, because the number of tracks is very different in the inclusive dataset used for the determination of the calibration factors and the exclusive dataset used to measure $|V_{\text{cb}}|$ (Figure B.2a and even more pronounced in Figure B.3).

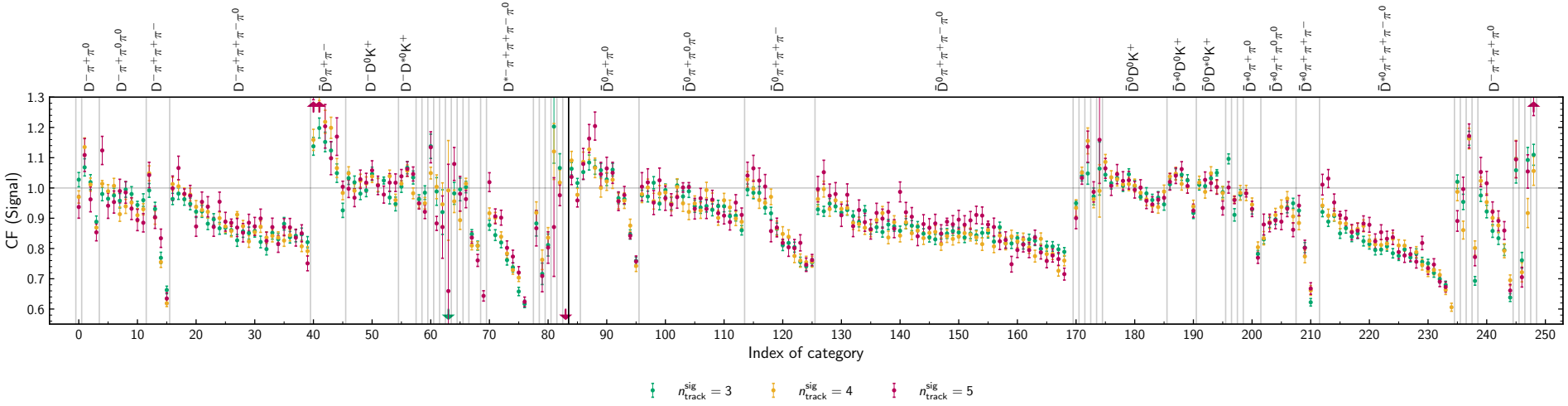
To check for any such effect, the inclusive dataset (both data and MC) is split up into five parts, for each of which we calculate a calibration factor as before:

(1) $n^{\text{sig}}_{\text{track}} \leq 2$

(2) $n^{\text{sig}}_{\text{track}} = 3$



(a) Average number of signal side tracks in the inclusive and exclusive datasets



(b) Calibration factors for $n_{\text{track}}^{\text{sig}} = 3, 4, 5$. For an aggregated comparison, see Figure B.4.

Figure B.2: Average number of signal side tracks and calibration factors for $n_{\text{track}}^{\text{sig}} = 3, 4, 5$.

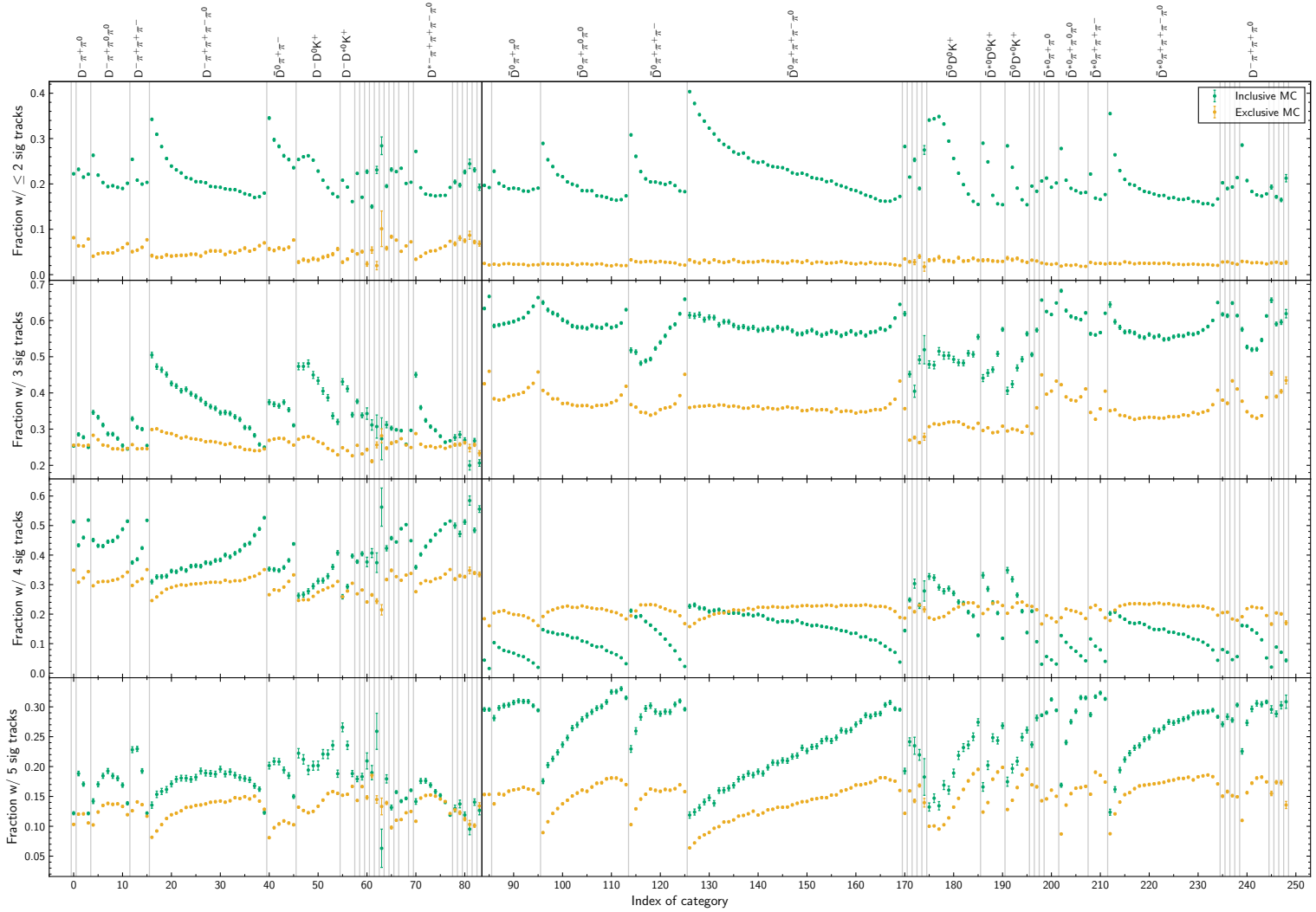


Figure B.3: Number of signal side tracks in the inclusive and exclusive datasets (detailed)

Categorization	N_{sdc}	N_1	$\lceil N_1/c_1 \rceil$	c_2	N_{cat}	B^0	B^+	Both
A	20.0	3.0	3	3.0	232	0.8954 ± 0.0022	0.9000 ± 0.0015	0.8988 ± 0.0013
B	3.5	4.0	4	3.5	340	0.8947 ± 0.0023	0.9051 ± 0.0015	0.9024 ± 0.0013
C	10.0	3.0	6	1.5	458	0.8916 ± 0.0023	0.9015 ± 0.0015	0.8989 ± 0.0013
D	3.0	3.0	4	2.0	413	0.8941 ± 0.0023	0.9035 ± 0.0015	0.9010 ± 0.0013
E	20.0	3.0	3	2.0	315	0.8933 ± 0.0023	0.9017 ± 0.0015	0.8995 ± 0.0013
F	20.0	3.0	6	1.5	458	0.8914 ± 0.0023	0.9010 ± 0.0015	0.8985 ± 0.0013
G	20.0	3.0	10	1.0	702	0.8907 ± 0.0023	0.8989 ± 0.0015	0.8968 ± 0.0013
H	20.0	∞	∞	0.3	1986	0.8904 ± 0.0023	0.8943 ± 0.0015	0.8933 ± 0.0013
I	20.0	3.0	20	1.0	949	0.8940 ± 0.0023	0.8991 ± 0.0015	0.8978 ± 0.0013
J	20.0	3.0	10	2.0	481	0.8914 ± 0.0023	0.8998 ± 0.0015	0.8976 ± 0.0013

Table B.1: Weighted averages of calibration factors based on different categorizations. N_{sdc}, N_1 and c_2 are specified in units of 10^4 events. See text for description of fits and averaging weights.

(3) $n_{\text{track}}^{\text{sig}} = 4$

(4) $n_{\text{track}}^{\text{sig}} = 5$

(5) $n_{\text{track}}^{\text{sig}} \geq 6$

For details regarding the calculation of $n_{\text{track}}^{\text{sig}}$, see Appendix D.5. For this study, the calibration factors were calculated with the 1D $|\vec{p}_\ell^{\,*}|$ fit (which performs better with lower statistics). We furthermore did not consider any of the systematic uncertainties other than MC statistics (as most of them would be shared between the different subsets). For a note regarding the track finding efficiency ratio between data and MC, see Appendix B.2.1.

The calibration factors for $n_{\text{track}}^{\text{sig}} = 3, 4, 5$ are shown in Figure B.2b, an aggregated comparison of the calibration factors is shown in Figure B.4. Relevant for the $|V_{\text{cb}}|$ measurement are $n_{\text{track}}^{\text{sig}} = 3, 4, 5$ (but we need the remaining calibration factors for a proper comparison via Equation (B.2.2)).

The calibration factors for $n_{\text{track}}^{\text{sig}} = 3$ and $n_{\text{track}}^{\text{sig}} = 4$ are well compatible with the calibration factor derived without any $n_{\text{track}}^{\text{sig}}$ cut (p-values > 0.5, see right side of Figure B.4). However, this is not the case for $n_{\text{track}}^{\text{sig}} = 5$.

We want to quantify this issue as an uncertainty on the signal track-averaged calibration factors (as calculated outside this side-study). For this, we will compare two different weighted averages of the signal track-specific calibrations.

Let us first define $\epsilon_i^{(k)}$ ($k = 1, \ldots, 5$; $i = 0, \ldots, N_{\text{cat}} - 1$) as the calibration factor for case k of the five cases enumerated above. Further define $f_i^{\text{incl}\,(k)}$ as the fraction of tracks of this subset in the total number of events in this category in inclusive MC, for example:

$$f_i^{\text{incl}\,(2)} = \frac{\#\{\text{incl. MC events in cat. } i \mid n_{\text{track}}^{\text{sig}} = 3\}}{\#\{\text{incl. MC events in cat. } i\}} \tag{B.2.1}$$

Note that $\sum_{k=1}^{5} f_i^{\text{excl}\,(k)} = 1$, i.e., the $f_i^{\text{incl}\,(k)}$ are (anti)correlated (which is taken into account in the assignment of the uncertainties). Similarly define $f_k^{\text{excl}\,(k)}$ for the fraction of events with a certain number of tracks in the *exclusive* dataset.

We can reproduce the "normal" calibration factor (without any $n^{\text{sig}}_{\text{track}}$ cut) by taking the weighted sum

$$\epsilon_i^{\text{incl}} = \sum_{k=1}^{5} f_i^{\text{incl}\,(k)} \epsilon_i^{(k)}. \tag{B.2.2}$$

However, we can now also build a calibration factor for the application to the exclusive dataset that weights the signal track specific calibration factors $\epsilon_i^{(k)}$ with the fractions of the *exclusive* dataset:

$$\epsilon_i^{\text{excl}} = \sum_{k=1}^{5} f_i^{\text{excl}\,(k)} \epsilon_i^{(k)}. \tag{B.2.3}$$

The reason that we define ϵ_i^{incl} as in (B.2.2), rather than simply calculation calibration factors without any cut on $n^{\text{sig}}_{\text{track}}$ is that in this way we can properly correlate its uncertainties with ϵ_i^{excl}. The two different averages ϵ_i^{excl} and ϵ_i^{incl} are shown in Figure B.5.

We now perform a χ^2 test with

$$\chi^2 = \sum_{i,j=0}^{N_{\text{cat}}-1} \frac{(\epsilon_i^{\text{excl}} - \epsilon_i^{\text{incl}})(\epsilon_j^{\text{excl}} - \epsilon_j^{\text{incl}})}{\text{Cov}(\epsilon_i^{\text{excl}} - \epsilon_i^{\text{incl}}, \epsilon_j^{\text{excl}} - \epsilon_j^{\text{incl}})}. \tag{B.2.4}$$

The initial p-value is very low: The results for ϵ_i^{incl} and ϵ_i^{excl} are systematically different and show a bias (see Figure B.6). To quantify this as a systematic uncertainty, we now add additional uncertainties on ϵ_i^{excl} until the p-values look reasonable:

$$\Delta\epsilon_i^{\text{incl}} \longrightarrow \Delta\epsilon_i^{\text{incl}} + \epsilon_i^{\text{incl}}(e_{\text{corr}} + e_{\text{uncorr}}), \tag{B.2.5}$$

where e_{corr} is fully correlated between all categories i. We furthermore look at one more quantity, which we will call *bias*

$$\text{bias} = \sum_{i=0}^{N_{\text{cat}}-1} \frac{\epsilon_i^{\text{excl}} - \epsilon_i^{\text{incl}}}{\Delta(\epsilon_i^{\text{excl}} - \epsilon_i^{\text{incl}})}, \tag{B.2.6}$$

where again, we take care of the proper correlations between the terms.

The results are shown in Figure B.7. From Figure B.7b, we conclude that a correlated uncertainty of 0.35 % is enough to bring the bias below one (i.e., less than 1σ of tension between the mean and zero). From Figure B.7a, we conclude that for this correlated uncertainty an uncorrelated uncertainty of 0.4 % is enough to bring the p-value to 10 %. Thus, we assign

$$e_{\text{corr}} = 0.35\,\%, \quad e_{\text{uncorr}} = 0.4\,\%. \tag{B.2.7}$$

B.2.1 Track finding efficiency ratio

As already mentioned in Section 9.5.1, the ratio of the track finding efficiency on data and MC, $r_{\text{track}} := \epsilon^{\text{Data}}_{\text{track}}/\epsilon^{\text{MC}}_{\text{track}}$ is compatible with unity only within an uncertainty of 0.35 %. Could the correlated uncertainty e_{corr} that we needed to assign simply be an artifact of $r_{\text{track}} \neq 1$? In this case, we would not need to assign an extra uncertainty because we usually only require a single lepton (for which we have assigned a tracking uncertainty in Section 9.5.1 already) and do not cut on the $n^{\text{sig}}_{\text{track}}$ variable. The signal track multiplicity specific calibration factors that we calculated in this side-study are

$$\frac{N^{\text{Data}}_{\text{sig},i}(n^{\text{sig}}_{\text{track reco}} = k)}{N^{\text{MC}}_{\text{sig},i}(n^{\text{sig}}_{\text{track reco}} = k)} = r^k_{\text{track}} \cdot \frac{N^{\text{Data}}_{\text{sig},i}(n^{\text{sig}}_{\text{track real}} = k)}{N^{\text{MC}}_{\text{sig},i}(n^{\text{sig}}_{\text{track real}} = k)}. \tag{B.2.8}$$

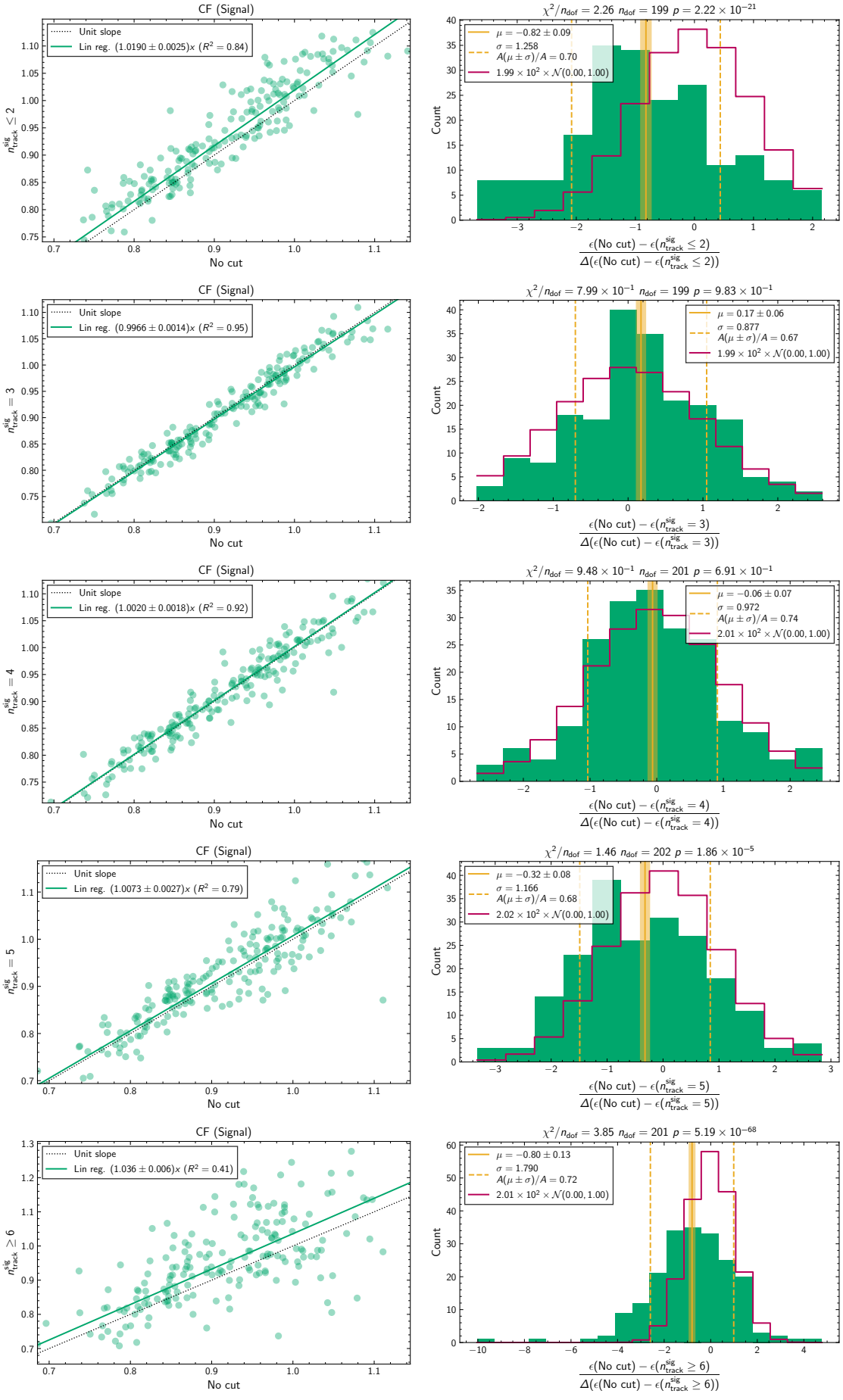


Figure B.4: Comparison of calibration factors derived for different signal side track multiplicities. The pull distributions and χ^2 tests ignore the correlation between the $|\vec{p}^{\,*}_{\ell}| = x$ and *no cut* dataset.

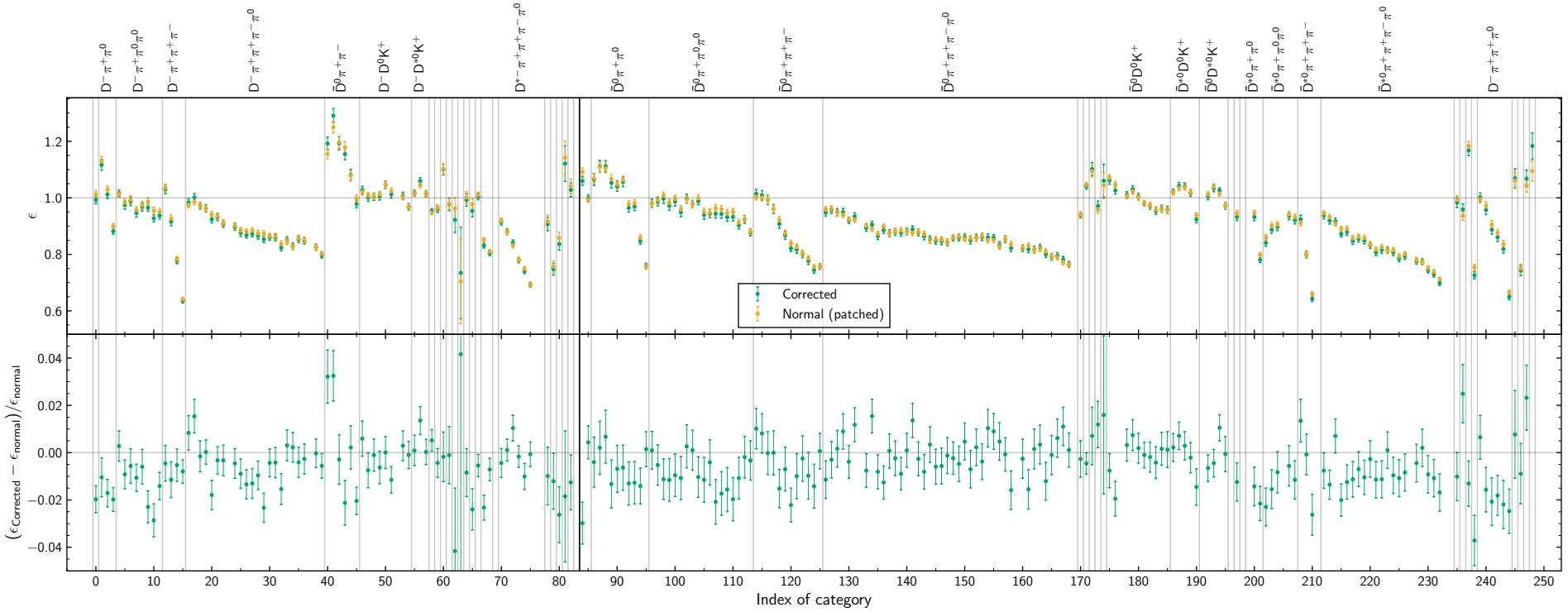
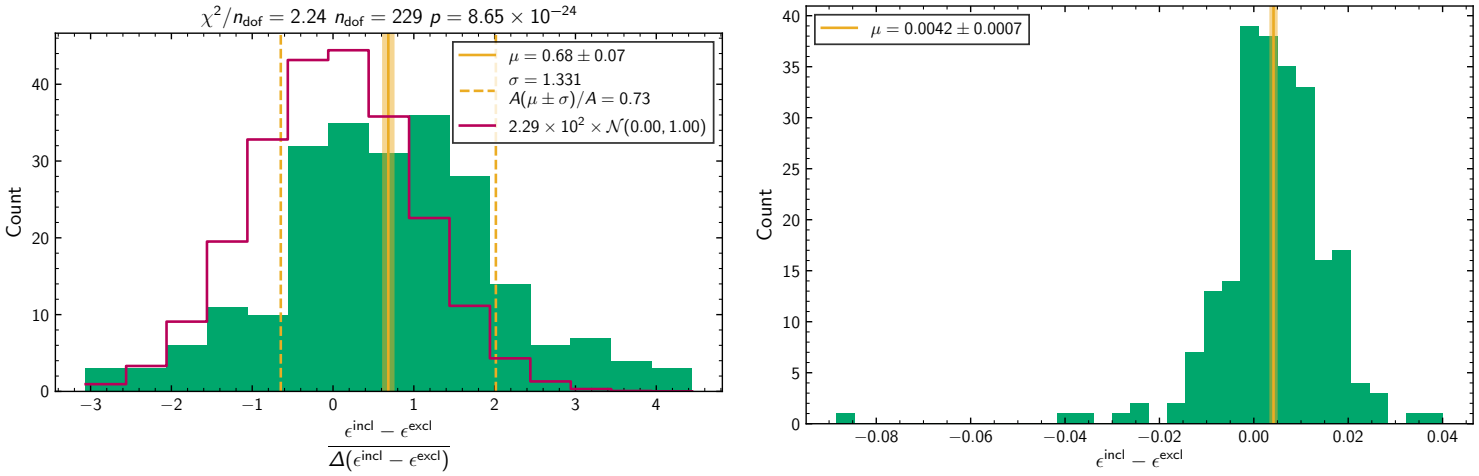


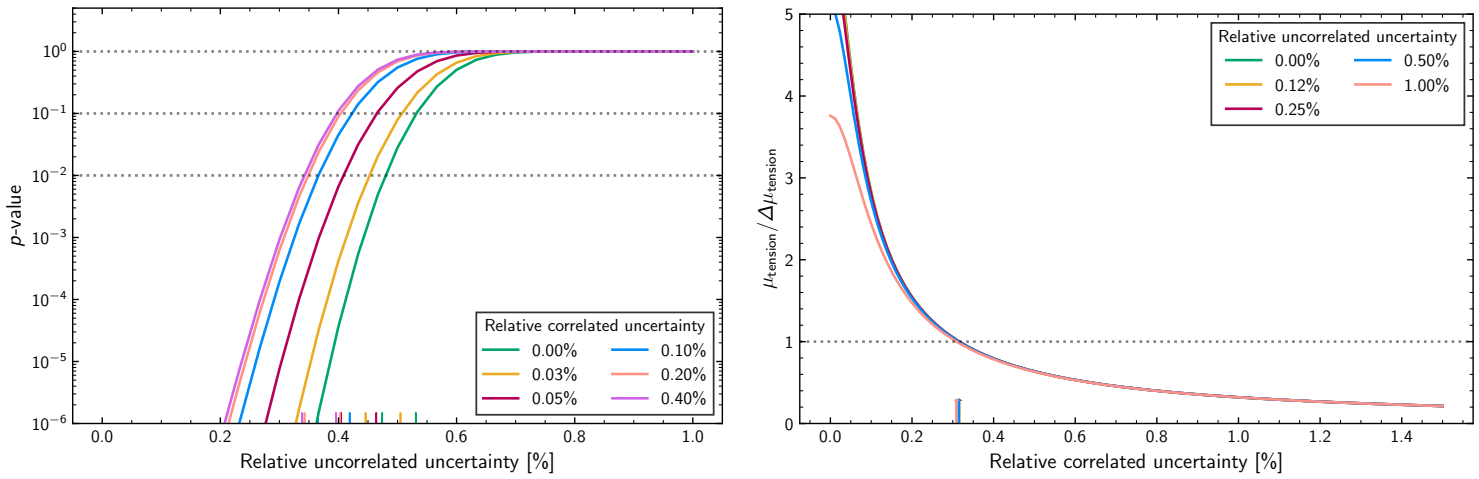
Figure B.5: Comparison of signal track weighted calibration factors ϵ_i^{excl} (denoted *corrected*) and ϵ_i^{incl} (denoted *normal/patched*).



(a) Pull plot

(b) Differences

Figure B.6: Comparing ϵ^{incl} and the signal-track weight corrected ϵ^{excl}. Correlations are taken into account for both the χ^2 test and uncertainties on the mean.



(a) p-value

(b) Bias

Figure B.7: Assigning a systematic uncertainty to the signal side track count dependency of the calibration factors. Note that the meaning of the different colors and the abscissa is switched in both plots. The colored lines on the abscissa denote the intersections with the dashed lines.

Assuming that $N^{\text{Data}}_{\text{sig},i}(n^{\text{sig}}_{\text{track real}} = k)/N^{\text{MC}}_{\text{sig},i}(n^{\text{sig}}_{\text{track real}} = k)$ does not depend on k, we expect to see a factor r between the different calibration factors. However, this is not the case (see Figure B.4): Both very low and high numbers of $n^{\text{sig}}_{\text{track}}$ deviate in the same direction. Thus, the observed effect cannot be explained by a difference in tracking efficiencies.

Appendix C

Closure tests for the calculation of calibration factors

This appendix presents various studies that validate the results presented in Chapter 10. However, as already discussed in Chapter 11, the current set of calibration factors does not account for differences between correctly and incorrectly reconstructed flavors of the B_{tag} meson. These tests are not sensitive to this effect.

Three studies test the physical validity and plausibility of the results:

- Validation with the exclusive $B \to D^{(*)}\ell\nu_\ell$ datasets (Appendix C.1): Are the calibration factors derived from the inclusive dataset reasonable when applied to the exclusive data? To confirm this, we derive calibration factors on the exclusive dataset itself and compare the results with those presented in Chapter 10. The results do indeed look consistent, except for the issue covered in Chapter 11.

- The calibration factors should not depend on the lepton flavor on the $X\ell\nu_\ell$ signal side. This is confirmed in Appendix C.2.

- Are the same events placed into different categories depending on the signal side reconstruction? Appendix C.3 answers this question with a clear *yes*. However, the differences are similar between data and MC. Therefore, they cancel in the calibration factors.

Two more studies test the validity of the implementation:

- In Appendix C.4 toy studies are presented to demonstrate the validity of the fitter on simulated data.

- In Appendix C.5 the calibration factors are applied to the inclusive data (that is, the same dataset with which they were calculated) before performing another fit to the signal component. The calibration factors calculated with this second fit should then be unity (which they are). This is another test for the validity of the implementation. We also show plots of the calibrated spectra in this section.

C.1 Validation of calibration factors with the exclusive datasets

This section compares the calibration factors derived from the inclusive dataset with calibration factors extracted from fits to the exclusive $B \to D^{(*)}\ell\nu_\ell$ dataset. If the assumptions of the tagging calibration are fulfilled, we expect to see similar calibration factors.

- Appendix C.1.1 compares weighted averages of the calibration factors shown in Chapter 10 with the results of calibration fits on the $B \to D^{(*)}\ell\nu_\ell$ datasets. The advantage of this strategy is that the comparison can be done for categories coarse enough to not be subject to large statistical uncertainties from the limited size of the $B \to D^{(*)}\ell\nu_\ell$ datasets.
- Appendix C.1.2 applies the usual calibration separately to the exclusive and inclusive datasets. Because of their limited size both $B \to D\ell\nu_\ell$ and $B \to D^*\ell\nu_\ell$ are combined and additional decay channels are considered. The calibration factors can then directly be compared.

C.1.1 Comparing weighted averages of calibration factors

The most straightforward comparison that we can do is to extract a *global* calibration factor on the exclusive dataset: We perform a single fit for a signal component and divide its yield by the MC expectation. This calibration factor $\bar{\epsilon}_{\text{sig}}$ is equal to the weighted average of calibration factors across categories:

$$\bar{\epsilon}_{\text{sig}} := \frac{N_{\text{sig}}^{\text{Data}}}{N_{\text{sig}}^{\text{MC}}} = \frac{\sum_{i=0}^{N_{\text{cat}}-1} N_{\text{sig},i}^{\text{Data}}}{\sum_{i=0}^{N_{\text{cat}}-1} N_{\text{sig},i}^{\text{MC}}} = \sum_{i=0}^{N_{\text{cat}}-1} \frac{N_{\text{sig},i}^{\text{MC}}}{\sum_{j=0}^{N_{\text{cat}}-1} N_{\text{sig},j}^{\text{MC}}} \frac{N_{\text{sig},i}^{\text{Data}}}{N_{\text{sig},i}^{\text{MC}}} = \sum_{i=0}^{N_{\text{cat}}-1} f_i\, \epsilon_{\text{sig},i}. \tag{C.1.1}$$

This means that we can compare $\bar{\epsilon}_{\text{sig}}^{\text{excl}}$ (extracted with a global fit on the exclusive dataset) with $\bar{\epsilon}_{\text{sig}}^{\text{incl}}$ (calculated from the calibration factors $\epsilon_{\text{sig},i}$ with the weights f_i from the exclusive dataset). If our categorization is of adequate granularity (as discussed in Appendix B.1), then $\bar{\epsilon}_{\text{sig}}^{\text{excl}} \approx \bar{\epsilon}_{\text{sig}}^{\text{incl}}$ should hold even though the $\mathcal{P}_{\text{FEI}}$ and C_{FEI} distributions are different between the datasets.

The results are shown in Table C.1. The table is divided into two sections of three lines:

1. The first line shows calibration factors derived from a $|\vec{p}_\ell^{\,*}|$ fit with five bins to the exclusive dataset (without separation into categories). We use the signal definition of the tagging calibration for the fit. That means that the signal component contains all decays with a correctly reconstructed, primary lepton from a $B\overline{B}$ decay. In particular, the fit does not distinguish between $B \to D\ell\nu_\ell$, $B \to D^*\ell\nu_\ell$ or $B \to D^{**}\ell\nu_\ell$ decays. The fitting procedure is thereby identical to the fallback fits of the tagging calibration (Section 9.2.6). The post-fit distributions are shown in Figure C.1.

   No additive systematic uncertainties are included in the fit. Based on the relative uncertainties of the isospin averaged branching ratios of $B \to D^{(*)}\ell\nu_\ell$, we assume 3 % ($B \to D\ell\nu_\ell$) resp. 2 % ($B \to D^*\ell\nu_\ell$) of systematic uncertainty on the calibration factors. This simplified assignment of systematic uncertainties is very conservative in the sense that it increases the significance of any tension that is found with the calibration factors extracted from $B \to X\ell\nu_\ell$.

2. The second line shows the weighted average of the calibration factors ϵ_{sig}^i that have been presented in Chapter 10 (these averages are therefore identical to the results shown in Table 10.1):

$$\bar{\epsilon} = \sum_{i=0}^{N_{\text{cat}}} \epsilon_{\text{sig}}^i f_i \quad \text{with} \quad f_i := \frac{N_{\text{excl,sig},i}^{\text{MC}}}{N_{\text{excl,sig}}^{\text{MC}}}, \tag{C.1.2}$$

   where $N_{\text{excl},c,i}^{\text{MC}}$ is the number of MC events for category i of the signal lepton component in the exclusive dataset and $N_{\text{excl},c}^{\text{MC}}$ is the sum thereof (such that $\sum_i f_i = 1$).

   The correct correlation of ϵ_{sig}^i is included in the calculation of the uncertainty on $\bar{\epsilon}$. Because the sum contains more than 200 terms, it is essentially reduced to the correlated part of the uncertainty of around 4 %.

	B^0	B^+	Both
$D\ell\nu$ calibration fit	0.90 ± 0.03	0.92 ± 0.03	0.91 ± 0.03
Incl. rwgt. to $D\ell\nu$ sig ℓ	0.93 ± 0.04	0.92 ± 0.04	0.92 ± 0.04
Tension	-0.53	-0.08	-0.21
$D^*\ell\nu$ calibration fit	0.85 ± 0.03	0.89 ± 0.02	0.88 ± 0.02
Incl. rwgt. to $D^*\ell\nu$ sig ℓ	0.92 ± 0.04	0.90 ± 0.04	0.91 ± 0.04
Tension	-1.25	-0.40	-0.69

Table C.1: Comparing averaged calibration factors from the inclusive dataset with the exclusive dataset. See text for explanation.

3. The values from Item 1 and Item 2 should agree within uncertainties. Therefore, the third line shows the *tension* $(a - b)/(\sigma_a \oplus \sigma_b)$.[1]

It is crucial to point out that B^0 and B^+ are defined via the tag side in this closure study (see Chapter 11).

Overall, the calibration factors derived from the inclusive dataset seem to perform well on the exclusive datasets: the tension rarely exceeds 1σ.

Generally, the calibration factors seem to be slightly too high when compared to the results from the exclusive datasets, resulting in negative pulls. Correspondingly, the tension shrinks further when only considering the resonant gap model, rather than the average between both gap models. This is shown in Table C.2.

Comparison in bins of $\mathcal{P}_{\text{FEI}}$

The same comparison strategy can also be applied in bins of $\mathcal{P}_{\text{FEI}}$. The results are shown in Figure 11.1.

Technical note

The exclusive dataset currently contains the w_{NSL} weight (see Equation (7.3.2)). Assuming that the tag side only includes hadronic B decays and the signal side only includes semileptonic decays, this weight has been removed from the exclusive MC before performing the fits. This also applies to the results of Appendix C.1.2.

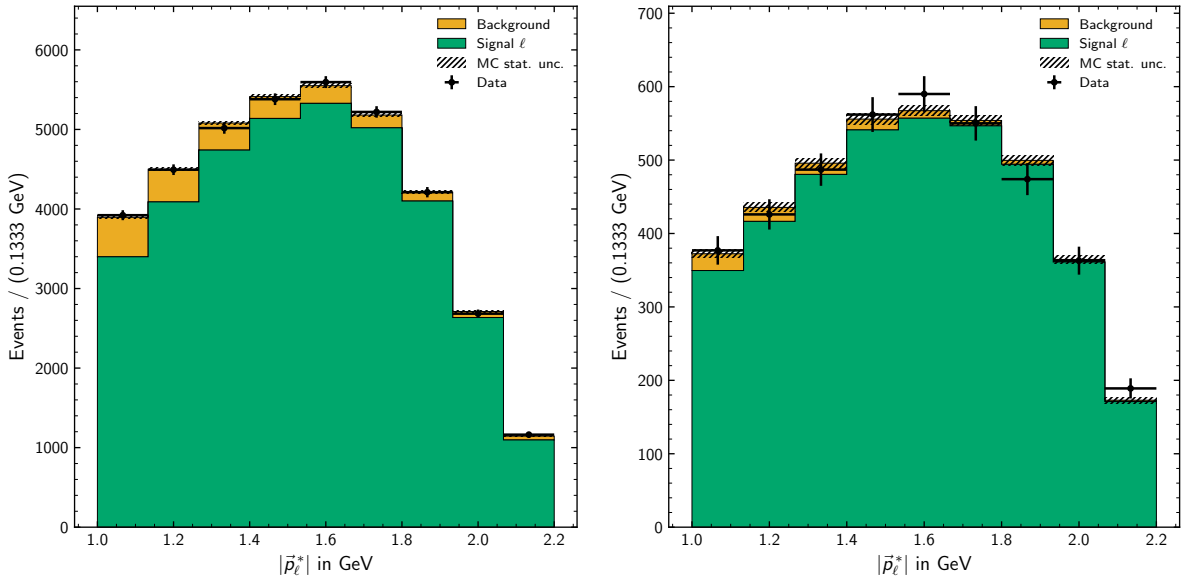
C.1.2 Comparing differential calibration factors

In this section, we perform direct comparisons of calibration factors extracted on the inclusive dataset with calibration factors extracted on the exclusive dataset. In order to reduce statistical uncertainties on the results on the exclusive dataset, both $B \to D\ell\nu_\ell$ and $B \to D^*\ell\nu_\ell$ are combined. Furthermore, the $D^* \to D\gamma$ decay is included in the $B \to D^*\ell\nu_\ell$ dataset. We use the categorization strategy of Section 9.1 with the parameters

$$N_{\text{sdc}} = 5 \times 10^3, \quad N_1 = 5 \times 10^3, \quad c_1 = 1 \times 10^3, \quad c_2 = 2 \times 10^3 \tag{C.1.3}$$

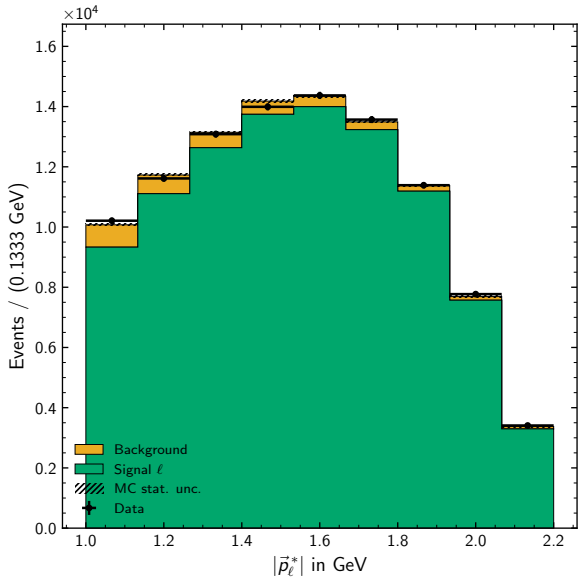
and apply it to the exclusive dataset. The same categories are then used in the subsequent fits to the inclusive dataset. The results are shown in Figure C.2. The scatter plot (Figure C.2b)

[1] The $B \to D^{(*)}\ell\nu_\ell$ branching ratio uncertainty is also included in the calibration uncertainty. However, there it only amounts to 0.6 % ($B \to D\ell\nu_\ell$) resp. 1.4 % ($B \to D^*\ell\nu_\ell$). We currently neglect the small amount of correlation that this induces.

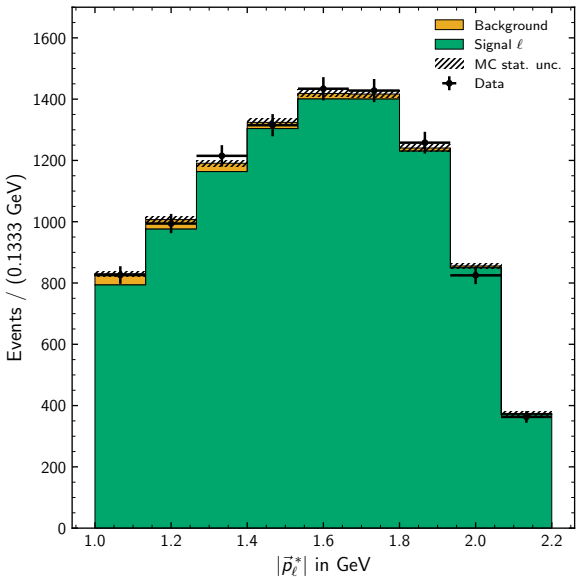


(a) $B^0 \to D^- \ell^+ \nu_\ell$

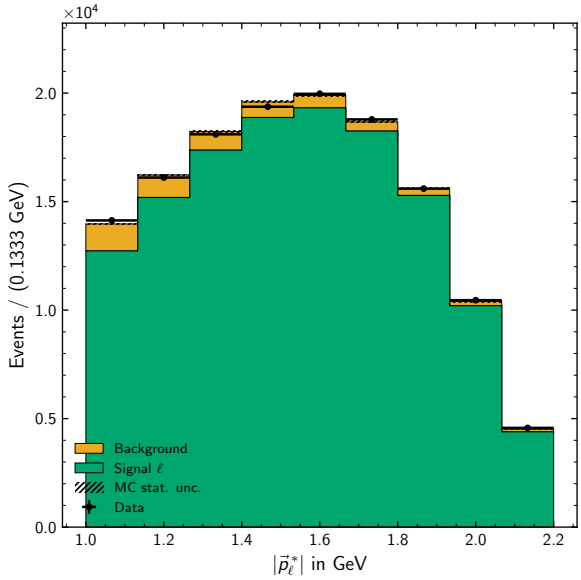
(b) $B^0 \to D^{*-} \ell^+ \nu_\ell$



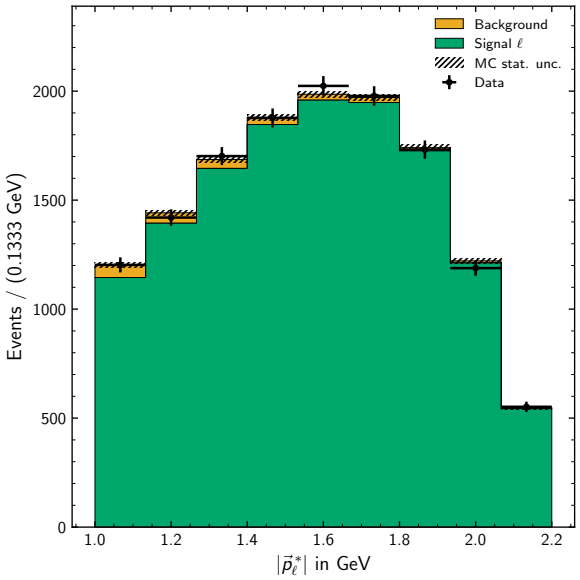
(c) $B^+ \to \overline{D}^0 \ell^+ \nu_\ell$



(d) $B^+ \to \overline{D}^{*0} \ell^+ \nu_\ell$



(e) $B \to D \ell \nu_\ell$



(f) $B \to D^* \ell \nu_\ell$

Figure C.1: Post-fit distribution of fits to exclusive data

	B^0	B^+	Both
$D\ell\nu$ calibration fit	0.90 ± 0.03	0.92 ± 0.03	0.91 ± 0.03
Incl. rwgt. to $D\ell\nu$ sig. ℓ	0.92 ± 0.04	0.91 ± 0.04	0.91 ± 0.04
Tension	-0.28	0.20	0.06
$D^*\ell\nu$ calibration fit	0.85 ± 0.03	0.89 ± 0.02	0.88 ± 0.01
Incl. rwgt. to $D^*\ell\nu$ sig. ℓ	0.90 ± 0.04	0.89 ± 0.04	0.89 ± 0.04
Tension	-1.03	-0.12	-0.42

Table C.2: Table equivalent to Table C.1, but using the calibration factors of the resonant gap model.

shows a systematic shift of 2 % between both sets of results, which is well below the correlated uncertainties (exceeding 4 % for the inclusive calibration alone). Comparing the results per category (Figure C.2a), we generally observe adequate agreement. However, some categories like $D^{*0}\pi^+\pi^+\pi^-\pi^0$ show systematic differences, especially for low $\mathcal{P}_{\text{FEI}}$. A possible explanation could be that the granularity in the sub decay channels is too coarse to cover the calibration factor and decay channel composition differences between both datasets. Furthermore, the recently discovered difference between calibration factors for correctly and incorrectly reconstructed tag flavors (Chapter 11) could affect these channels more than others.

C.2 Lepton flavor specific calibration factors

The calibration factors should not depend on whether we reconstruct an e or μ on the signal side. All differences should be covered by PID uncertainties.

To check this, we split up data and MC by the flavor of the lepton. We only consider the resonant gap model and only use fallback fits. The only systematic uncertainties that are considered are the lepton efficiencies (because all other systematic uncertainties would be assumed to be 100 % correlated between both lepton flavors).

The individual calibration factors are shown in Figure C.3a, the pull distribution between both sets of calibration factors in Figure C.3b. The mean of the pull distribution is 0.2 ± 0.5 (the considerable uncertainty is the consequence of the lepton efficiency being correlated between categories). Consequently, no overall bias is apparent.

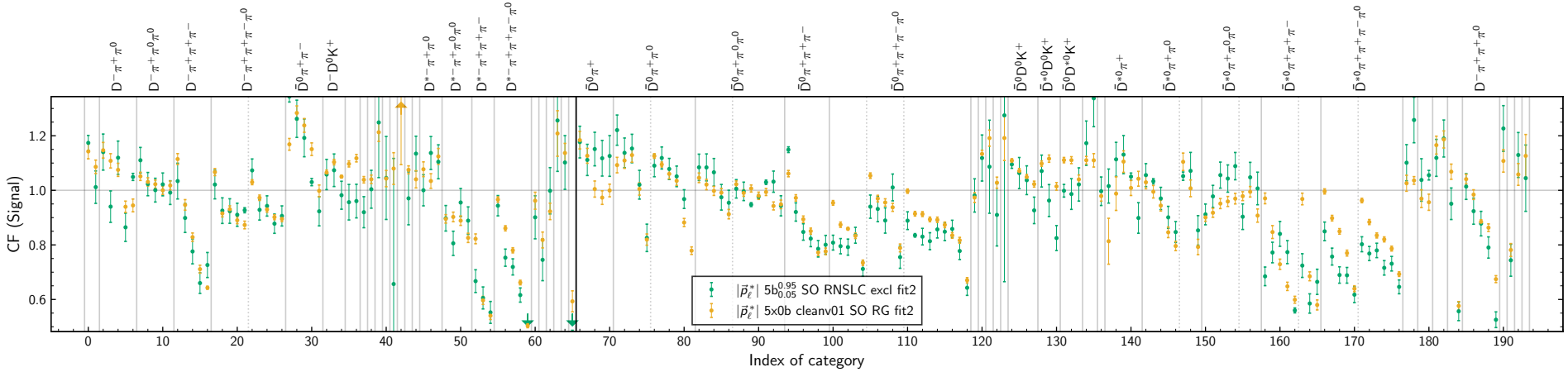
While the p-value is only 3 %, this seems to be mostly due to four outliers which are highlighted by blue bands in Figure C.3a. Excluding these brings the p-value to 34 % (Figure C.3c), which is acceptable.

C.3 Category migration

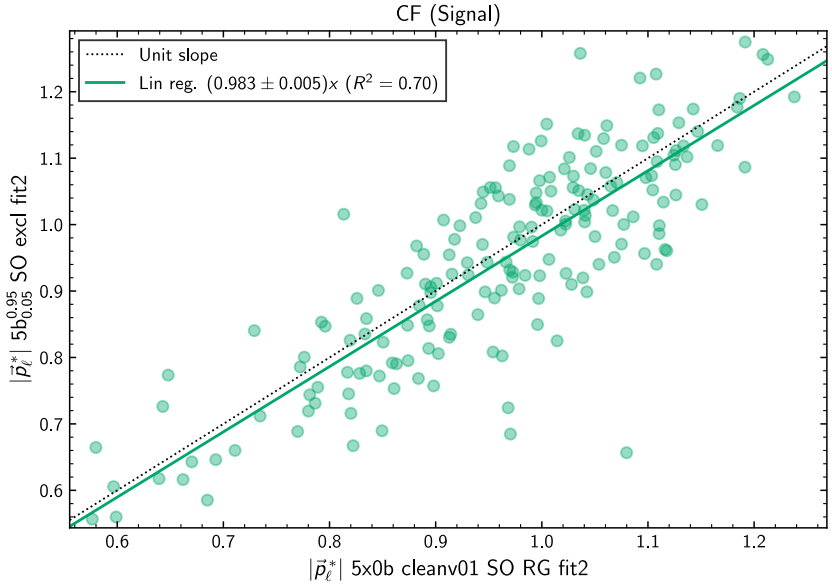
In this side-study, we investigate the impact of the signal side reconstruction on the B_{tag} category (FEI B_{tag} reconstruction mode and signal probability).

As described in more detail in Section 8.1, the reconstruction of an event E is equivalent[2] to the following steps

[2]Technically, B_{sig} and B_{tag} are reconstructed independently and then combined (which discards all candidate pairs that have "overlap").

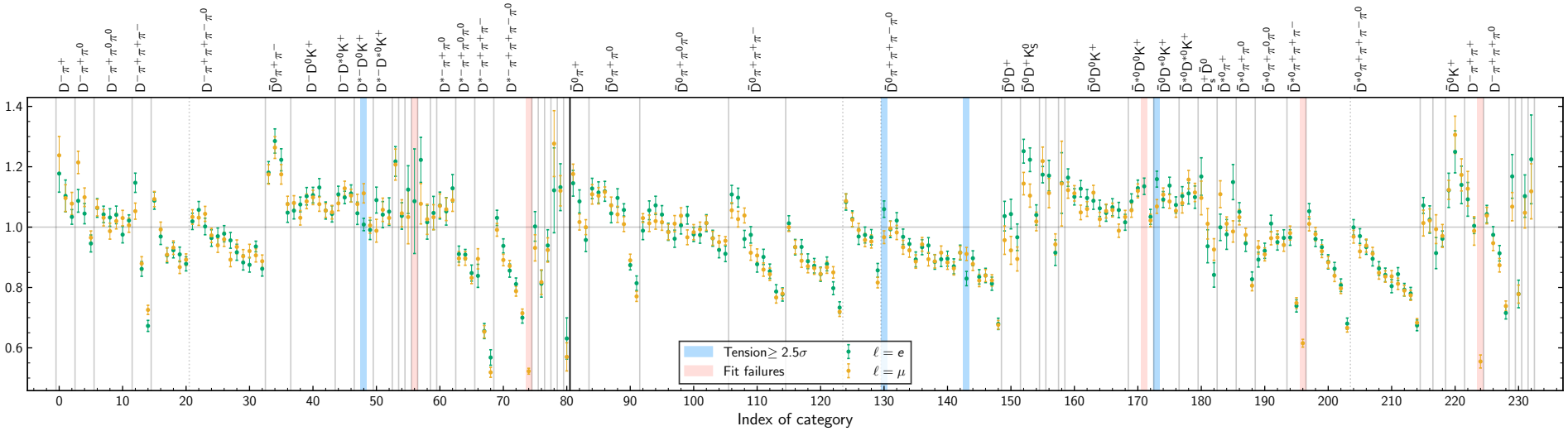


(a) Results for each category. The green (yellow) error bars show the results for the exclusive (inclusive) dataset.

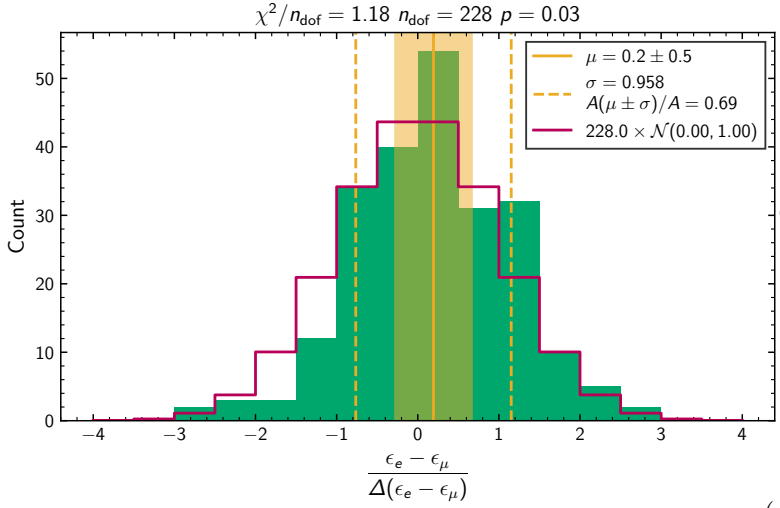


(b) Scatter plot comparing the inclusive calibration (x-axis) with those the exclusive calibration (y-axis).

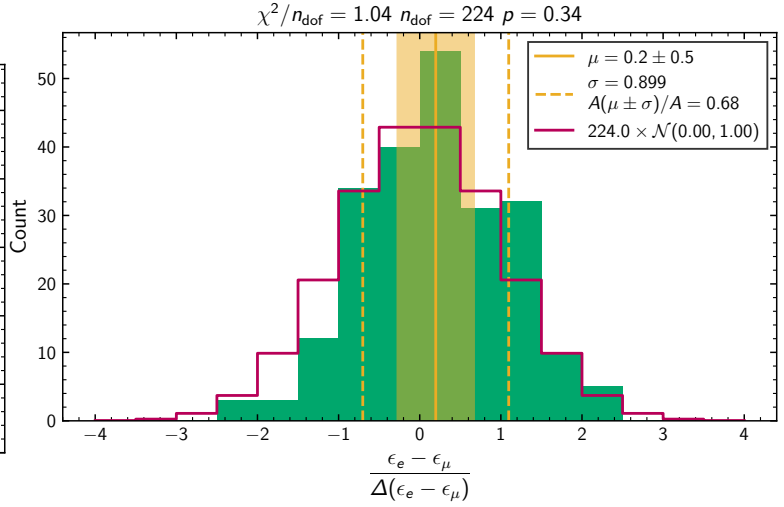
Figure C.2: Comparing calibration factors measured on the exclusive dataset with those measured on the inclusive dataset (resonant gap model). Only statistical uncertainties are shown.



(a) Calibration factors for each category. Categories with more than 2.5σ of tension between the electron and the muon calibration factor are highlighted in blue. Categories where one of the two fails to be fitted are highlighted in salmon.



(b) Pull distribution of both results



(c) Pull distribution of both results without calibration factors highlighted in blue

Figure C.3: Comparing calibration factors derived with only electrons or only muons

1. Signal side candidates are reconstructed. Only a single lepton is reconstructed in the inclusive data set; in the exclusive data set, a $B \to D^* \ell \nu_\ell$ decay is reconstructed.

2. Tag side candidates are reconstructed from the ROE of step 1.

3. A best candidate selection is applied.

If we ignore any cuts that might discard the best candidate, this maps the event to a single candidate with an FEI reconstruction mode C_{FEI} and classifier output $\mathcal{P}_{\text{FEI}}$. Because of the different reconstruction strategies in the tagging calibration and the $|V_{cb}|$ analysis, the same event could be reconstructed with $(C^{\text{incl}}_{\text{FEI}}, \mathcal{P}^{\text{incl}}_{\text{FEI}}) \neq (C^{\text{excl}}_{\text{FEI}}, \mathcal{P}^{\text{excl}}_{\text{FEI}})$. Therefore, the same event could be placed into two different categories $C^{\text{incl}} \neq C^{\text{excl}}$ (depending on the categorization). In the following, we will refer to this as *migration* (an event migrates from one tag category to another depending on the reconstruction of the signal side). We will write $f^{\text{incl}}(E) = C^{\text{incl}}$ ($f^{\text{excl}}(E) = C^{\text{excl}}$) with f^{incl} (f^{excl}) the function that maps the event E to its category based on the inclusive (exclusive) reconstruction.

Let us recall that the calibration procedure should ultimately estimate the quantity

$$\epsilon_i^{\text{excl}} := \frac{N^{\text{Data}}_{\text{excl},i}}{N^{\text{MC}}_{\text{excl},i}} := \frac{\#\{E \in \Omega^{\text{Data}}_{\text{excl}} \mid f^{\text{excl}}(E) = C_i\}}{\#\{E \in \Omega^{\text{MC}}_{\text{excl}} \mid f^{\text{excl}}(E) = C_i\}}, \tag{C.3.1}$$

where $\Omega^{\text{Data/MC}}_{\text{excl}}$ are the exclusive ($|V_{cb}|$) data sets and we ignored possible efficiency differences between data and MC on the signal side.

The idea of this tagging calibration procedure is to measure the same fraction on an (almost) independent data set $\Omega^{\text{Data/MC}}_{\text{incl}}$, that is[3]

$$\frac{\#\{E \in \Omega^{\text{Data}}_{\text{incl}} \mid f^{\text{excl}}(E) = C_i\}}{\#\{E \in \Omega^{\text{MC}}_{\text{incl}} \mid f^{\text{excl}}(E) = C_i\}}. \tag{C.3.2}$$

However, what we actually measured so far, is

$$\epsilon_i^{\text{incl}} := \frac{N^{\text{Data}}_{\text{incl},i}}{N^{\text{MC}}_{\text{incl},i}} := \frac{\#\{E \in \Omega^{\text{Data}}_{\text{incl}} \mid f^{\text{incl}}(E) = C_i\}}{\#\{E \in \Omega^{\text{MC}}_{\text{incl}} \mid f^{\text{incl}}(E) = C_i\}}. \tag{C.3.3}$$

If the signal side reconstruction affects the reconstruction of the B_{tag} and f^{incl} is significantly different from f^{excl}, then the quantity calculated in Equation (C.3.2) (with which we want to approximate ϵ_i^{excl}) is not guaranteed to match ϵ_i^{incl} from Equation (C.3.3). On the other hand, the two quantities are also not guaranteed to differ, as two effects play in our favor:

i. We only *count* the events in a certain category, i.e., for symmetric absolute[4] migration (as much migration *to* a category as *from* a category), the migration cancels and numerator and denominator stay the same.

ii. We are only interested in the ratio of data over MC, i.e., for similar migration in data and MC, the effects will tend to cancel.

In fact, this is what we will observe: There is substantial migration, but the effects seem to cancel to the extent that no further corrections are necessary.

[3] In practice, we perform fits to calculate the fraction of Equations (C.3.2) and (C.3.3) only with the signal component, but this subtlety is ignored in this section.

[4] I.e., for migration measured in the absolute number of events. See Appendix D.6.2 for a mathematical description in terms of migration matrices.

	Events	F. excl.	F. incl.
Data	193903	91.36%	1.74%
MC	725528	86.78%	1.50%

Table C.3: Overlap sample. *F. excl.* (*F. incl.*) denotes the fraction of events of the exclusive (inclusive) data sample (after several cuts, see text) that are in the overlap sample.

C.3.1 Methodology

We investigate the *event overlap* between the inclusive and exclusive samples, i.e., every event that is reconstructed and passes the respective cuts in both the inclusive reconstruction and the exclusive reconstruction. Here, the exclusive dataset includes both $B \to D\ell\nu_\ell$ and $B \to D^*\ell\nu_\ell$ decays.

The sample event counts are summarized in Table C.3. It should be noted that we have applied an additional $|\vec{p}_\ell^{\,*}| > 0.75\,\text{GeV}$ cut to match the cuts of the inclusive dataset and selected only the first stream of the exclusive MC (because the inclusive MC only includes this stream). Similarly, the rare MC has been excluded from the numbers in the overlap (because it is not contained in the inclusive data set), and the non-resonant gap model is used in both datasets.

As expected, the overwhelming fraction of exclusive events can be found in the inclusive dataset because the inclusive reconstruction is much looser than the exclusive one. A more detailed version of Table C.3 is given in the appendix (Table D.3).

We define

$$N_{Y\,i}^{\text{ovl}\,X} := \#\{E \in \Omega_{\text{ovl}}^X \mid f^Y(E) = C_i\}, \tag{C.3.4}$$

$$\epsilon_{Y\,i}^{\text{ovl}} := \frac{N_{Y\,i}^{\text{ovl Data}}}{N_{Y\,i}^{\text{ovl MC}}} \tag{C.3.5}$$

where X stands for Data or MC and Y for inclusive or exclusive. Note that we do not need to calculate an explicit migration matrix here. However, it can provide some additional insight into the migration effects. For this reason, migration matrices are introduced in Appendix D.6.

To evaluate the statistical compatibility of $\epsilon_{\text{incl}\,i}^{\text{ovl}}$ and $\epsilon_{\text{excl}\,i}^{\text{ovl}}$, we will consider the pull

$$g_i := \frac{\epsilon_{\text{incl}\,i}^{\text{ovl}} - \epsilon_{\text{excl}\,i}^{\text{ovl}}}{\sqrt{(\Delta\epsilon_{\text{incl}\,i}^{\text{ovl}})^2 + (\Delta\epsilon_{\text{excl}\,i}^{\text{ovl}})^2}}. \tag{C.3.6}$$

We perform a χ^2 test with N_{cat} degrees of freedom with $\chi^2 = \sum_i g_i^2$.

We should note that $\epsilon_{\text{incl}\,i}^{\text{ovl}}$ and $\epsilon_{\text{excl}\,i}^{\text{ovl}}$ are not strictly independent from each other because $N_{\text{excl}\,i}^{\text{ovl}\,X}$ and $N_{\text{incl}\,i}^{\text{ovl}\,X}$ are histograms of two variables on the *same* dataset. To avoid further complicating this side-study, we randomly split the overlap dataset (data and MC) into two parts. One of these parts is used to calculate $N_{\text{excl}\,i}^{\text{ovl}\,X}$ and the other to calculate $N_{\text{incl}\,i}^{\text{ovl}\,X}$. In this way, both variables are independent, and no correlations need to be taken into account to calculate the uncertainties on g_i.

Besides the p-value of the χ^2 test, we consider a second quantity, which we call *bias*: $b = \mu/\Delta\mu$, where μ is the mean of the difference $\epsilon_{\text{incl}\,i}^{\text{ovl}} - \epsilon_{\text{excl}\,i}^{\text{ovl}}$.

While the random split simplifies our calculations, it has the disadvantage that it introduces another source of randomness into the results: the p-value and the bias will vary to some extent based on the random split. For this reason, we repeat the same experiment for different splits and show distributions of the results.

Quantities that are not sensitive to the treatment of these correlations (like simple plots of histograms) are calculated without the random split.

In the following sections, we investigate

1. The migration in the FEI classifier $\mathcal{P}_{\text{FEI}}$ (Appendix C.3.2),
2. The migration in the FEI reconstruction channels C_{FEI} (Appendix C.3.3),
3. The migration in the categories C_i used in Chapter 10 (Appendix C.3.4).

The reason that we first look at migration in $\mathcal{P}_{\text{FEI}}$ and C_{FEI} separately is that the results are easier to visualize and understand qualitatively (in particular, the complete migration matrices given in Appendix D.6). They also have slightly more statistical power (because fewer bins are being used).

In each of these cases, we will show that $\epsilon^{\text{ovl}}_{\text{excl}\,i}$ and $\epsilon^{\text{ovl}}_{\text{incl}\,i}$ are statistically compatible. This is summarized in Appendix C.3.5.

C.3.2 Migration in the FEI classifier

The categories are defined as 30 quantiles in $\mathcal{P}^{\text{incl}}_{\text{FEI}}$ in data. The resulting histograms $N^{\text{ovl}\,X}_{Y\,i}$ are shown in Figures C.4a and C.4b. $N^{\text{ovl}\,\text{Data}}_{\text{incl}\,i}$ is flat by the definition of the categories. From Figures C.4a and C.4b, we can immediately recognize that migration occurs in both data and MC (else the inclusive and exclusive distributions should match – do note the cropped ordinate though). However, the effects seem to approximately cancel between data and MC so that the calibration factors for each category do not seem to systematically differ between inclusive and exclusive categorization (Figures C.4c to C.4e).

To test this quantitatively, we now calculate g_i using the randomly split dataset and calculate the p-value from the χ^2 test and the bias. Repeating this procedure leads to the distributions shown in Figure C.5. Note that we do not necessarily expect the p-value distribution to be flat and the bias distribution to be centered around zero if our hypothesis is correct: we only have one dataset that we investigate. By repeating the random splitting, we only check the influence of this source of randomness on the results. Nonetheless, the p-value distribution *is* flat in this case, with its mean close to 0.5, indicating a very good agreement of $\epsilon^{\text{ovl}}_{\text{incl}\,i}$ and $\epsilon^{\text{ovl}}_{\text{excl}\,i}$. The mean centers around -0.9, so there is less one standard deviation of tension between the means of $\epsilon^{\text{ovl}}_{\text{incl}\,i}$ and $\epsilon^{\text{ovl}}_{\text{excl}\,i}$, which is acceptable ($\sim 30\,\%$ probability under the null hypothesis). We conclude that we do not see evidence for the $\mathcal{P}_{\text{FEI}}$ categorization migration having a systematic effect on the calibration factors.

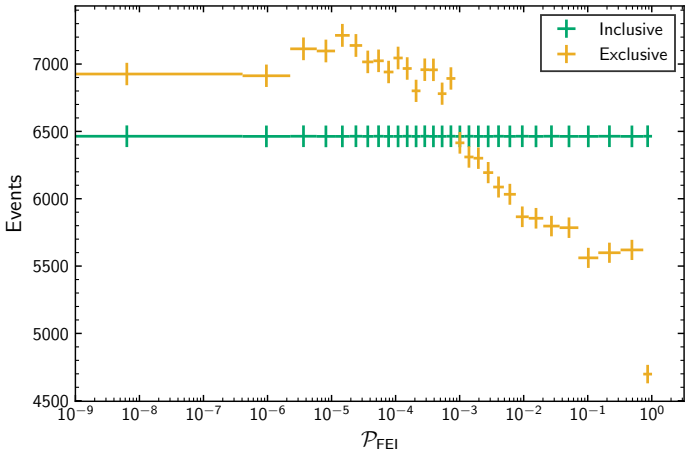
More details about the migration are presented in Appendix D.6.3: In general, the migration is very large, especially on MC, but tends to cancel by the effects of Items i and ii.

C.3.3 Migration in the FEI reconstruction channel

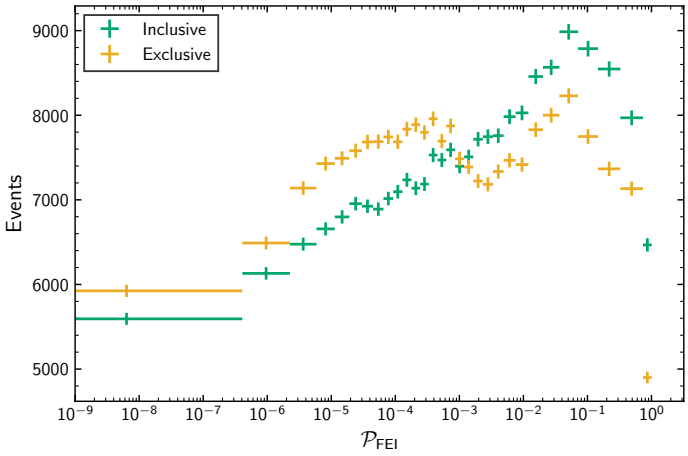
We proceed precisely as in the last section, this time choosing 54 categories for the 54 reconstruction modes C_{FEI}. The histograms of $N^{\text{ovl}\,\text{Data}}_{Y\,i}$ and $N^{\text{ovl}\,\text{MC}}_{Y\,i}$ are shown in Figures C.6a and C.6b. Some channels show clear significant migration, but again, this seems to cancel out in the calibration factors (Figures C.6c to C.6e).

Performing the statistical tests on the randomly split datasets leads to the distributions shown in Figure C.7. The average p-value is $74\,\%$ and the mean of $\epsilon^{\text{ovl}}_{\text{incl}\,i} - \epsilon^{\text{ovl}}_{\text{excl}\,i}$ shows a deviation of less than 0.5σ from zero on average. We conclude that we do not see evidence that the C_{FEI} category migration systematically affects the calibration factors.

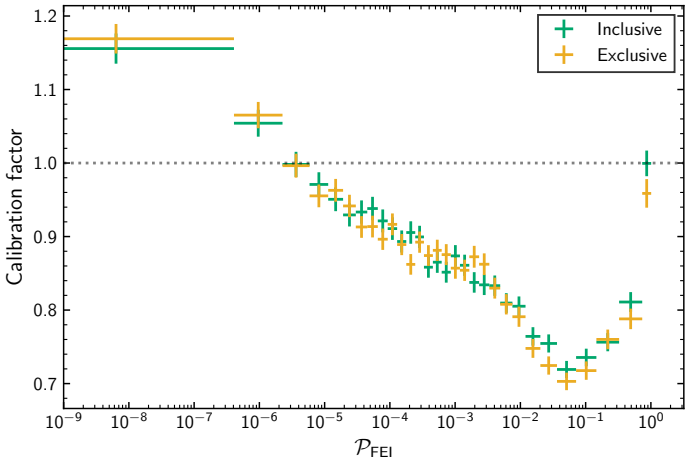
More details about the underlying migration are presented in Appendix D.6.4.



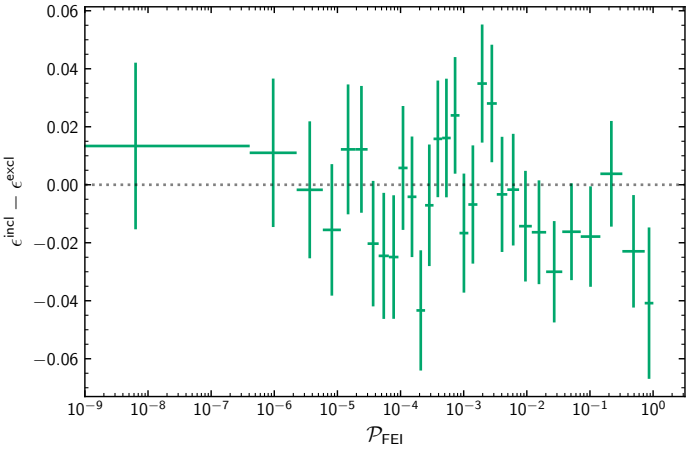
(a) $N_{Y\,i}^{\text{ovl Data}}$



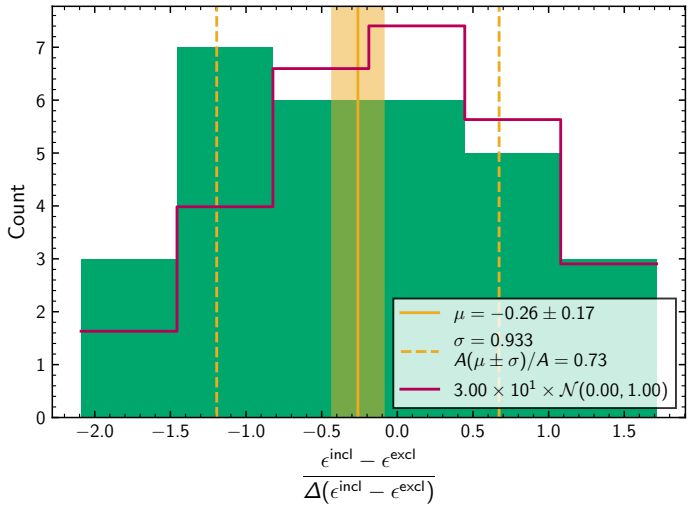
(b) $N_{Y\,i}^{\text{ovl MC}}$



(c) $\epsilon_{\text{incl}\,i}^{\text{ovl}}$ vs. $\epsilon_{\text{excl}\,i}^{\text{ovl}}$

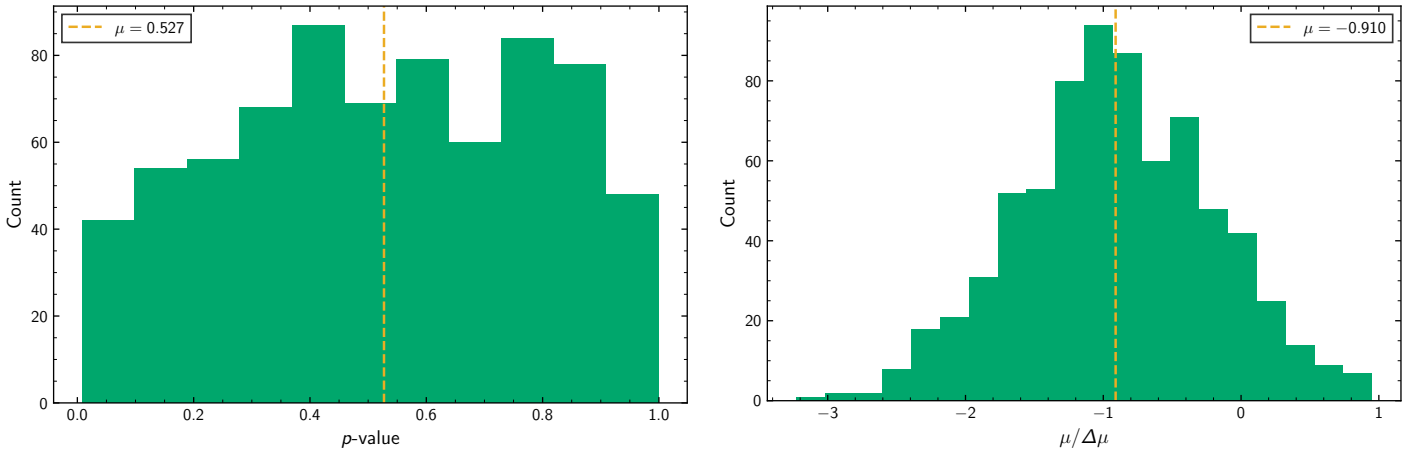


(d) $\epsilon_{\text{incl}\,i}^{\text{ovl}} - \epsilon_{\text{excl}\,i}^{\text{ovl}}$



(e) Pull distribution

Figure C.4: Investigating the categorization in $\mathcal{P}_{\text{FEI}}$ on the overlap dataset. No correlations considered, no random split performed.



(a) p-values (b) Bias

Figure C.5: Testing the statistical compatibility of $\epsilon^{\text{ovl}}_{\text{incl}\,i}$ and $\epsilon^{\text{ovl}}_{\text{excl}\,i}$ for categorization in $\mathcal{P}_{\text{FEI}}$.

Migration in	$\langle p \rangle$	$\langle \mu/\Delta\mu \rangle$
$\mathcal{P}_{\text{FEI}}$	53%	−0.91
C_{FEI}	74%	0.44
$\mathcal{P}_{\text{FEI}}$ and C_{FEI}	82%	0.42

Table C.4: Summary of the results of the category migration studies.

C.3.4 Migration in the FEI classifier and reconstruction channel

Finally, we use 249 categories in both the reconstruction channel and the FEI classifier and repeat the same procedure.

The histograms of $N^{\text{ovl Data}}_{Y\,i}$ and $N^{\text{ovl MC}}_{Y\,i}$ are shown in Figure C.10. Significant migration is visible, but again, seems to cancel in the calibration factors (Figures C.8 and C.10). This is further corroborated by the statistical tests Figure C.9: On average, a p-value of 82 % is observed and a bias of 0.42. We conclude that no statistically significant impact on the calibration factors by category migration can be observed.

More details about the underlying migration are presented in Appendix D.6.5.

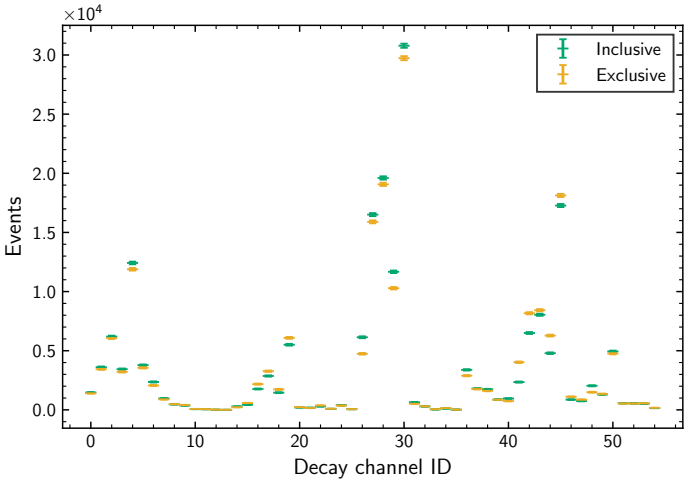
C.3.5 Summary

The findings of the three sub-studies are summarized in Table C.4. We conclude from the p-values that the observed effects are consistent with statistical fluctuations. Thus no correction of category migration effects is necessary in the tagging calibration.

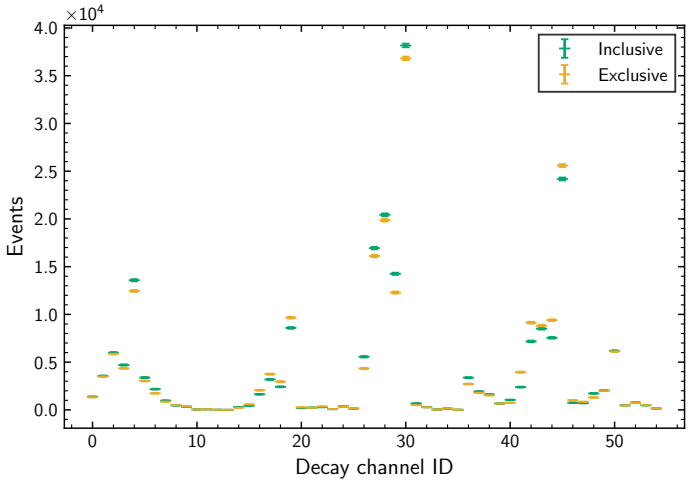
C.4 Toy studies

To validate the fit procedure described in Section 9.2, we perform toy experiments. For this, we generate bootstrapped datasets from MC.

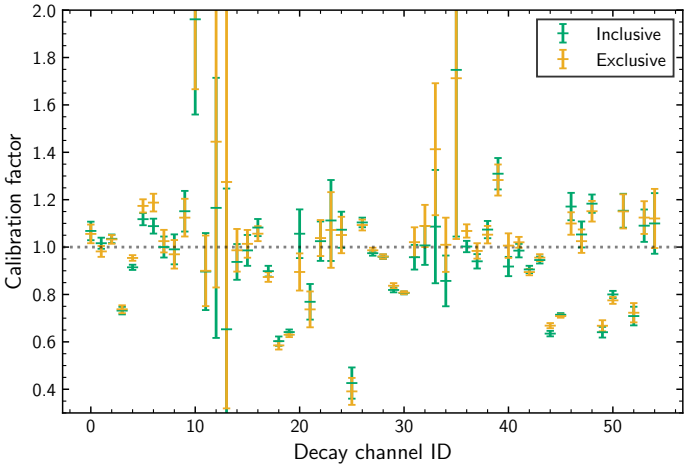
As a first step, we sample with replacement all MC events with probabilities given by their weight until we reach a dataset with equal total weight. This new dataset does not have any event weights anymore (the heavier events of the weighted dataset were more likely to be sampled multiple times instead).



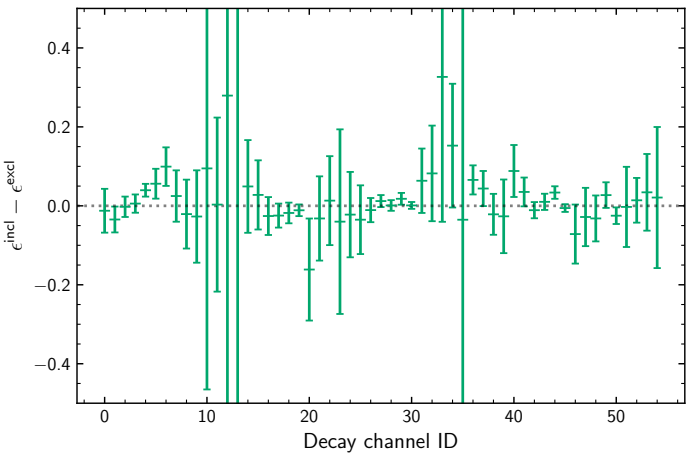
(a) $N_{Y\,i}^{\text{ovl Data}}$



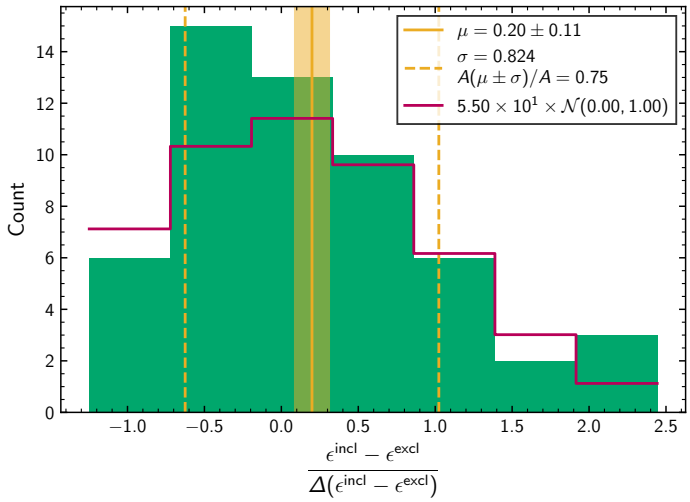
(b) $N_{Y\,i}^{\text{ovl MC}}$



(c) $\epsilon_{\text{incl}\,i}^{\text{ovl}}$ vs. $\epsilon_{\text{excl}\,i}^{\text{ovl}}$

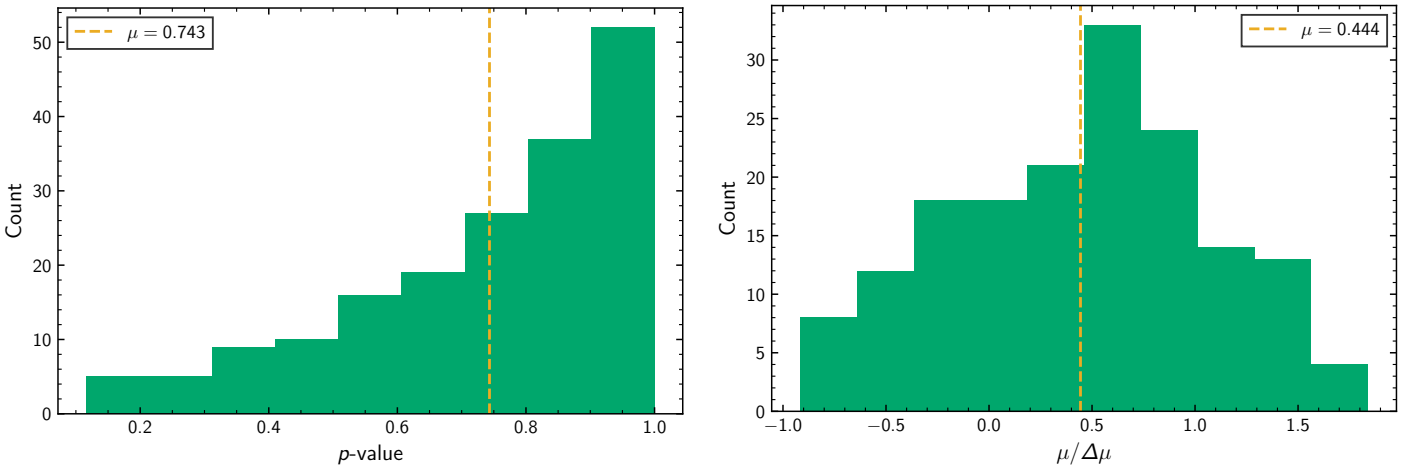


(d) $\epsilon_{\text{incl}\,i}^{\text{ovl}} - \epsilon_{\text{excl}\,i}^{\text{ovl}}$



(e) Pull distribution

Figure C.6: Investigating the categorization in $\mathcal{P}_{\text{FEI}}$ on the overlap dataset. No correlations considered, no random split performed.



(a) p-values (b) Bias

Figure C.7: Testing the statistical compatibility of $\epsilon^{\text{ovl}}_{\text{incl}\,i}$ and $\epsilon^{\text{ovl}}_{\text{excl}\,i}$ for categorization in C_{FEI}.

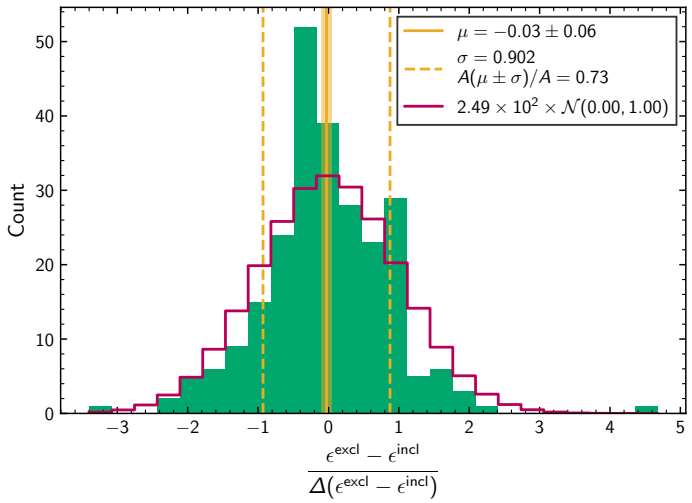
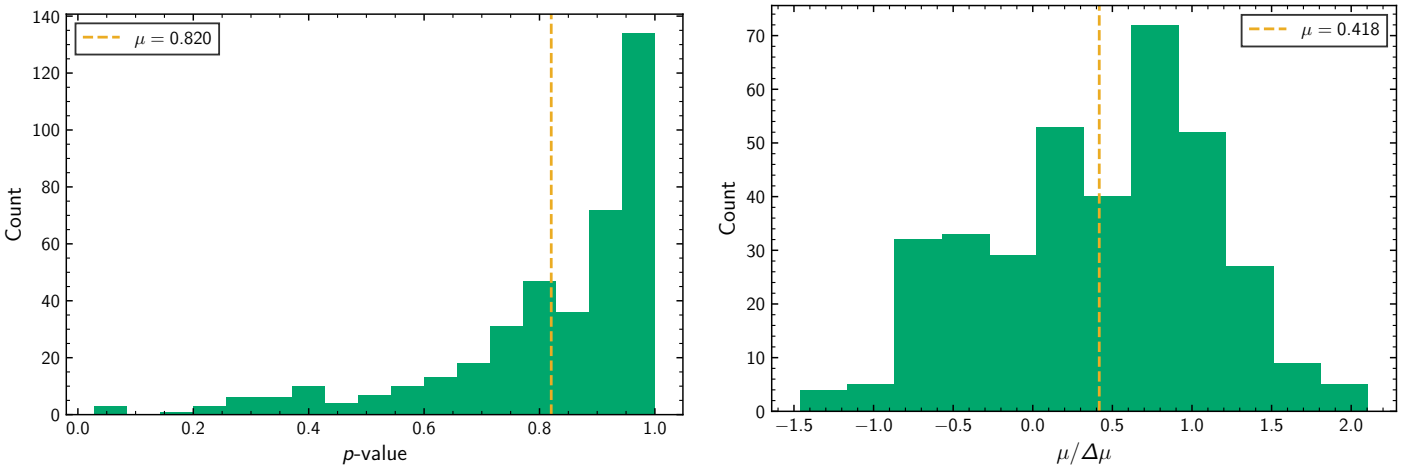


Figure C.8: Pull distribution of $\epsilon^{\text{ovl}}_{\text{incl}\,i} - \epsilon^{\text{ovl}}_{\text{excl}\,i}$ in $(C_{\text{FEI}}, \mathcal{P}_{\text{FEI}})$ categories.



(a) p-values (b) Bias

Figure C.9: Testing the statistical compatibility of $\epsilon^{\text{ovl}}_{\text{incl}\,i}$ and $\epsilon^{\text{ovl}}_{\text{excl}\,i}$ for categorization in $(C_{\text{FEI}}, \mathcal{P}_{\text{FEI}})$.

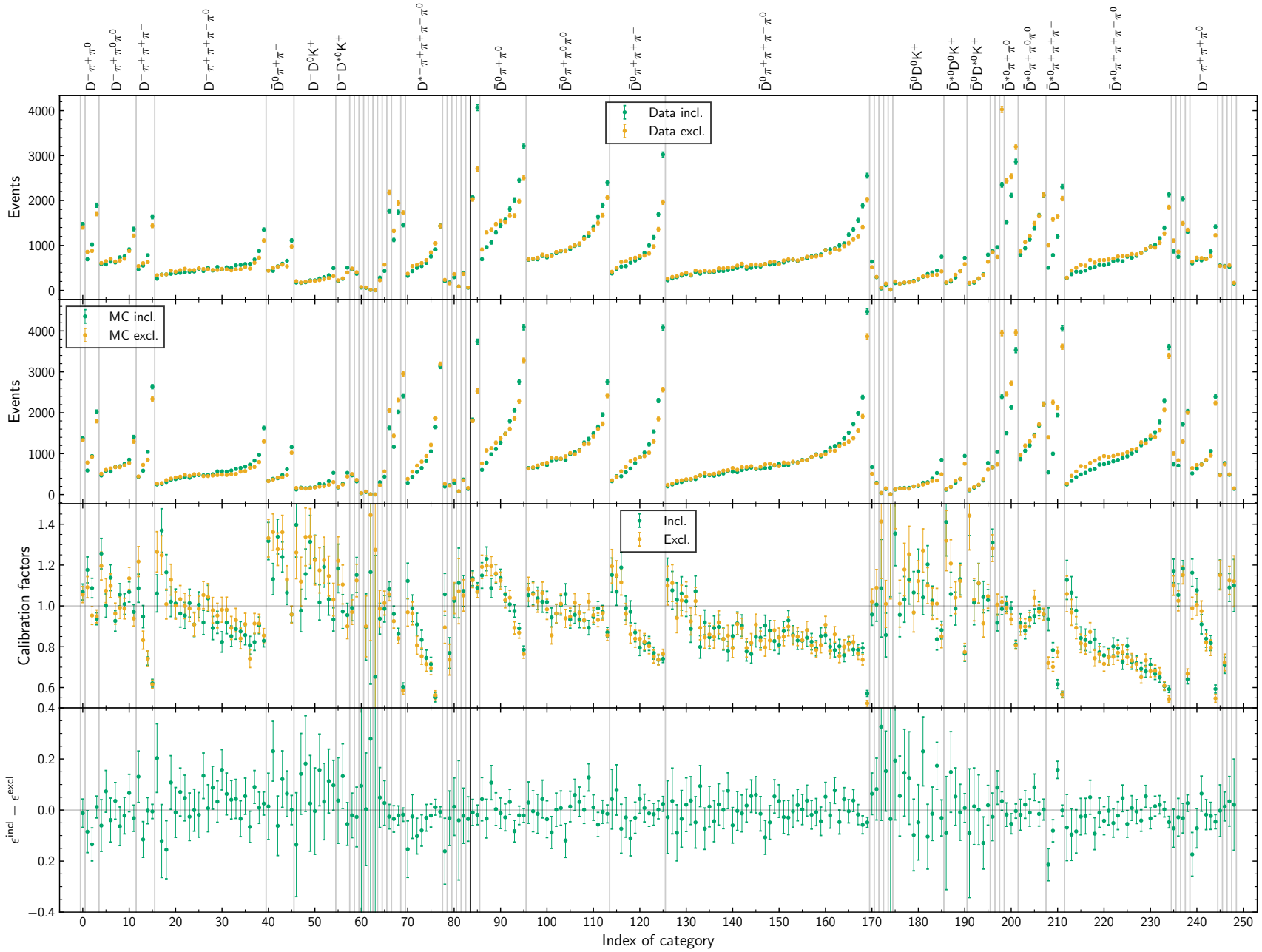


Figure C.10: Migration in $(C_{\text{FEI}}, \mathcal{P}_{\text{FEI}})$ categories in the overlap samples.

We now add Poisson weights to each event, that is weights $w \sim \text{Pois}(1)$, i.e., integer weights $k \in \mathbb{N}_0$ with a probability

$$\text{P}(w = k) = \frac{1}{\text{e} \cdot k!}, \tag{C.4.1}$$

where e is Euler's constant. This procedure creates a new dataset which we take as MC. Adding a different set of Poisson weights, we obtain yet another dataset, which we will assume to be data.

Using this newly created MC and data dataset, we fit and calculate calibration factors as outlined in Sections 9.2 and 9.3 but without any systematic uncertainties: All nuisance parameters are fixed during the fit.[5]

Because both datasets were generated from the same distribution, we expect to see the calibration factors ϵ_c^i to have median $\mu_2(\{\epsilon_{\text{sig}}^i \,|\, i = 0, \ldots, N_{\text{cat}} = 1\}) = 1$. This is clearly the case (left figures in Figure C.11). We further expect the pull distributions

$$\text{pull}_i = \frac{\epsilon_c^i - 1}{\Delta\epsilon_c^i} \tag{C.4.2}$$

to be Gaussian distributions $\mathcal{N}(\mu = 0, \sigma = 1)$. As we can see in Figure C.11, the pull distribution for signal calibration factors is a bit more narrow than expected, which means that we slightly overestimate the uncertainty $\Delta^{\text{stat}}\epsilon_c^i$. This can also be seen in Figure C.12 which compares the uncertainty from the standard error of the distribution of all ϵ_{sig}^i for fixed i with the mean of the statistical uncertainty as reported by the fitter: the uncertainty as estimated from the standard error of the bootstrapped dataset fluctuates around 1.0 %, while the uncertainty as reported from the fitter fluctuates around 1.2 %. The pull distributions for the two background components look perfect.

Because this makes the estimate of the fitter slightly more conservative and because the difference in uncertainties only has a negligible impact on the total uncertainty, this is not an issue.

Furthermore, we can check the calculation of the p-values of the fitter: As expected, Figure C.13 shows a completely flat distribution.

By assigning additional weights w_c to the different components in the bootstrapped data sample, we can also check the linearity of the fits. This is shown in Figure C.14. In all cases, even in the case of scaling fakes and secondaries separately (which is usually covered by an additional uncertainty), the deviation of the median of the calibration factors from the expectation is of $\mathcal{O}(10^{-4})$ or less.

C.5 Refitting calibrated distribution

As another closure check, we can apply the calibration factors ϵ_{sig}^i (as shown in Chapter 10) to the full inclusive dataset and perform another fit to the signal component. We then expect to see a unit signal calibration factor (we still need to do a fit for this because the background components have not been calibrated, so we cannot simply compare data and MC normalizations). The corresponding post-fit distributions are shown in Figure C.15.

Using only statistical uncertainties, we obtain

$$\epsilon_{\text{sig}}^{\text{B}^0} = 1.005 \pm 0.005, \qquad \epsilon_{\text{sig}}^{\text{B}^+} = 0.998 \pm 0.004, \tag{C.5.1}$$

thus we indeed see closure.

[5] Without fixing the nuisance parameters, the templates could still be adjusted because of the statistical uncertainty on the templates.

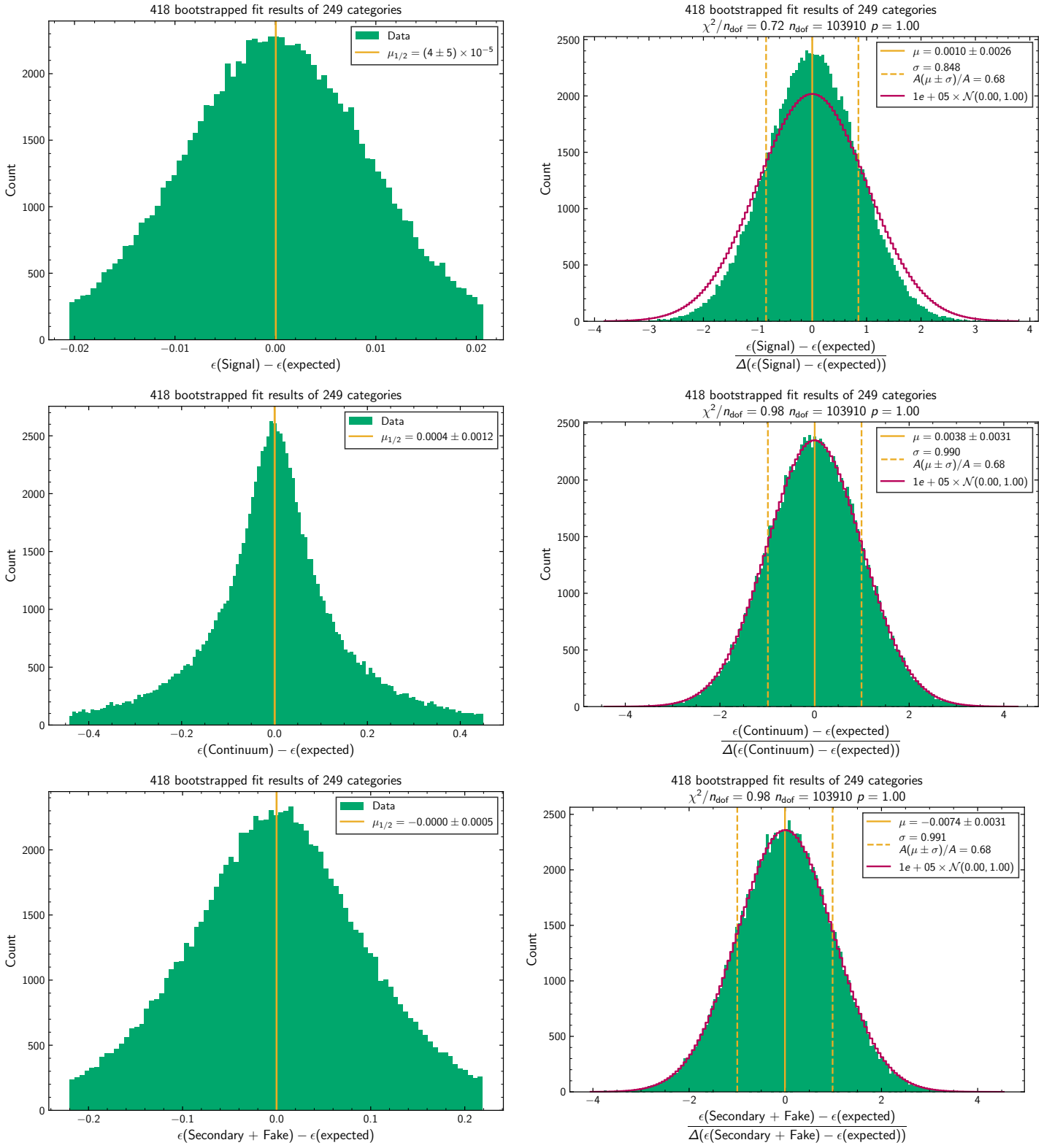


Figure C.11: Fitted calibration factors ϵ_i^c for the bootstrapped dataset. In all three cases, $\epsilon(\text{expected}) = 1$.

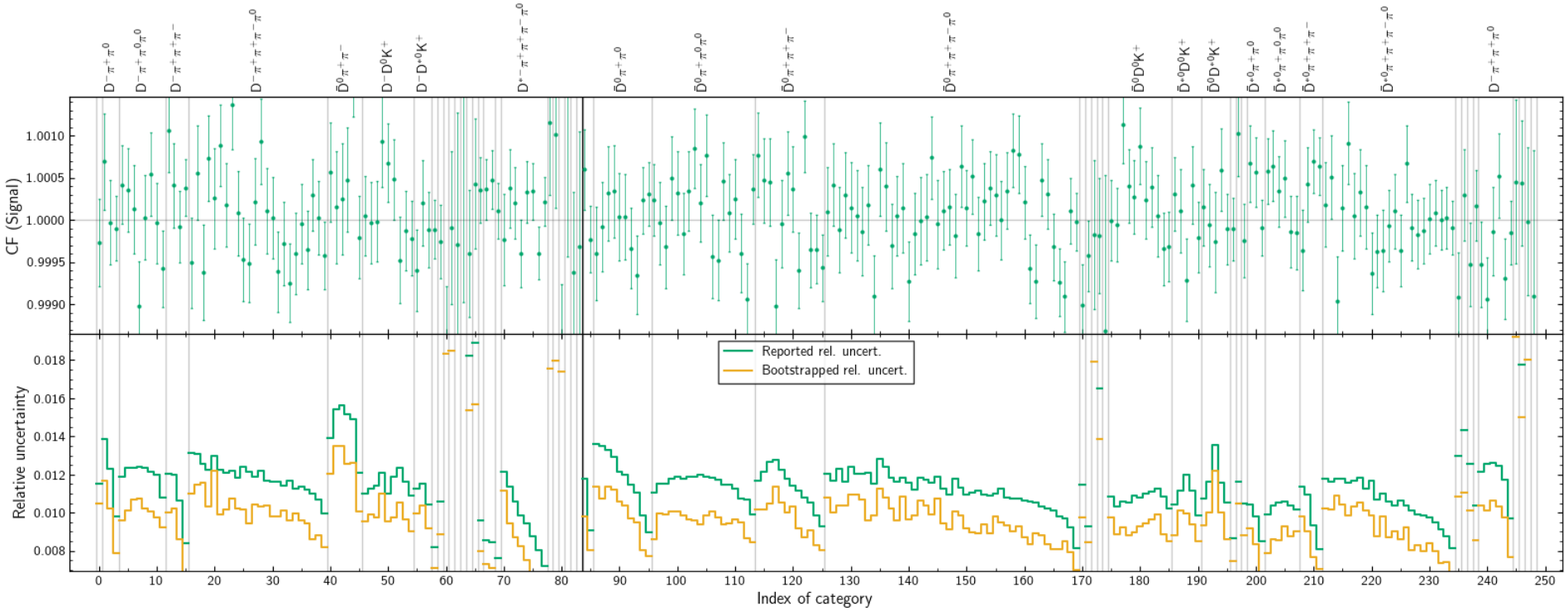
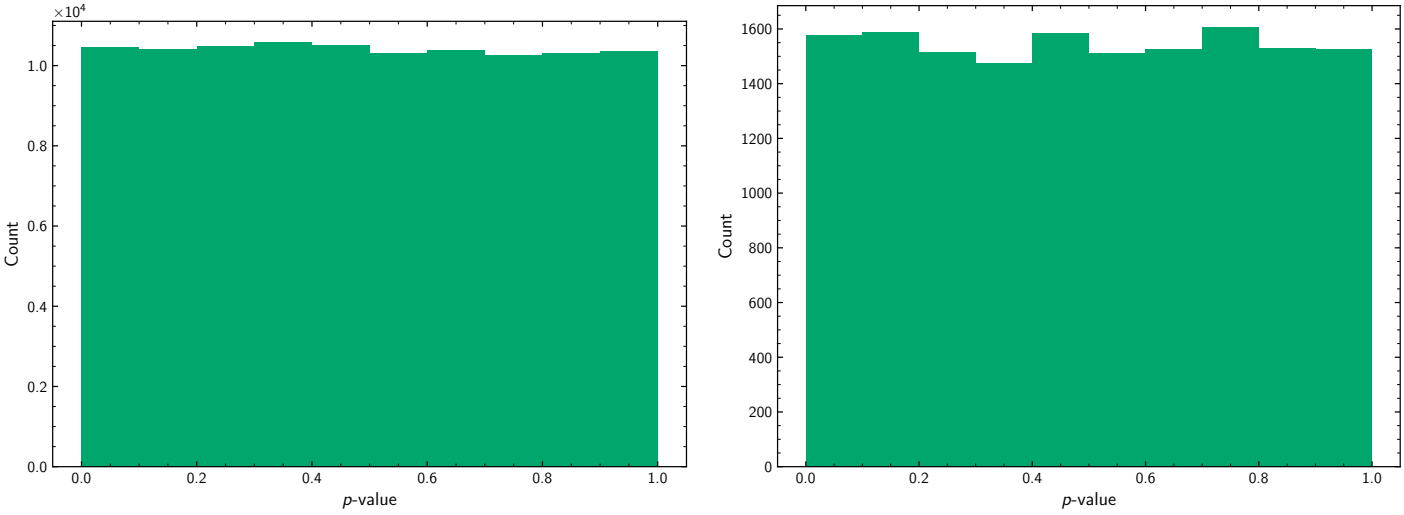


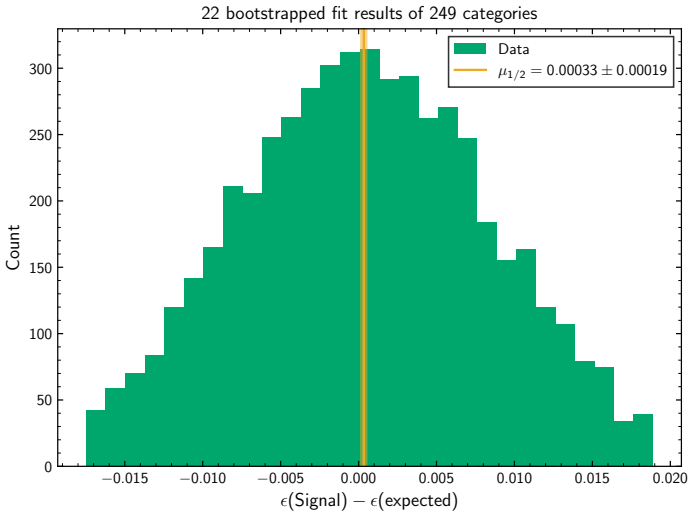
Figure C.12: Calibration results for bootstrapped data per category. The upper figure shows the mean and standard error of the distribution of $\epsilon_{\text{sig}}^{i}$ (fixed i) for the different bootstrappings. The bottom figure compares the standard error from the upper plot with the mean of the statistical uncertainty as reported by the fitter.



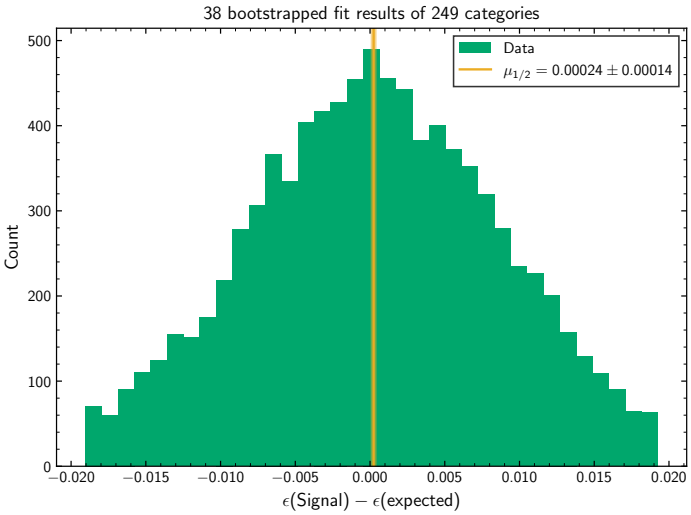
(a) Primary fit

(b) Fallback fit

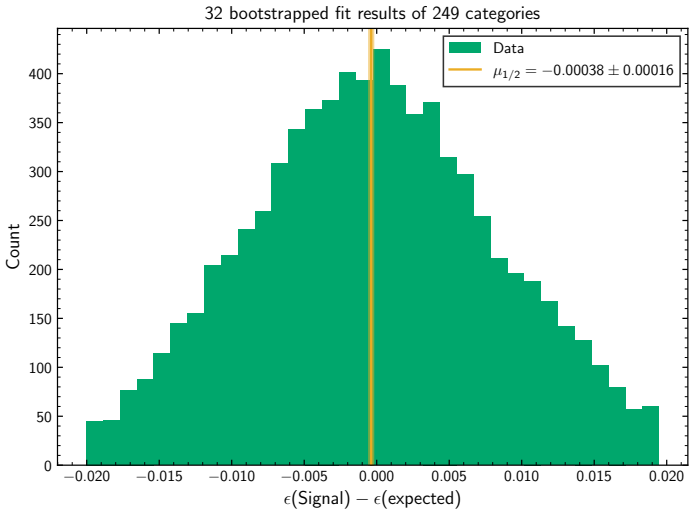
Figure C.13: Distribution of p-values from fits to bootstrapped data (combined for all categories).



(a) $w_{\text{sig}} = 0.9$

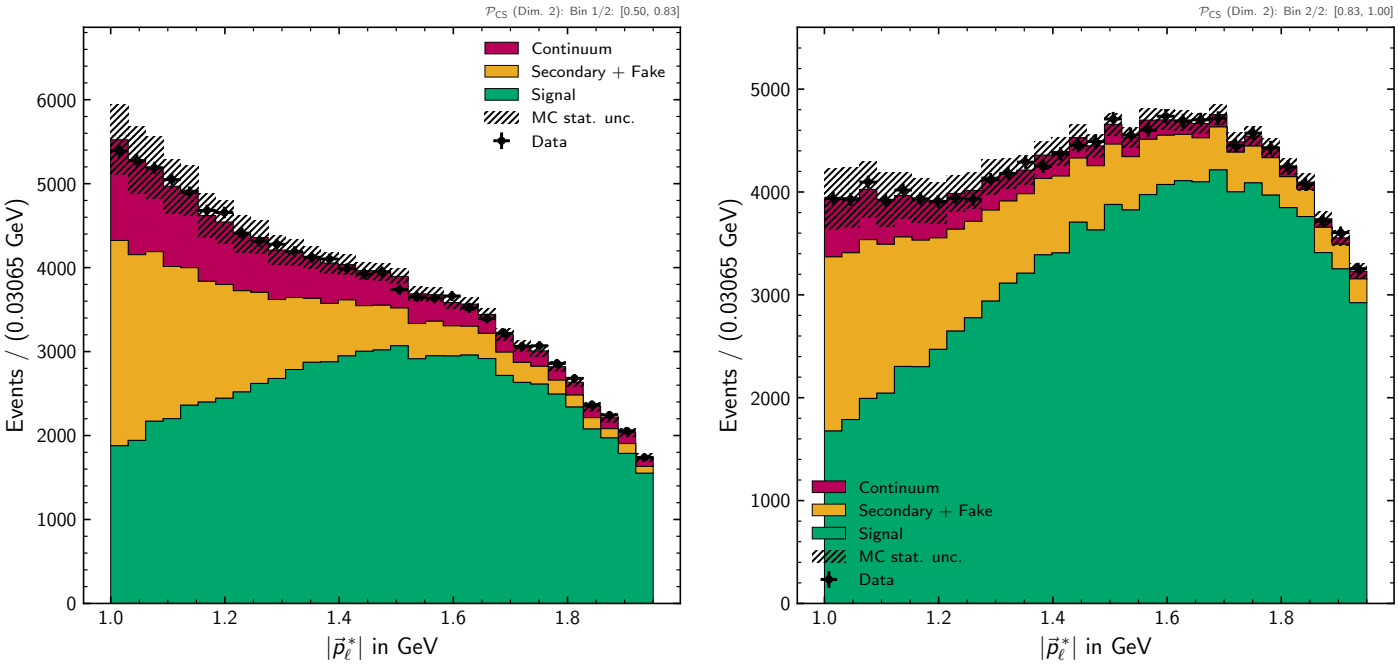


(b) $w_{\text{sig}} = 0.9$, $w_{\text{continuum}} = 0.9$, $w_{\text{fake+sec}} = 1.2$

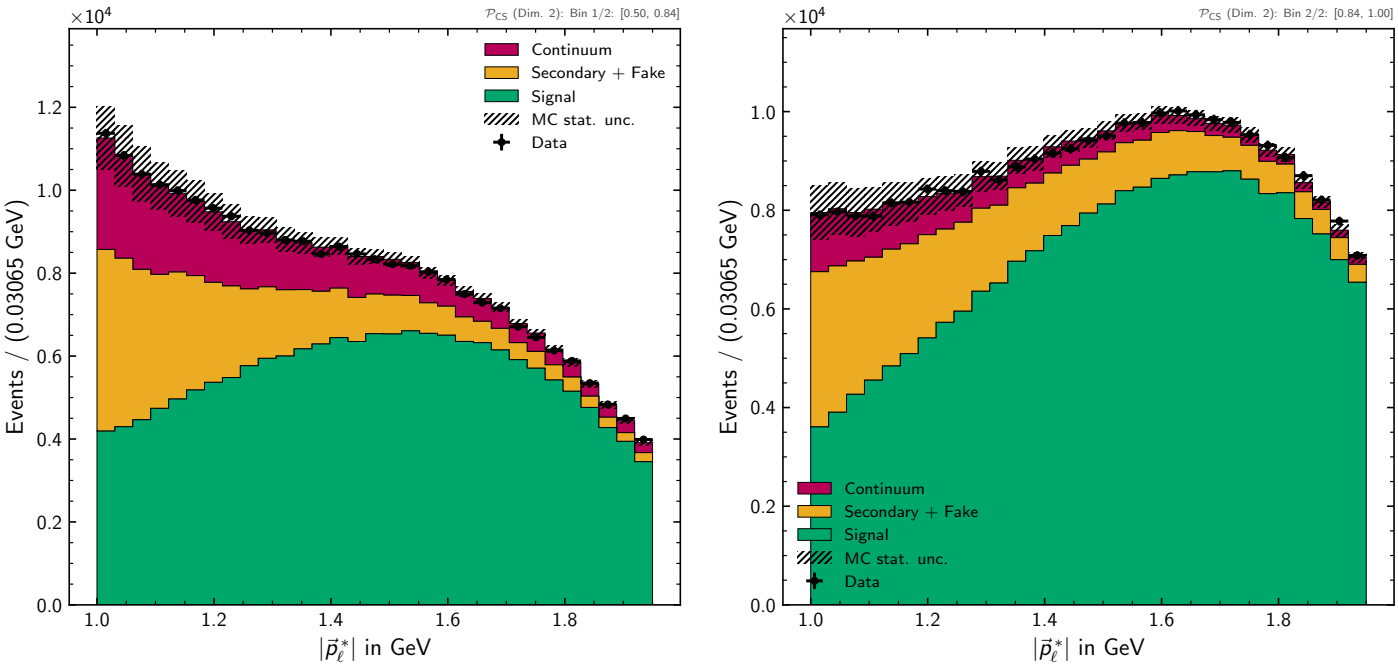


(c) $w_{\text{sig}} = 0.9$, $w_{\text{continuum}} = 1.3$, $w_{\text{fake}} = 1.1$, $w_{\text{secondary}} = 1.2$

Figure C.14: Comparing calibration factors with the expectation in bootstrap experiments with additional weights on the different components.



(a) Post-fit distribution for B^0



(b) Post-fit distribution for B^+

Figure C.15: Re-fitting the inclusive dataset with already applied calibration factors: Post-fit distributions.

Appendix D

Supplementary material and additional side studies for the calibration

This chapter contains a smorgasbord of additional material left out in various chapters of this thesis to help readability. It also includes the *tag track side-study*: In this study, we split up the D reconstruction modes for several sizable reconstruction modes by their number of tracks and compare the calibration factors. The observations were previously used to assign an uncertainty on the potentially neglected impact of sub-decay channels, which is now covered by the study shown in Appendix B.1. It is, however, still instructive because it provides some direct side-to-side comparison of the calibration factors for different sub-decay channels and shows a very large (8 %!) difference.

D.1 Additional material for the dataset and reconstruction

D.1.1 Samples

Several details regarding the MC datasets

- For generic $b \to c$ MC in the tagging calibration, we use stream 0 for experiments ≥ 31 (new tracking) and stream 10 for experiments < 31 (old tracking) to match the tracking used in the experimental data. For the $|V_{cb}|$ analysis 0-9 and 10-19 are used.
- $B \to u\ell\nu$ Belle special MC has been produced at a $20\times$ higher cross section than the reality. A 1/20 scaling weight is applied accordingly. A tiny fraction ($< 0.01\,\%$) of $B^0 \to u\ell\nu$ events cannot be processed and leads to segmentation faults in the processing. These problematic events cannot easily be excluded. As a workaround, the $B^0 \to u\ell\nu$ MC is processed in very small chunks of events. A weight of $w_{\text{miss}} \approx 1.31$ balances out the number of chunks that fail to process because problematic events have been encountered.

D.1.2 Branching ratios

The updated D meson branching ratios below one percent are listed in Table D.1.

D.1.3 Reconstruction

Technical details regarding the reconstruction presented in Chapter 8.

- Track smearing and track momentum smearing corresponds to the `SmearTrack=2` settings in newer versions of `b2bii` or `Smear_trk\2` in `basf`,
- The FEI is set up with the `feiv4/Belle1_2017_convertedMC_Track14_2` training.

Decay	BR [%]
$K^-\pi^+\eta$	0.92
$\bar{K}^0\pi^-\pi^-$	0.91
$\pi^+\pi^+\pi^-\pi^-$	0.76
$K^0K^-\pi^+$	0.66
$K^-\pi^+\eta'$	0.64
$K^-\rho^+\pi^-$	0.63
$\bar{K}_0^{*0}\pi^-$	0.59
$K^-\pi^+\rho^0$	0.53
$\rho^-\pi^+$	0.52
$K^-\pi^+\pi^+\pi^-\pi^-$	0.49
$\bar{K}^0K^0\pi^+\pi^-$	0.49
$\bar{K}^0K^+K^-$	0.48
$\bar{K}^0K^+\pi^-$	0.43
$\pi^+\pi^-\pi^+\pi^-\pi^-$	0.42
K^+K^-	0.41
$K_1^-\pi^+$	0.39
$K^+K^-\pi^+\pi^-\pi^-$	0.31
$\pi^-e^+\nu_e$	0.29
$K^+K^-\pi^-\pi^-$	0.28
$\pi^-\mu^+\nu_\mu$	0.27
$K_0^{*-}\pi^+$	0.27
$K^{*+}K^-$	0.25
$\phi\bar{K}^0$	0.20
$f_0'\bar{K}^0$	0.12
$f_0'\bar{K}^0$	0.12
$K^-\pi^+\pi^-\pi^-\pi^-$	0.00
$K^+K^-\bar{K}^0\pi^-$	0.00
$K^-\pi^-e^+\nu_e$	0.00
$\bar{K}^0\pi^-\pi^-\pi^-$	0.00
$K^-\pi^-\mu^+\nu_\mu$	0.00
$\pi^+\pi^-\pi^-\pi^-\pi^-$	0.00

(a) D^0 branching ratios

Decay	BR [%]
$K^+K^-\pi^+$	0.99
$\bar{K}^{*0}\rho^+$	0.93
$K^{*-}\pi^+\pi^+$	0.93
$K^-\rho^+\pi^+$	0.70
$\bar{K}^0\pi^+\pi^+\pi^-\pi^-$	0.67
$\bar{K}^0K^+$	0.61
$\bar{K}_2^{*0}e^+\nu_e$	0.50
$\bar{K}_2^{*0}\mu^+\nu_\mu$	0.50
$\eta'\pi^+$	0.50
$\pi^+\pi^-\pi^-$	0.47
$\pi^+\pi^-\pi^-\pi^-$	0.39
$\bar{K}^{*0}\omega\pi^+$	0.39
$K^-\pi^+\pi^+\pi^-\pi^-$	0.39
$\bar{K}^0\omega\pi^+$	0.39
$\eta\pi^+$	0.38
$\pi^-e^+\nu_e$	0.37
$\pi^-\mu^+\nu_\mu$	0.35
$\pi^+\pi^+\pi^-$	0.33
$\phi\pi^+$	0.32
$\bar{K}^0\pi^+\pi^-\pi^-\pi^-$	0.27
$\bar{K}^{*0}K^+$	0.25
$\rho^0\mu^+\nu_\mu$	0.24
$\bar{K}^{*0}\rho^0\pi^+$	0.23
$\rho^0e^+\nu_e$	0.22
$\omega\mu^+\nu_\mu$	0.17
$\omega e^+\nu_e$	0.17
$\pi^+\pi^+\pi^+\pi^-\pi^-$	0.17
$\eta\pi^+\pi^-$	0.14
$\pi^-\pi^+$	0.12
$\bar{K}^{*0}\pi^+\pi^+\pi^-$	0.12
$\tau^+\nu_\tau$	0.12
$\eta\mu^+\nu_\mu$	0.11
$\eta e^+\nu_e$	0.11
$\bar{K}_1^{0\prime}e^+\nu_e$	0.11
$\bar{K}''^{*0}\pi^+$	0.10
$K^+K^{*-}\pi^+$	0.08
$\bar{K}^{*0}\eta\pi^+$	0.00
$\bar{K}^0\pi^-e^+\nu_e$	0.00
$\bar{K}^{*0}K^{*+}$	0.00
$\bar{K}^0\eta\pi^+$	0.00
$\bar{K}^0K^{*0}\pi^+$	0.00
$\bar{K}^0\pi^+\pi^-\pi^-$	0.00
$\bar{K}^0\pi^-\mu^+\nu_\mu$	0.00
$\pi^+\pi^+\pi^+\pi^-\pi^-\pi^-$	0.00
$K^{*-}\rho^+\pi^+$	0.00

(b) D^+ branching ratios

Table D.1: D branching ratios. Only branching ratios below 1 % are shown. The large branching ratios are shown in Table 7.5.

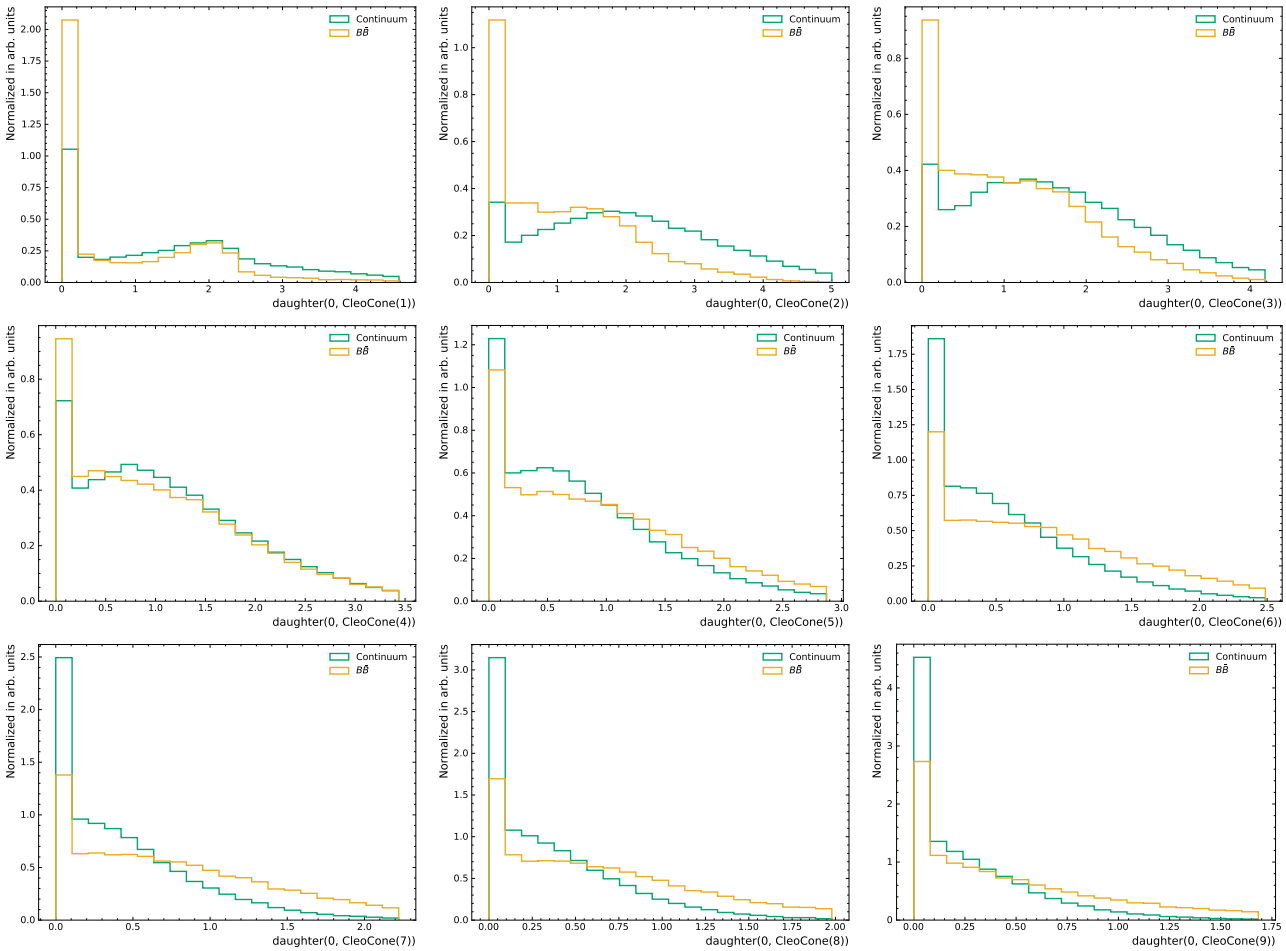


Figure D.1: Cleo cones.

D.2 Additional plots for continuum suppression

This appendix shows plots of all variables used in the continuum suppression. Note that they are not calibrated/fitted, so discrepancies between data and MC are expected.

- Cleo cones are shown in Figure D.1 and Figure D.4,
- KSFW Variables are shown in Figure D.2 and Figure D.5,
- Other variables are shown in Figure D.3 and Figure D.6.

D.3 Categorization

D.3.1 Uncertainties related to the granularity of the categorization

As mentioned in Section 9.1, choosing categories for the calibration factors is a bias-variance tradeoff. The more granular the categorization becomes, the higher the variance.

This tradeoff can be seen in Figure D.7, which plots the statistical and systematic uncertainties of calibration factors vs. the respective number of data events. As expected, the systematic uncertainty is relatively independent, but the statistical uncertainty follows a $\alpha/\sqrt{n}$ distribution ($\alpha > 0$ a constant and n the number of events). The green dashed lines are the fitted $\alpha/\sqrt{n}$ curves to the available data.

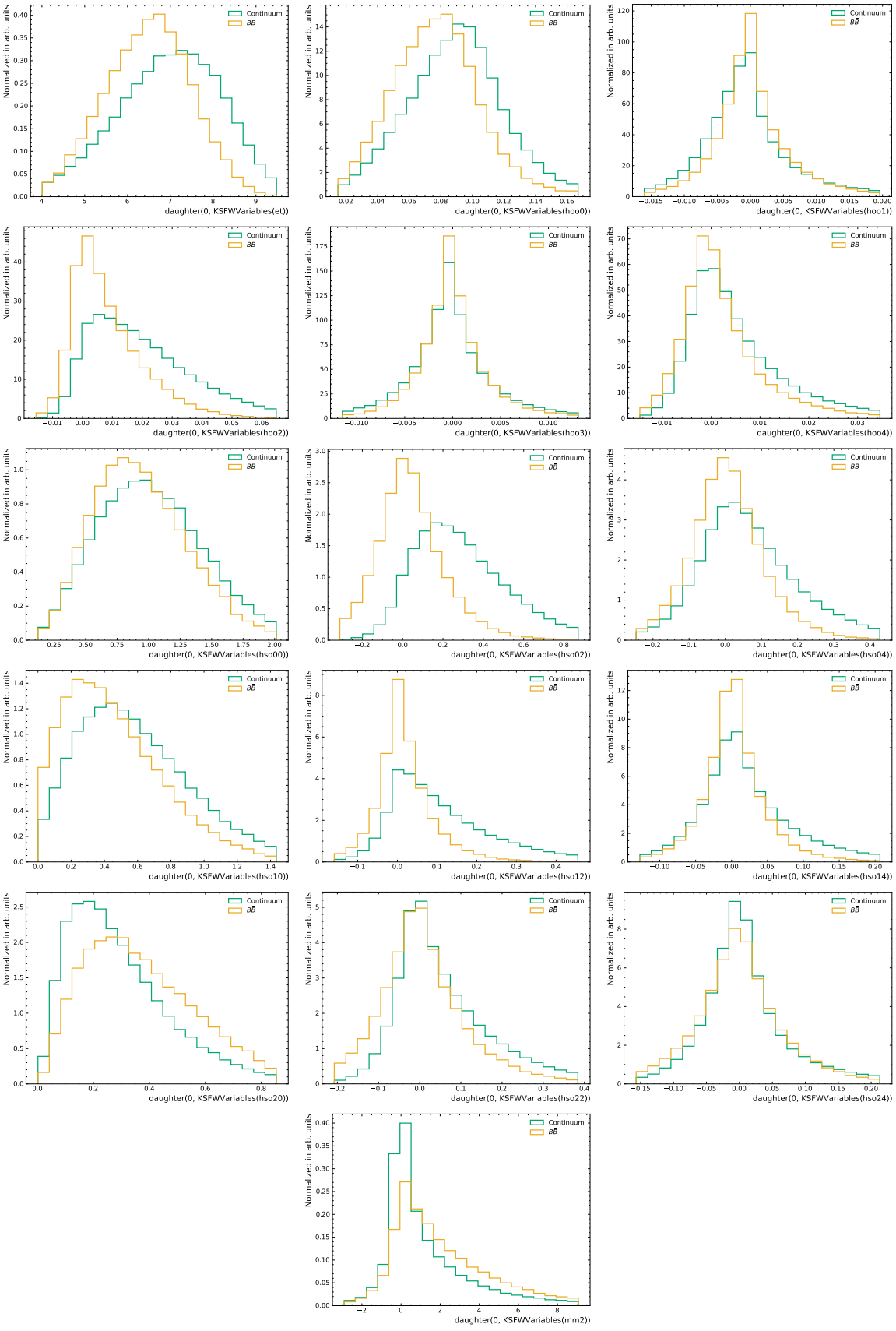


Figure D.2: KSFW variables.

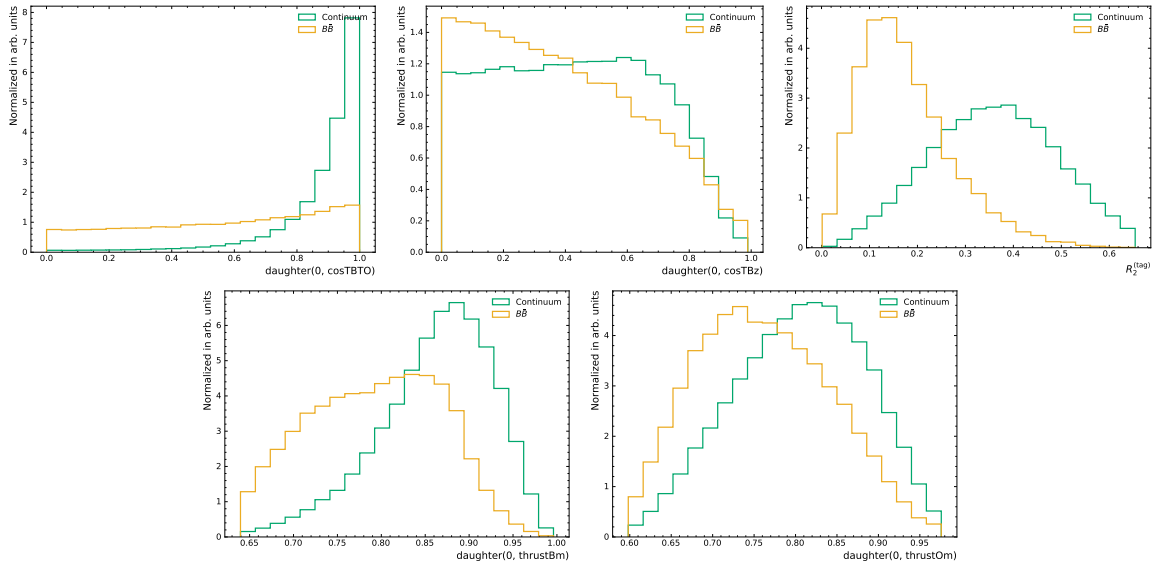


Figure D.3: Other variables used in the continuum suppression BDT.

These fitted curves are combined with the median systematic uncertainty in Figure D.8, showing the total uncertainty on the calibration factors. While the total uncertainty is dominated by systematic uncertainties for $n > 10^4$, the statistical uncertainty leads to a perceivable increase in the total uncertainty for $n < 10^4$.

D.3.2 Fraction of events per category

Figure D.9 shows the number of MC events per category for different components and datasets. In particular, these counts will be the weights used in weighted averages of calibration factors as shown in Table 10.1 and used more extensively in Appendix B.1 and Appendix C.1.

D.4 Calibration factors and tag side tracks

In the categorization (Section 9.1) used for the results of Chapter 10, most of the reconstruction modes are only distinguished up to the first daughter particle level. An exception mode is only made for some of the most dominating reconstruction modes, for example, $\mathrm{B}_{\mathrm{tag}} \to \overline{\mathrm{D}}^0\pi^+\pi^+\pi^-\pi^0$. However, even there, only two significant D modes are considered separately, and the remaining 13 reconstruction modes are considered together (see Table D.2). By considering different categorizations in Appendix B.1, we have already assigned an uncertainty that might arise from this (necessary) simplification when applying the results to the $\mathrm{B} \to \mathrm{D}^*\ell\nu_\ell$ dataset.

This side-study was an earlier attempt to determine such an uncertainty with a different method (access to information of sub-decay channels was not available for technical reasons at the time). Rather than considering concrete sub-decay channels, the decay channels were sub-divided according to their number of tracks ($n_{\mathrm{track}}^{\mathrm{tag}}$, shown in Figure D.10). This division effectively split the D decay channels into two parts.

However, this study is still very instructive:

- A different option to split the sub-decay channels is used. At equal statistical uncertainties, there is a direct tradeoff between the number of bins in $\mathcal{P}_{\mathrm{FEI}}$ and the granularity of the dissection of sub-decay channels. Therefore, investigating different divisions of the sub-decay channels into groups can be very helpful. Considering the large differences observed when splitting by $n_{\mathrm{track}}^{\mathrm{tag}}$, this grouping seems to be particularly effective.

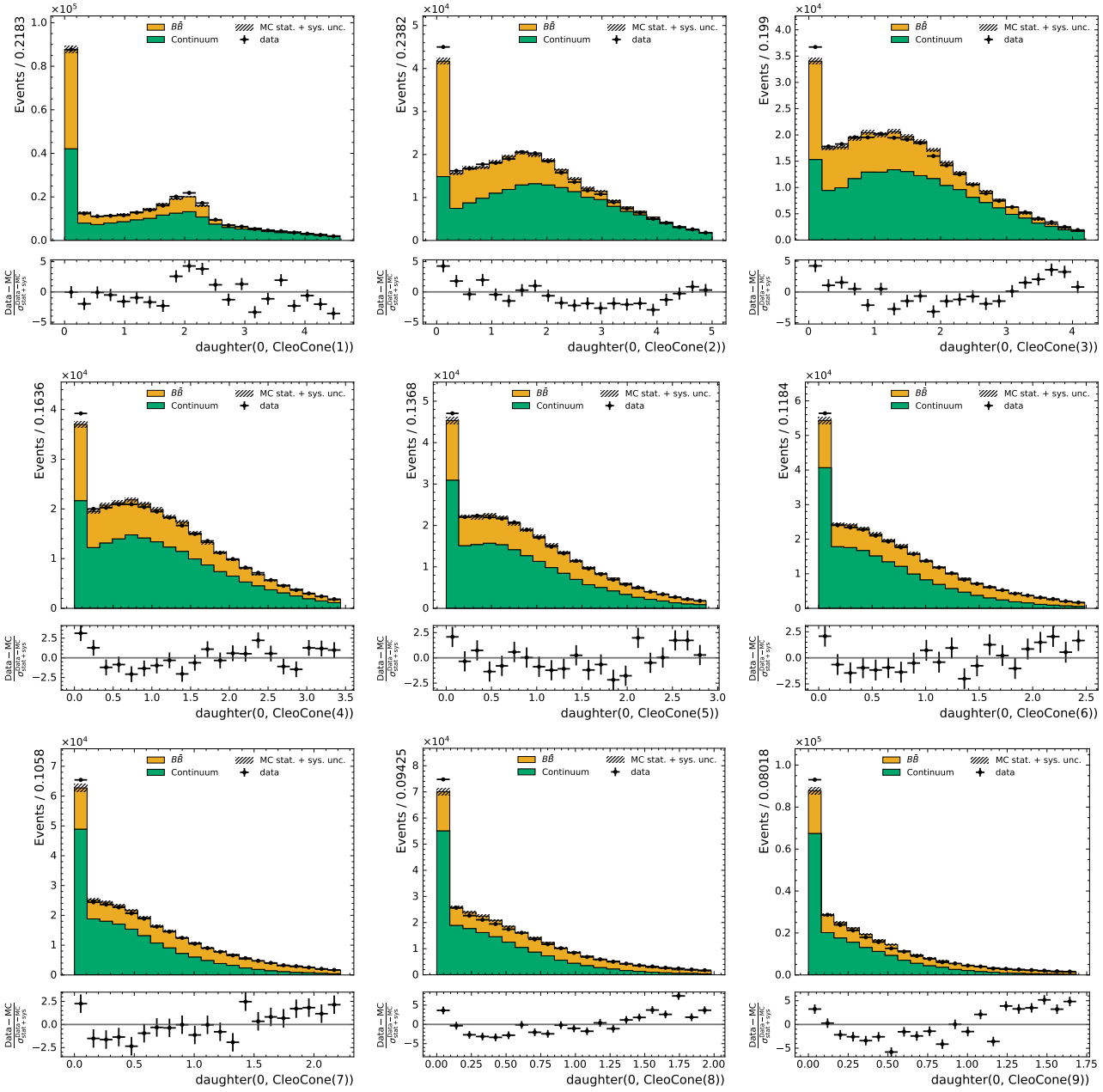


Figure D.4: Cleo cones. The plots do not include calibration.

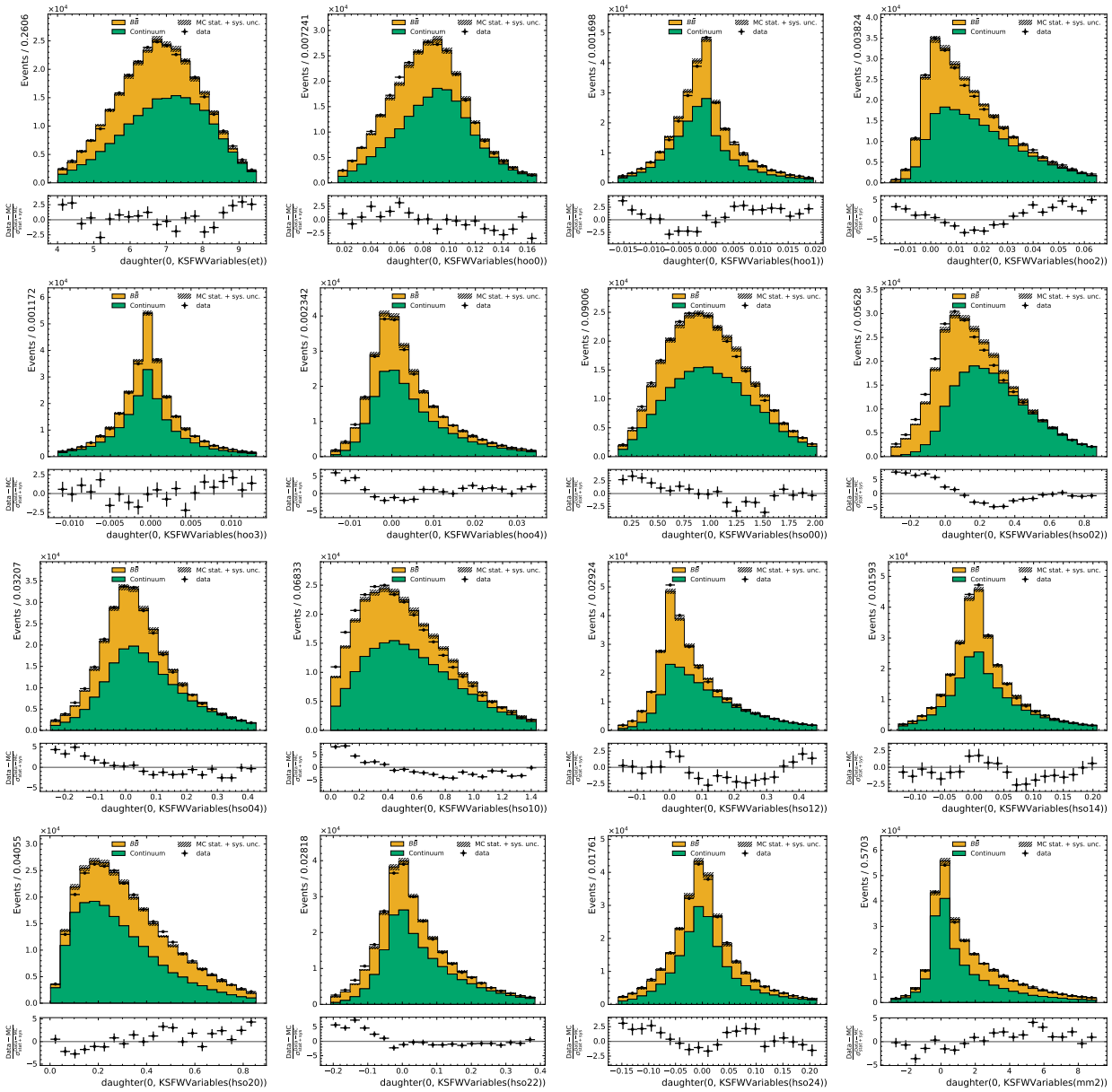


Figure D.5: KSFW variables. The plots do not include calibration.

- We quantify the difference in calibration factors between sub-decay channels and the difference between the relative fraction of sub-decay channels between the inclusive and the exclusive datasets. The difference in calibration factors is found to be as high as 8 %! It is only because the relative fractions of the two sub-decay channel groups are very similar between the datasets that this does not become a very problematic issue.

As we can see in Figure D.10b, the B^0 decays are dominated mainly by one value of $n_{\text{track}}^{\text{tag}}$ and are thus less suited for our side-study. From the B^+ decays, we pick the following two subsets:

(a) $\overline{D}^0\pi^+\pi^+\pi^-\pi^0$, $\overline{D}^{*0}\pi^+\pi^+\pi^-\pi^0$: For each of these decay channels, $n_{\text{track}}^{\text{tag}} = 5$ and $n_{\text{track}}^{\text{tag}} = 7$ are the most frequent number of tag tracks and both occur relatively often.

(b) $\overline{D}^0\pi^+\pi^0$, $\overline{D}^0\pi^+\pi^0\pi^0$, $\overline{D}^{*0}\pi^+\pi^0$, $\overline{D}^{*0}\pi^+\pi^0\pi^0$. Here $n_{\text{track}}^{\text{tag}} = 3$ and $n_{\text{track}}^{\text{tag}} = 5$ occur most often.

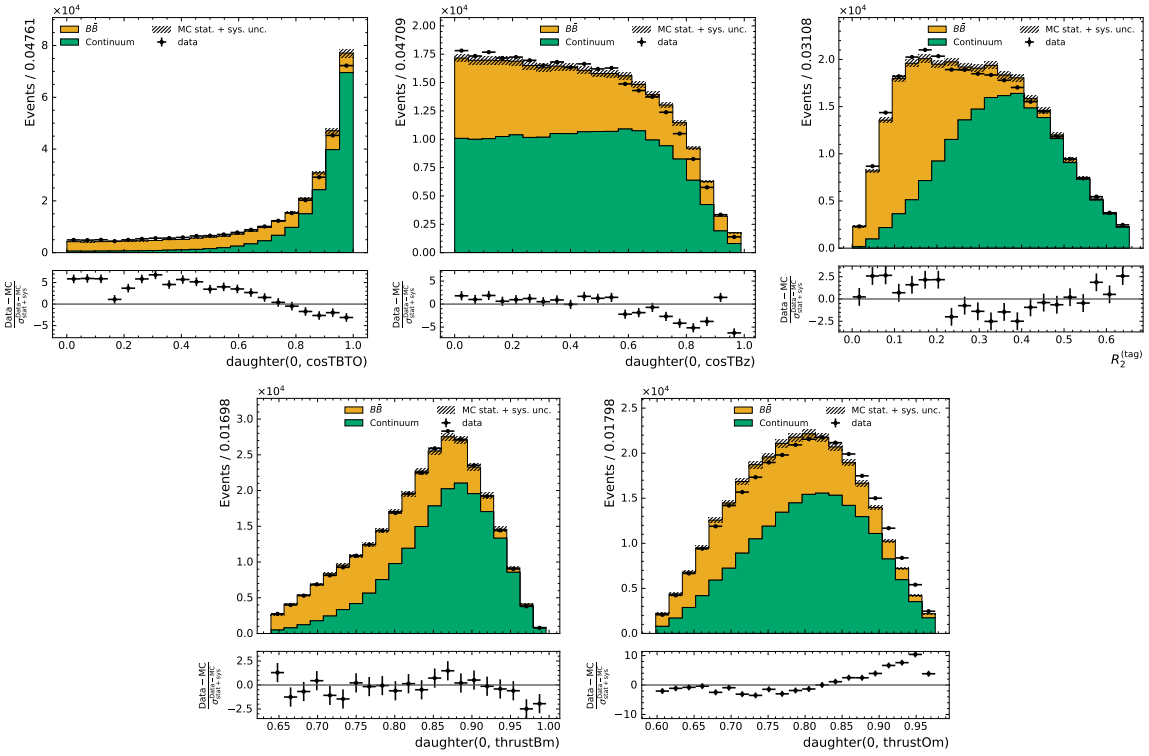
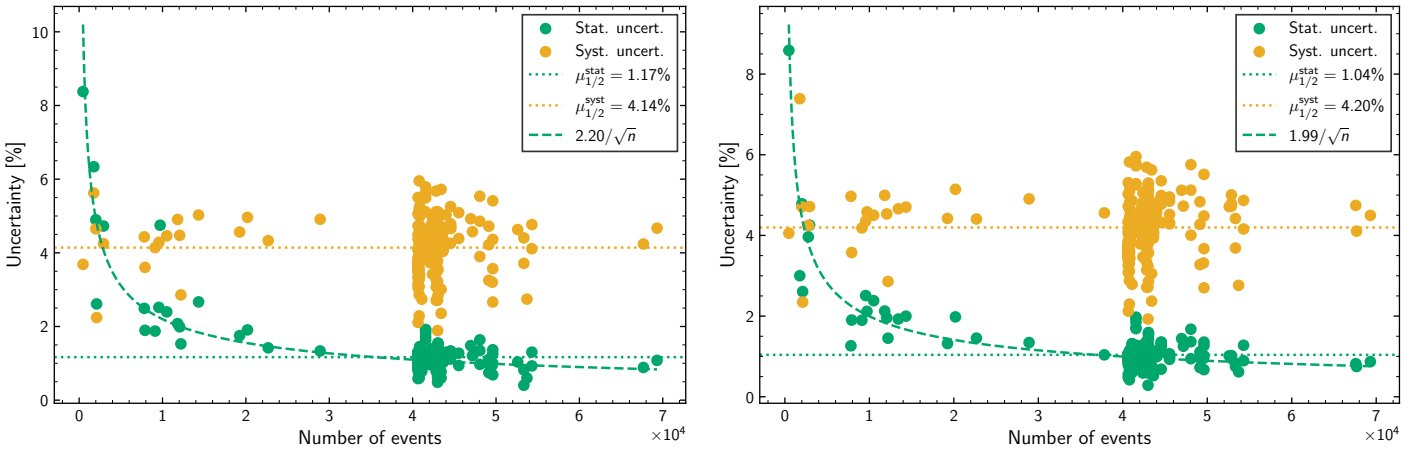


Figure D.6: Other variables used in the continuum suppression BDT. The plots do not include calibration.



(a) Primary fit (9 bins in $|\vec{p}_\ell^{\,*}|$)

(b) Fallback fit (9 bins in $|\vec{p}_\ell^{\,*}|$)

Figure D.7: Statistical and systematic uncertainty on the calibration factors vs. number of data events.

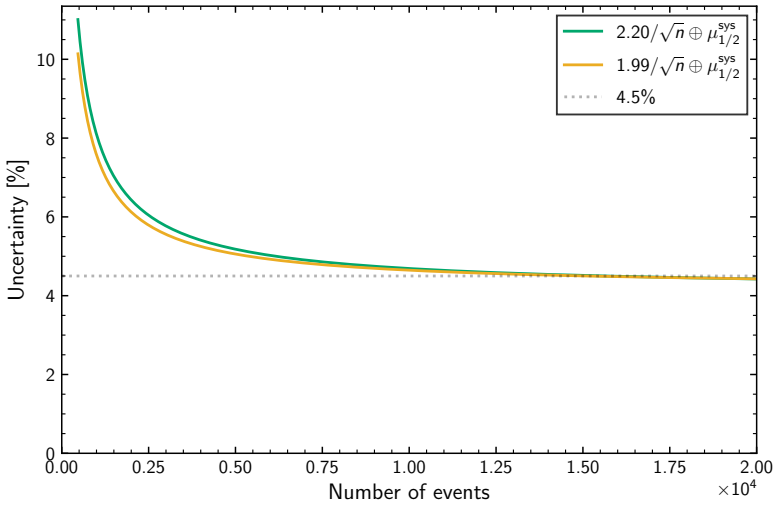


Figure D.8: Projection of total uncertainty on the calibration factors vs. number of data events.

D^{*+}	BR	D^{*0}	BR	D^+	Trk	BR	D^0	Trk	BR
$D^0\pi^+$	67.7	$D^0\pi^0$	64.7	$K^-\pi^+\pi^+$	3	9.4	$K^-\pi^+$	2	3.9
$D^+\pi^0$	30.7	$D^0\gamma$	35.3	$K^-\pi^+\pi^+\pi^0$	3	6.3	$K^-\pi^+\pi^0$	2	14.4
$D^+\gamma$	1.6			$K^-K^+\pi^+$	3	1.0	$K^-\pi^+\pi^0\pi^0$	2	8.9
				$K^-K^+\pi^+\pi^0$	3	0.7	$K^-\pi^+\pi^+\pi^-$	4	8.2
				$\pi^+\pi^0$	1	0.1	$K^-\pi^+\pi^+\pi^-\pi^0$	4	4.3
				$\pi^+\pi^+\pi^-$	3	0.3	$\pi^-\pi^+$	2	0.1
				$\pi^+\pi^+\pi^-\pi^0$	3	0.1	$\pi^-\pi^+\pi^0$	2	1.5
				$K^0_S\pi^+$	1/3	1.6	$\pi^-\pi^+\pi^0\pi^0$	2	1.0
				$K^0_S\pi^+\pi^0$	1/3	7.4	$\pi^-\pi^+\pi^+\pi^-$	4	0.8
				$K^0_S\pi^+\pi^+\pi^-$	3/5	3.1	$K^0_S\pi^0$	0/2	1.2
				$K^+K^0_SK^0_S$	1/3/5	0.3	$K^0_S\pi^+\pi^-$	2/4	2.8
							$K^0_S\pi^+\pi^-\pi^0$	2/4	5.2
							K^-K^+	2	0.4
							$K^-K^+\pi^0$	2	0.3
							$K^-K^+K^0_S$	2/4	0.4

Table D.2: Decay channels of $D^{(*)}$ mesons used in the reconstruction by the FEI. The *Trk* column counts the number of tracks in the decay of the decay channel to the left, the *BR* column lists the branching ratio in percent. Note that K^0_S can decay as $\pi^0\pi^0$ (30.7 %) or $\pi^+\pi^-$ (69.2 %).

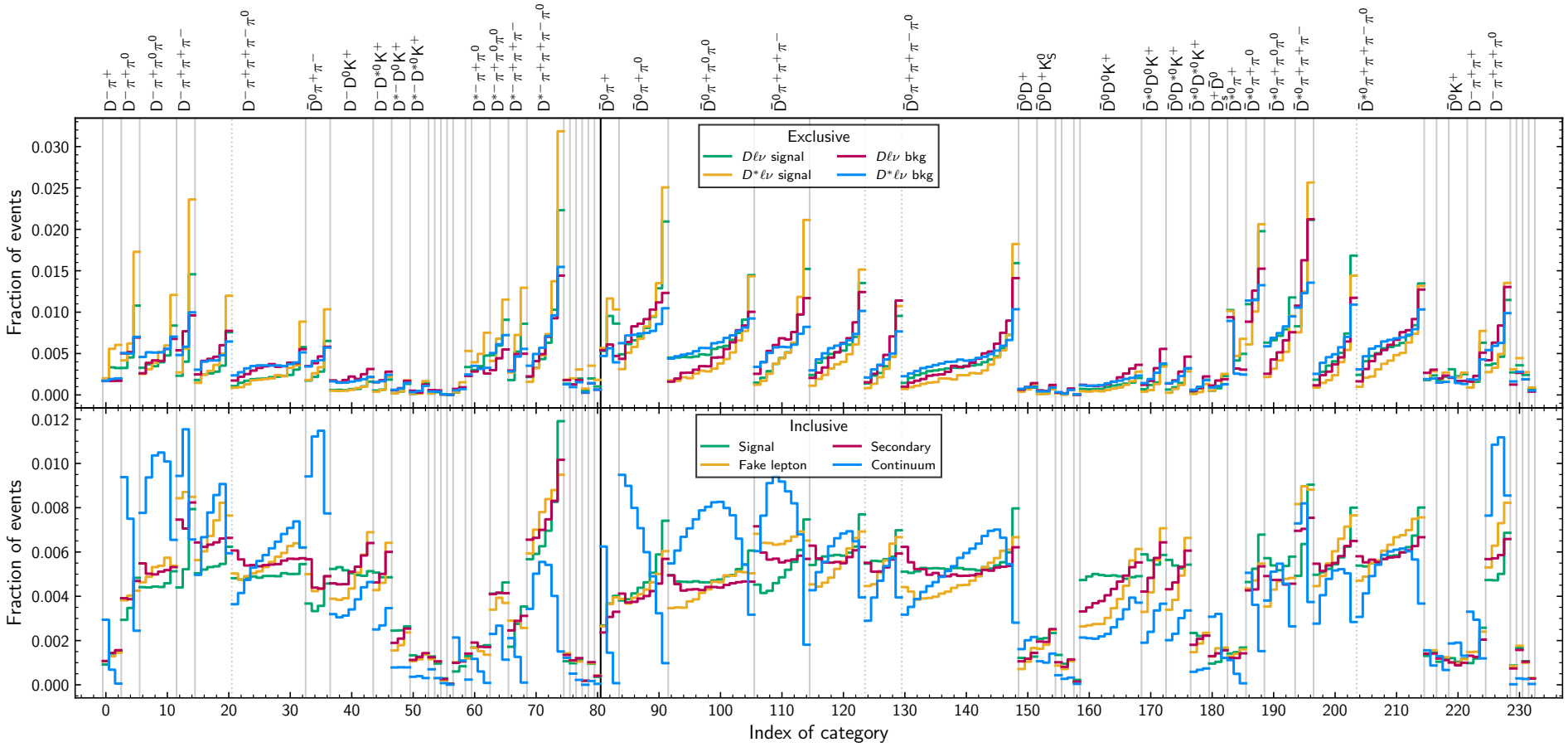
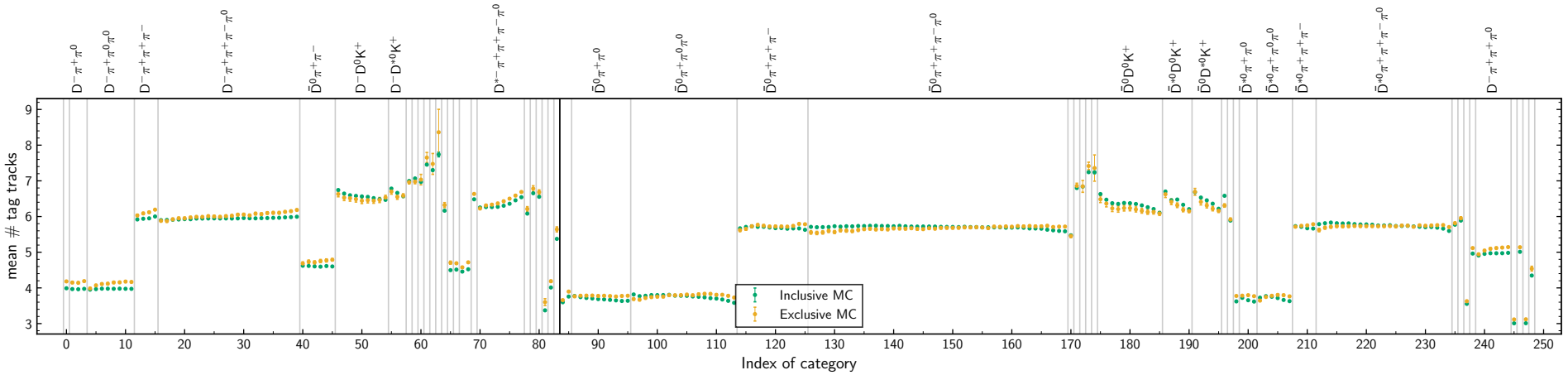
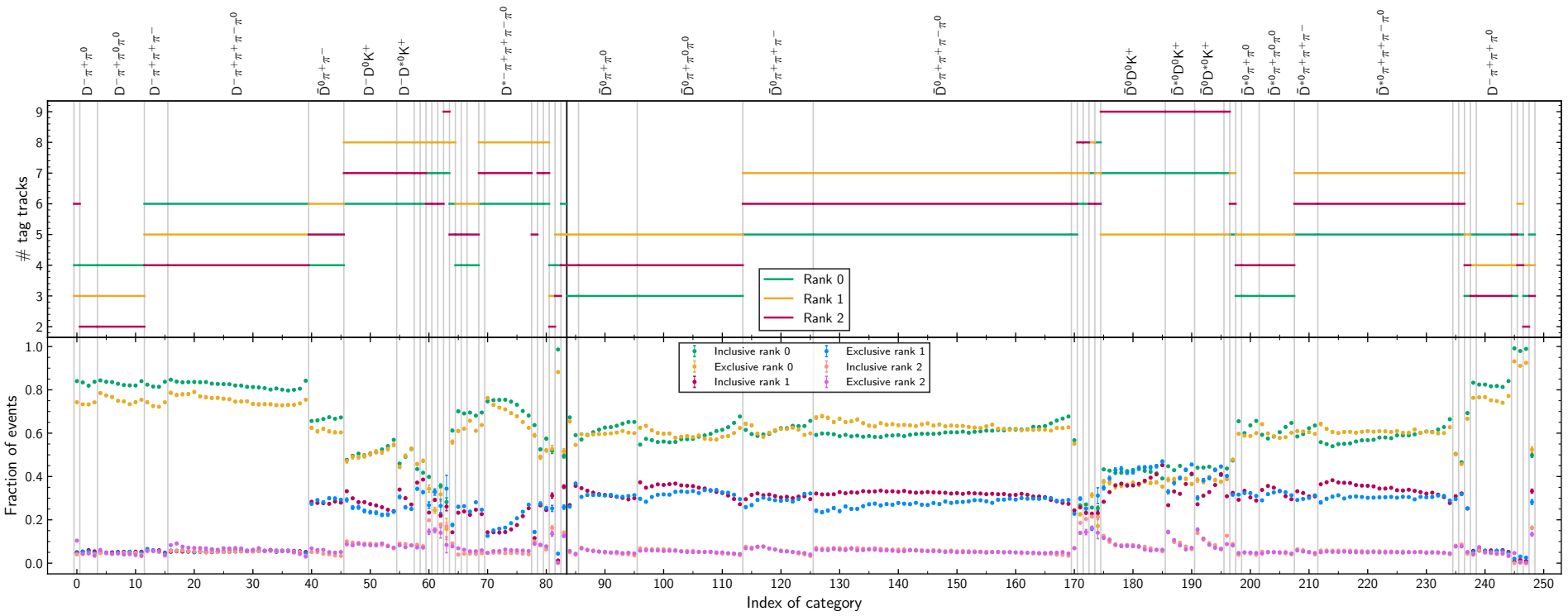


Figure D.9: Fraction of MC events in categories for both the exclusive and inclusive datasets. The fractions are used as weights for the calculation of average calibration factors, such as presented in Table 10.1 and Appendix B.1.

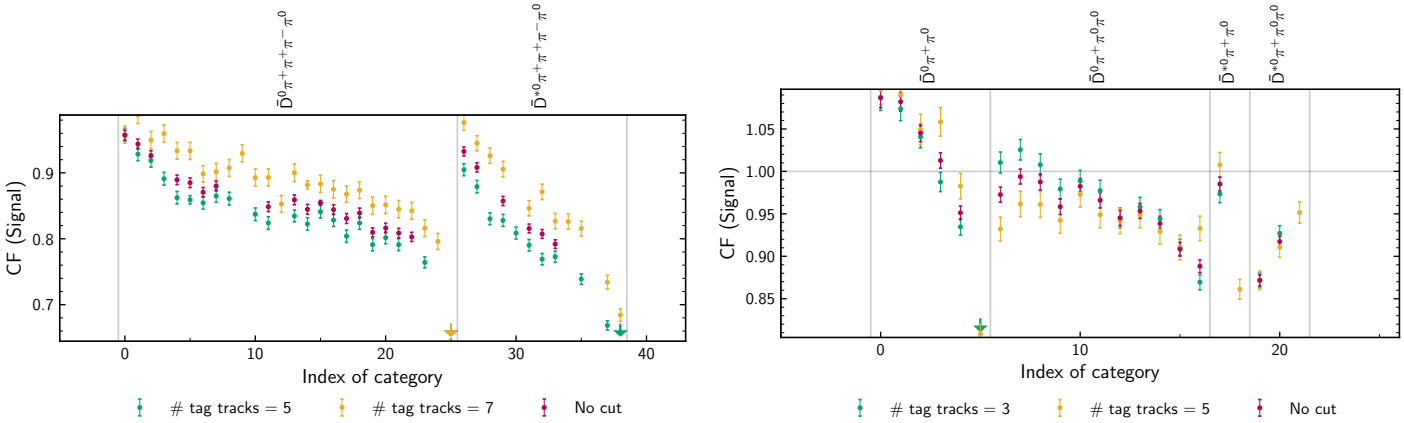


(a) Mean number of tag tracks.



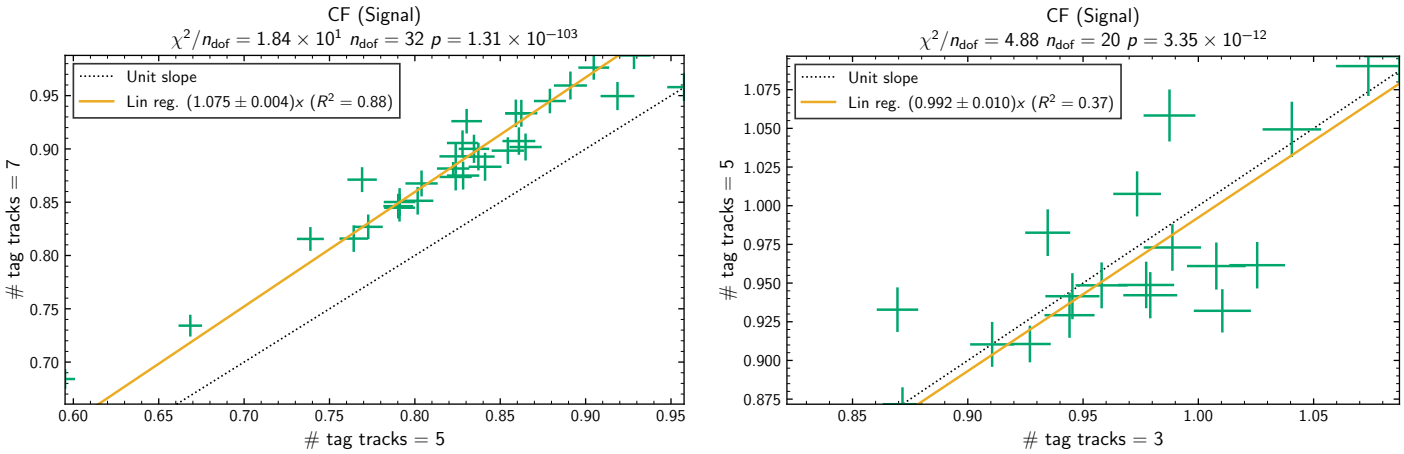
(b) Fraction occupied by the most popular numbers of tag tracks. The upper half of the figure shows the three most popular $n_{\text{track}}^{\text{tag}}$ values for the inclusive (MC) dataset, the lower half shows their frequencies in the inclusive and exclusive dataset.

Figure D.10: Number of tag tracks in the inclusive and exclusive dataset.



(a) $n_{\text{track}}^{\text{tag}} = 5, 7$ (b) $n_{\text{track}}^{\text{tag}} = 3, 5$

Figure D.11: Calibration factors for specific numbers of tag tracks. *No cut* refers to calibration factors derived without a $n_{\text{track}}^{\text{tag}}$ cut (i.e., in the "normal" way).



(a) $n_{\text{track}}^{\text{tag}} = 5, 7$ (b) $n_{\text{track}}^{\text{tag}} = 3, 5$

Figure D.12: Calibration factors for specific numbers of tag tracks.

Because $D^{*+} \to D^0\pi^+$ and $D^{*+} \to D^0\gamma$ are not distinguishable by their number of tracks, this essentially amounts to splitting the D^0 decay modes into two parts (see Table D.2), ignoring only the $D^0 \to K_S^0\pi^0$ decay. The D^0 decays with two tracks have a higher branching fraction than those with four tracks (which are further suppressed by a factor of $0.85^2 \approx 72\,\%$ because of the track finding efficiency of the two additional tracks), which can also be seen in Figure D.11b. The tiny, but non-vanishing fraction of reconstruction modes with an even number of tag tracks is likely due to incorrect counting (see Appendix D.5).

We now consider the subsets of decay channels of case a (b) in the inclusive dataset and further split it up in $n_{\text{track}}^{\text{tag}} = 5$ and $n_{\text{track}}^{\text{tag}} = 7$ ($n_{\text{track}}^{\text{tag}} = 3$ and $n_{\text{track}}^{\text{tag}} = 5$). For each of these datasets, we extract calibration factors using the 1D $|\vec{p}_\ell^{\,*}|$ fits (without systematic uncertainties other than MC statistics). The results are shown in Figures D.11 and D.12.

While case b mostly shows a decorrelating effect with relatively little bias, a shows a very clear shift: The calibration factor for $n_{\text{track}}^{\text{tag}} = 7$ is almost 8 % higher than for $n_{\text{track}}^{\text{tag}} = 5$ (Figure D.12a)! This is also interesting, because both a and b consider the same split in the D^0 decay modes: The impact of the D reconstruction modes cannot be factored out, but is different for different decay modes. Similar to the argument of Appendix B.2.1, this also definitely cannot be explained

with a difference in tracking efficiency between data and MC, as the deviations of case a and b run in different directions (and the effect in case of a is an order of magnitude too large).

We now want to quantify the effect of this shift on the averaged calibration factors. For this, we adopt almost the same procedure as in Appendix B.2: Let $\epsilon_i^{(1)}$ and $\epsilon_i^{(2)}$ be the calibration factors calculated with $n_{\text{track}}^{\text{tag}} = 5$ and $n_{\text{track}}^{\text{tag}} = 7$ ($n_{\text{track}}^{\text{tag}} = 3$ and $n_{\text{track}}^{\text{tag}} = 5$) and $f_i^{\text{incl}\,(k)}$ and $f_i^{\text{excl}\,(k)}$ ($k = 1, 2$) defined analogue to Equation (B.2.1). Further define $f^{\text{incl}\,(0)} := 1 - f^{\text{incl}\,(1)} - f^{\text{incl}\,(2)}$ ($f^{\text{excl}\,(0)}$ analogue) and let us denote the normal calibration factor without any $n_{\text{track}}^{\text{tag}}$ cut as $\epsilon_i^{(0)}$. Analogue to Equations (B.2.2) and (B.2.3), we define

$$\epsilon_i^{\text{incl}} = \sum_{k=0}^{2} f_i^{\text{incl}\,(k)} \epsilon_i^{(k)}, \tag{D.4.1a}$$

$$\epsilon_i^{\text{excl}} = \sum_{k=0}^{2} f_i^{\text{incl}\,(k)} \epsilon_i^{(k)}, \tag{D.4.1b}$$

where the only difference to Appendix B.2 is that we have substituted the average $\epsilon_i^{(0)}$ for the very small fraction of events that are not covered by $k = 1$ and $k = 2$. Again ϵ_i^{incl} is almost identical to the calibration factors without any $n_{\text{track}}^{\text{tag}}$ cut and is only calculated in this way to take care of correlations with ϵ_i^{excl}. As before, we also take care of (anti)correlations between the $f_i^{\text{incl/excl}\,(k)}$ of different k.

The results are shown in Figure D.13. The deviations barely exceed 0.4 % and generally decline with $\mathcal{P}_{\text{FEI}}$. The fact that we do not have a bigger issue is thanks to the fact that f_i^{incl} is relatively close to f_i^{excl} (Figure D.10b).[1]

To quantify the impact on the calibration factors, we now continue in the same way as before to assign uncertainties, adding additional uncertainties and performing χ^2 tests. This is shown in Figure D.14. Let us first consider case a. A correlated uncertainty of 0.2 % brings the bias significantly below 1σ. An additional uncorrelated uncertainty of 0.08 % then easily lifts the p-value above 10 %. Now considering case b, the same uncertainties bring the p-value close to one and the bias below 0.3.

Under the assumption that no decay modes are more impacted by the sub-decay channels than the ones we just investigated, we thus consider the uncertainties

$$e_{\text{corr}} = 0.2\,\%, \quad e_{\text{uncorr}} = 0.08\,\%. \tag{D.4.2}$$

as sufficient to cover the tension discovered in this side-study. Compared to the total uncertainty on the calibration factors, this is negligible. Therefore, we conclude that the issue of the neglected D sub-decay channels does not have a significant impact on the use case at hand: Not because sub-decay channels do not influence the calibration factors, but because the distribution of sub-decay channels is similar between the calibration dataset and the analysis dataset. This might, however, not be true for all analyses that use the FEI!

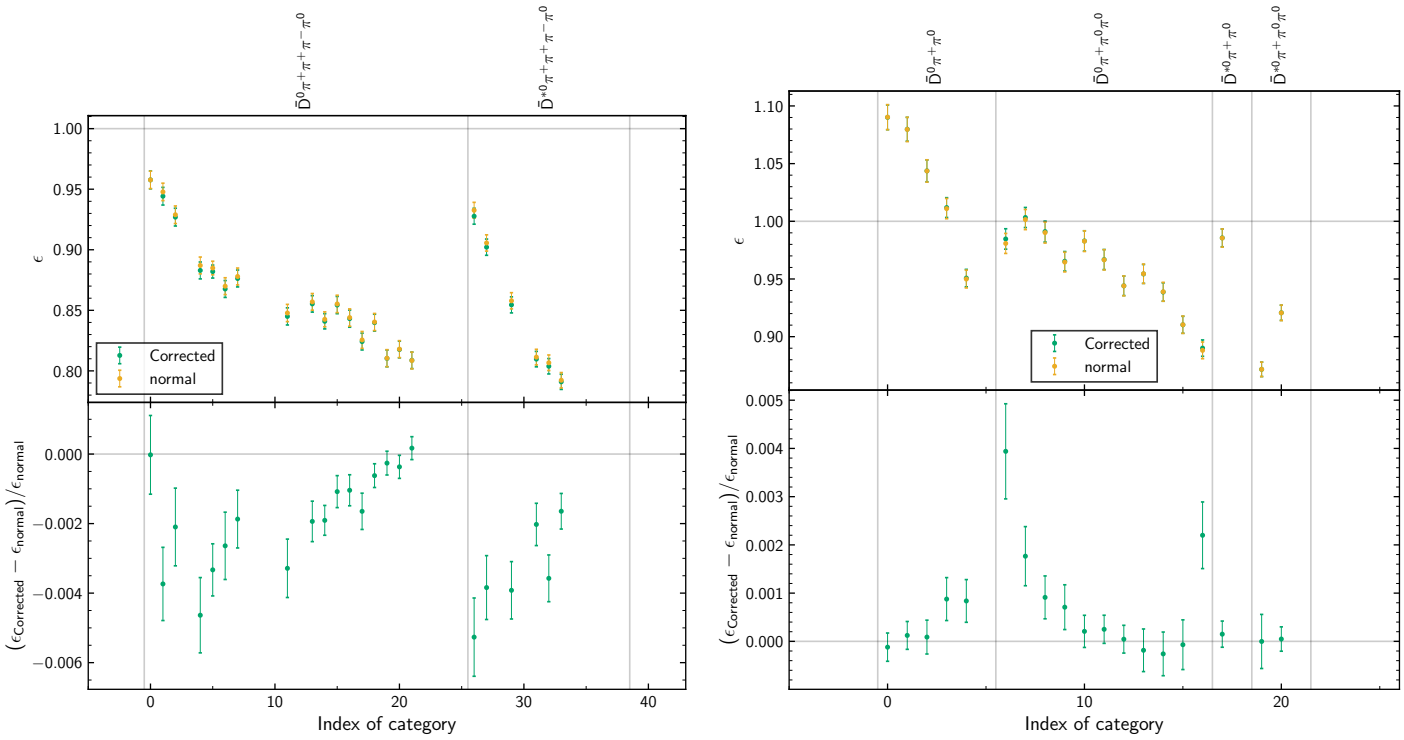
D.5 Counting tracks

In the inclusive sample, the number of tracks is calculated as

$$n_{\text{track}}^{\text{sig}} := n_{\text{track}}^{\text{ROE}} + 1, \tag{D.5.1a}$$

$$n_{\text{track}}^{\text{tag}} := n_{\text{track}} - n_{\text{track}}^{\text{sig}}, \tag{D.5.1b}$$

[1] As a quick back-of-the-envelope calculation: We see 8 % deviation between the two calibration factors in case a, but weighing that by a difference between f^{excl} and f^{incl} of 5 % brings us to around 0.4 %.



(a) $n_{\text{track}}^{\text{tag}} = 5, 7$ (b) $n_{\text{track}}^{\text{tag}} = 3, 5$

Figure D.13: Calibration factors built from a weighted average of results for different tag track multiplicities. *Corrected* refers to ϵ_i^{excl} (defined in Equation (D.4.1b)), *normal* to ϵ_i^{incl} (defined in Equation (D.4.1a)).

where $n_{\text{track}}^{\text{ROE}}$ is the number of tracks in the ROE (remember, this is everything but the lepton and the B_{tag}) with the appropriate ROE mask applied, i.e., with the requirements $\text{d}r < 2\,\text{cm}$, $|\,\text{d}z| < 4\,\text{cm}$ and `trackCurlerCloneIndicator` $= 0$. $n_{\text{track}}^{\text{sig}}$ is calculated with a cut only on the impact parameters ($\text{d}r < 2\,\text{cm}$, $|\,\text{d}z| < 4\,\text{cm}$). The same track quality requirement is used by the FEI (see Keck, 183). Because of the inconsistent use of the `trackCurlerCloneIndicator` $= 0$ cut, $n_{\text{track}}^{\text{tag}}$ might in rare cases differ from the number of tracks used by the FEI (see discussion regarding wrong track multiplicity parity in Appendix D.4).

In the exclusive sample, Equation (D.5.1b) still hold true, but $n_{\text{track}}^{\text{sig}}$ is calculated directly from the knowledge of the reconstructed decay channel. n_{track} is calculated identically to the inclusive sample. Subtle differences in the requirements on the signal tracks (such as the `trackCurlerCloneIndicator` $= 0$ requirement) might cause $n_{\text{track}}^{\text{tag}}$ to be different from the number of tracks used by the FEI in rare cases.

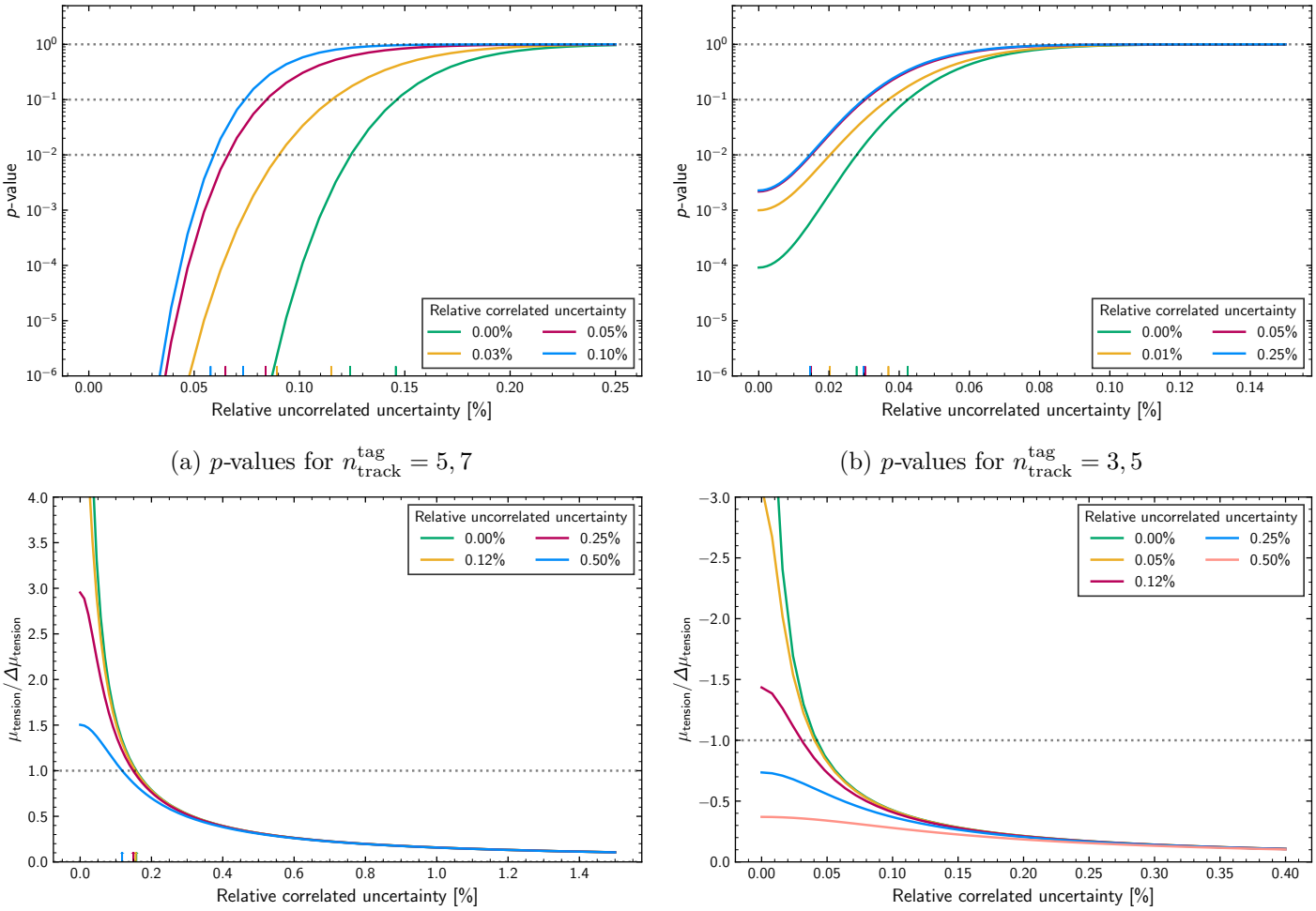
D.6 Additional material for the category migration study

This appendix expands on details about the side-study discussed Appendix C.3.

D.6.1 The overlap dataset

The overlap percentages in MC are presented in more detail in Table D.3.

One might ask why we do not find around 10 % percent of the events of the exclusive dataset in the inclusive dataset, given that the reconstruction in the inclusive dataset is strictly looser (after we apply the $|\vec{p}_\ell^{\,*}|$ cut to exclusive MC and select the same streams and data samples). To investigate this, we consider the exclusive dataset, apply all selections that are unique to the



(a) p-values for $n_{\text{track}}^{\text{tag}} = 5, 7$

(b) p-values for $n_{\text{track}}^{\text{tag}} = 3, 5$

(c) Bias for $n_{\text{track}}^{\text{tag}} = 5, 7$

(d) Bias for $n_{\text{track}}^{\text{tag}} = 3, 5$

Figure D.14: Assigning uncertainties in the tag track multiplicity side-study. Note that the meaning of the different colors and the abscissa is switched in the upper and lower plots. The colored lines on the abscissa denote the intersections with the dashed lines.

inclusive dataset and compare kinematic distributions between the events that are also present in the inclusive dataset, and the rest.

- The $m_{\text{bc}}^{\text{tag}}$ distribution of the non-overlapping events (Figure D.15a) is found to be strongly biased to low values. This might be due to the subtlety about the order of operations between the $|\vec{p}_\ell^{\,*}|$ cut and the best candidate selections that was described in Section 8.2.2.

- The bias in $|\vec{p}_\ell^{\,*}|$ (Figure D.15c) might be due to slightly decorrelated lepton momenta between the inclusive and exclusive dataset: Low-momentum leptons are sometimes filtered out by the $|\vec{p}_\ell^{\,*}|$ cut in only one of the datasets.

- The bias in the total photon energy $\sum E_\gamma$ (Figure D.15d) could stem from the requirement of having less than 18 photons in the inclusive dataset (which is not applied in the exclusive dataset). Unfortunately the photon count variable was not available in the exclusive dataset, but the shift towards higher photon energies suggests that this is indeed the case.

- The bias in $\Delta E_{\text{B}_{\text{tag}}}$ (Figure D.15b) might be related/correlated to the other effects.

Sample	Events	F. excl.	F. incl.
Charged	136652	91.05%	3.14%
Mixed	58896	90.16%	1.62%
$B^+ \to \bar{D}_0^{*0} e^+ \nu_e$	36668	87.77%	1.89%
$B^+ \to \bar{D}_0^{*0} \mu^+ \nu_\mu$	34049	88.14%	1.76%
$B^+ \to \bar{D}_1^0 e^+ \nu_e$	32513	85.99%	1.69%
$B^+ \to \bar{D}_2^{*0} e^+ \nu_e$	32339	86.01%	1.59%
$B^+ \to \bar{D}_1'^0 e^+ \nu_e$	32034	86.06%	1.68%
$B^+ \to \bar{D}_2^{*0} \mu^+ \nu_\mu$	31344	86.92%	1.52%
$B^+ \to \bar{D}_1^0 \mu^+ \nu_\mu$	31071	86.55%	1.61%
$B^+ \to \bar{D}_1'^0 \mu^+ \nu_\mu$	30917	86.61%	1.60%
$B^0 \to D_0^{*-} e^+ \nu_e$	23837	86.01%	1.45%
$B^0 \to D_1^- e^+ \nu_e$	23788	85.31%	1.49%
$B^0 \to D_1'^- e^+ \nu_e$	23392	85.29%	1.45%
$B^0 \to D_2^{*-} e^+ \nu_e$	23115	84.97%	1.33%
$B^0 \to D_1'^- \mu + \nu_\mu$	22668	85.94%	1.35%
$B^0 \to D_2^{*-} \mu^+ \nu_\mu$	22466	85.51%	1.26%
$B^0 \to D_1^- \mu^+ \nu_\mu$	22364	85.59%	1.38%
$B^0 \to D_0^{*-} \mu^+ \nu_\mu$	22326	86.45%	1.33%
$B^+ \to \bar{D}_1^0 (\to D\pi\pi) \ell^+ \nu_\ell$	21463	85.52%	1.66%
$B^0 \to D_1^- (\to D\pi\pi) \ell^+ \nu_\ell$	12299	85.00%	1.12%
Gap $B^+ \to \bar{D}^0 \eta \ell^+ \nu_\ell$	11430	81.06%	1.98%
Gap $B^+ \to \bar{D}^{*0} \eta \ell^+ \nu_\ell$	8469	77.78%	1.85%
Gap $B^+ \to D^* \pi\pi \ell^+ \nu_\ell$	6364	79.84%	0.82%
Gap $B^0 \to D^- \eta \ell^+ \nu_\ell$	5402	79.75%	1.05%
Gap $B^0 \to D^* \pi\pi \ell^+ \nu_\ell$	4343	79.50%	0.59%
Gap $B^0 \to D^{*-} \eta \ell^+ \nu_\ell$	4273	78.66%	0.97%
Charm	3091	67.14%	0.26%
$B^+ \to u \ell \nu$	2728	87.21%	0.16%
Gap $B^+ \to D\pi\pi \ell^+ \nu_\ell$	1713	81.18%	0.87%
$B^0 \to u \ell \nu$	1254	64.04%	0.13%
UDS	1172	64.47%	0.09%
Gap $B^0 \to D\pi\pi \ell^+ \nu_\ell$	1088	81.38%	0.61%

Table D.3: Overlap by MC data type.

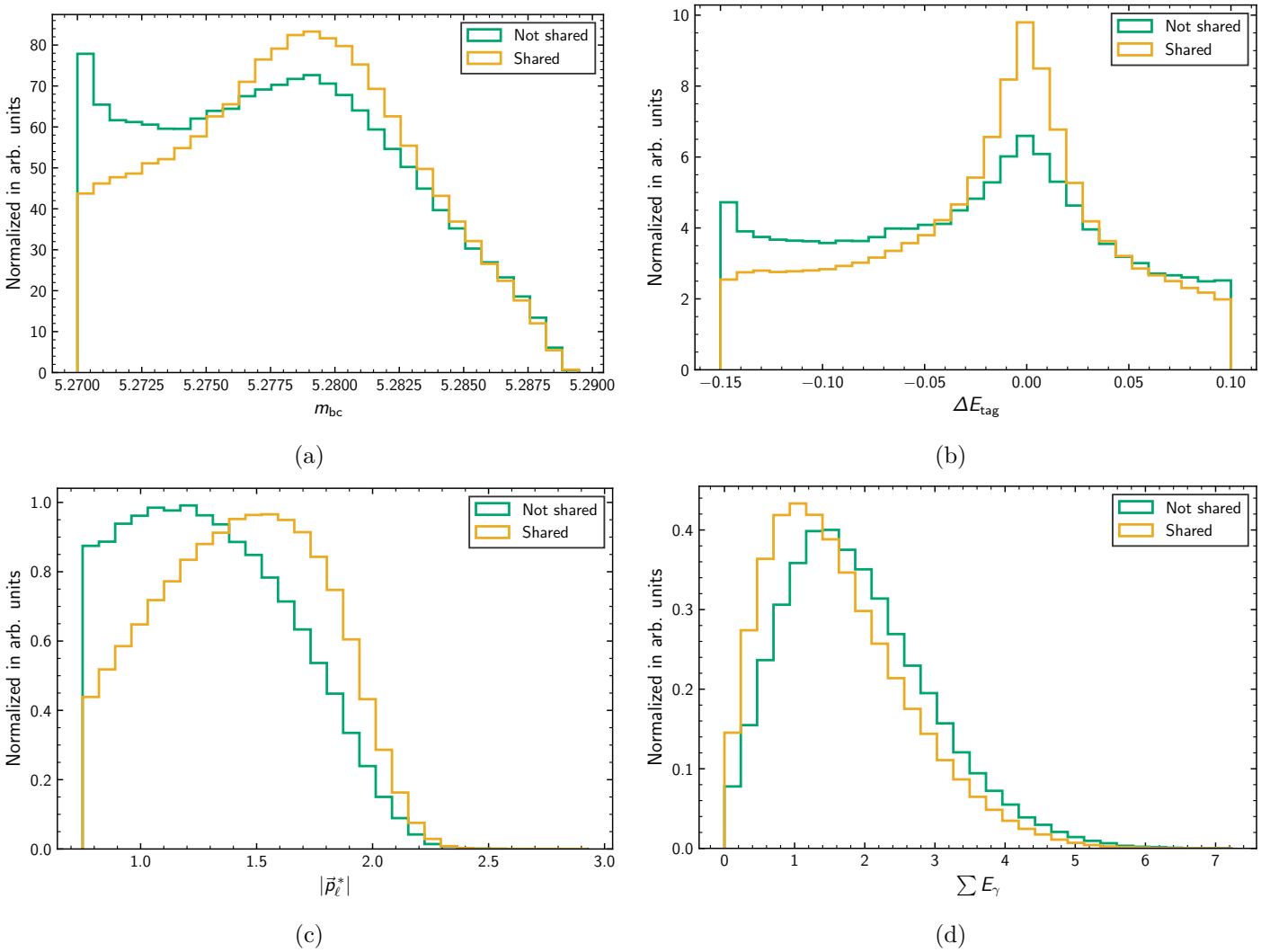


Figure D.15: Kinematic distributions that significantly differ between the overlapping and non-overlapping part of exclusive MC (already including cuts on stream, data samples and $|\vec{p}_\ell^*|$).

D.6.2 Migration matrices

Because we are only interested in the resulting event counts in each category (see remark i, page 182), we did not need to introduce migration matrices in Appendices C.3.2 to C.3.4. However, they do allow us to give a more detailed picture and to understand if the positive and negative migration contributions cancel, or if there is no migration in the first place.

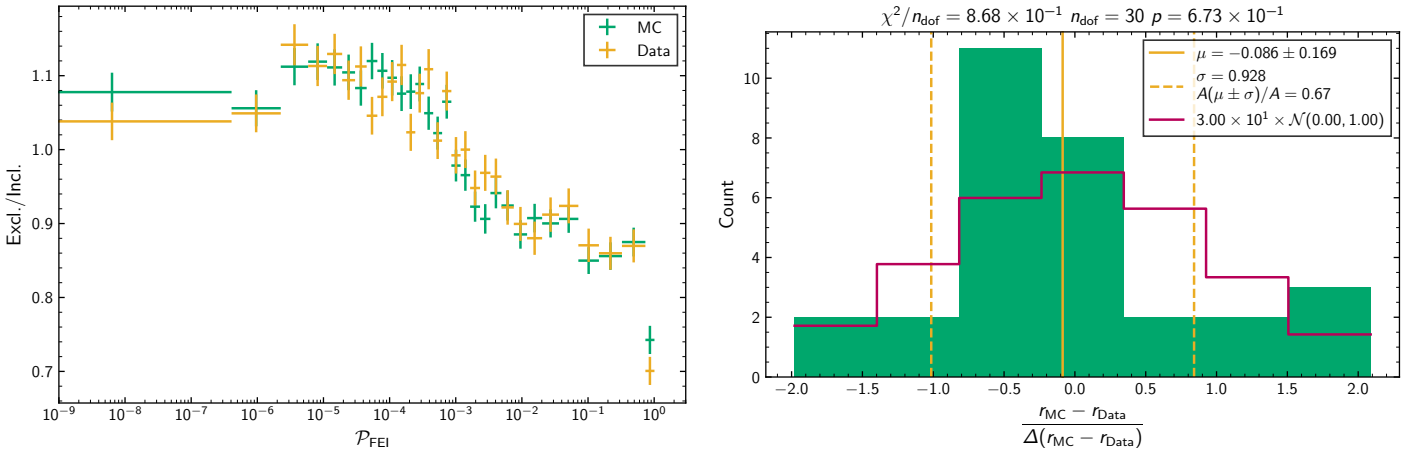
For N_{cat} categories let us first define

$$\widetilde{\mathcal{M}}_{ij} := \#\{E \in \Omega_{\text{ovl}} \mid f^{\text{excl}}(E) = C_i \ \wedge f^{\text{incl}}(E) = C_j\}, \tag{D.6.1}$$

where Ω_{ovl} is the overlap dataset. We can recover the marginal distributions as

$$\begin{aligned} \sum_{i=0}^{N_{\text{cat}}-1} \widetilde{\mathcal{M}}_{ij} &= N^{\text{ovl}}_{\text{incl}\,j}, \\ \sum_{j=0}^{N_{\text{cat}}-1} \widetilde{\mathcal{M}}_{ij} &= N^{\text{ovl}}_{\text{excl}\,i}. \end{aligned} \tag{D.6.2}$$

With this definition of "migration", we can illustrate remark i from page 182: If $\widetilde{\mathcal{M}} = \widetilde{\mathcal{M}}^\intercal$, then by Equation (D.6.2), $N^{\text{ovl}}_{\text{incl}\,i} = N^{\text{ovl}}_{\text{excl}\,i}$, no matter the size of the elements of $\widetilde{\mathcal{M}}$. However, it is



(a) r^X

(b) Pulls between r^{Data} and r^{MC}.

Figure D.16: r^X and pulls for categorization in $\mathcal{P}_{\text{FEI}}$ only.

clear that $\widetilde{\mathcal{M}}$ is not a suitable measure for the relative amount of migration that occurs: We need a normalization to become independent of the sample sizes. In the following, we normalize $\widetilde{\mathcal{M}}$ across columns, that is $\mathcal{M}_{ij} := \widetilde{\mathcal{M}}_{ij}/N_j^{\text{incl}}$. The resulting migration matrix has the property that

$$\sum_{j=0}^{N_{\text{cat}}-1} \mathcal{M}_{ij} N^{\text{ovl}}_{\text{incl}\,j} = N^{\text{ovl}}_{\text{excl}\,i}. \tag{D.6.3}$$

If $N^{\text{ovl}}_{\text{incl}\,i} \approx \text{const.}$, then the symmetry $\mathcal{M} = \mathcal{M}^\intercal$ still implies that $N^{\text{ovl}}_{\text{incl}\,i} \approx N^{\text{ovl}}_{\text{excl}\,i}$, but generally, symmetry of $\mathcal{M}$ no longer implies that the migration effects cancel in the counts.

To quantify the impact of the migration on the distributions, we introduce the ratio

$$r_i^X := \frac{N^{\text{ovl}\,X}_{\text{excl}\,i}}{N^{\text{ovl}\,X}_{\text{incl}\,i}}, \tag{D.6.4}$$

which will be referred to as *effective* migration. Note that

$$r_i^{\text{MC}} = r_i^{\text{Data}} \quad \Longleftrightarrow \quad \epsilon^{\text{ovl}}_{\text{excl}\,i} = \epsilon^{\text{ovl}}_{\text{incl}\,i}, \tag{D.6.5}$$

where "=" means statistical compatibility.

D.6.3 Migration in the FEI classifier

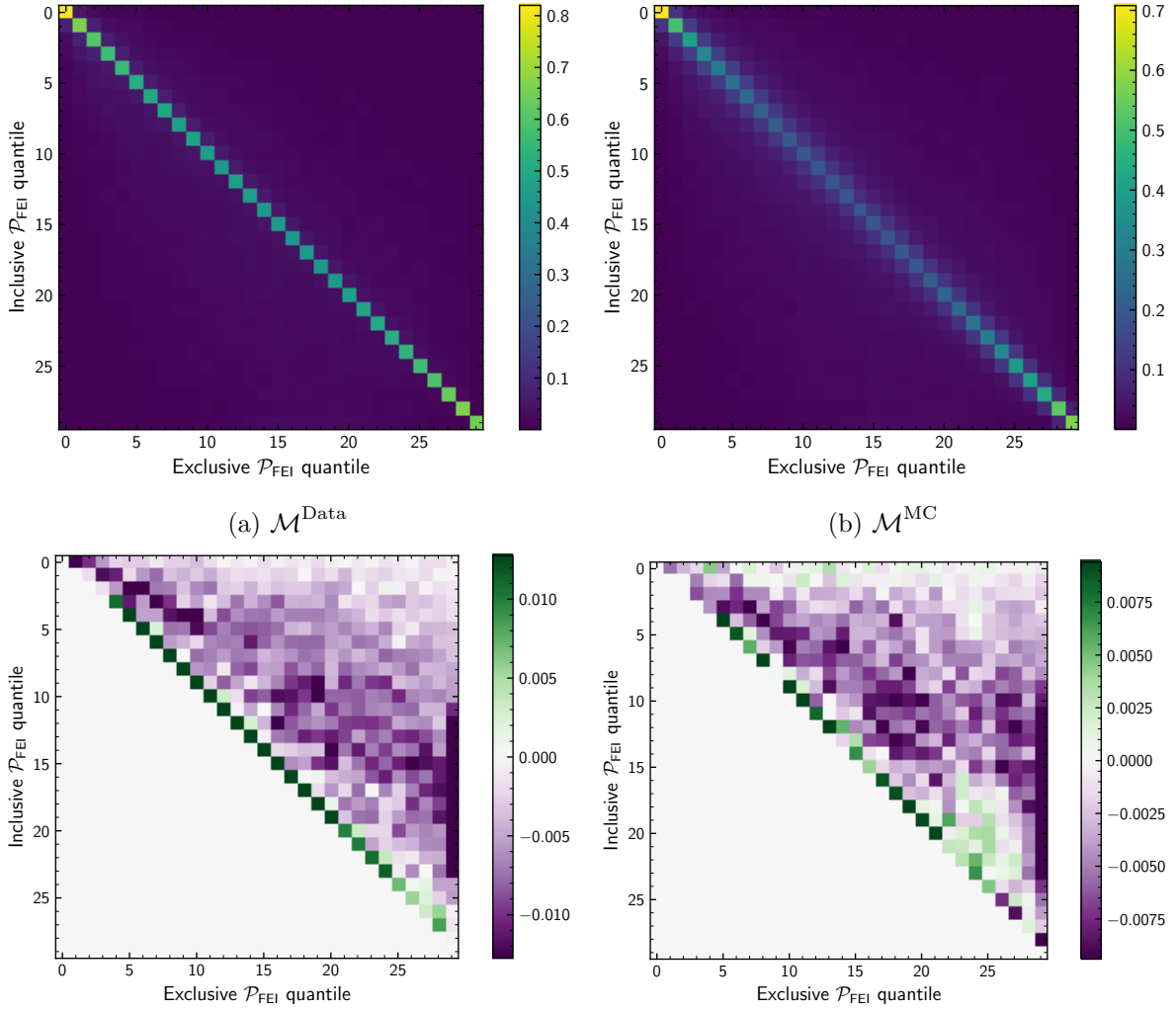
This section extends the information provided in Appendix C.3.2.

While the effective migration r_i^X (for the most part) remains within 10 % around unity (Figure D.16a), the corresponding migration matrix has diagonal elements that drop to $\approx 50\,\%$ (data) resp. $\approx 20\,\%$ (MC) (see Figure D.18a).

The migration matrices and their asymmetries are shown in Figure D.17. It is clearly visible that significantly more migration occurs in MC than in data. Both matrices are relatively symmetric with the absolute asymmetries $\mathcal{M} - \mathcal{M}^\intercal$ being smaller than 1.3 percent points (0.79 percent points) in data (MC) (for the next-to-diagonal elements, see also Figure D.18b).

D.6.4 Migration in the FEI reconstruction channel

This section extends the information provided in Appendix C.3.3.



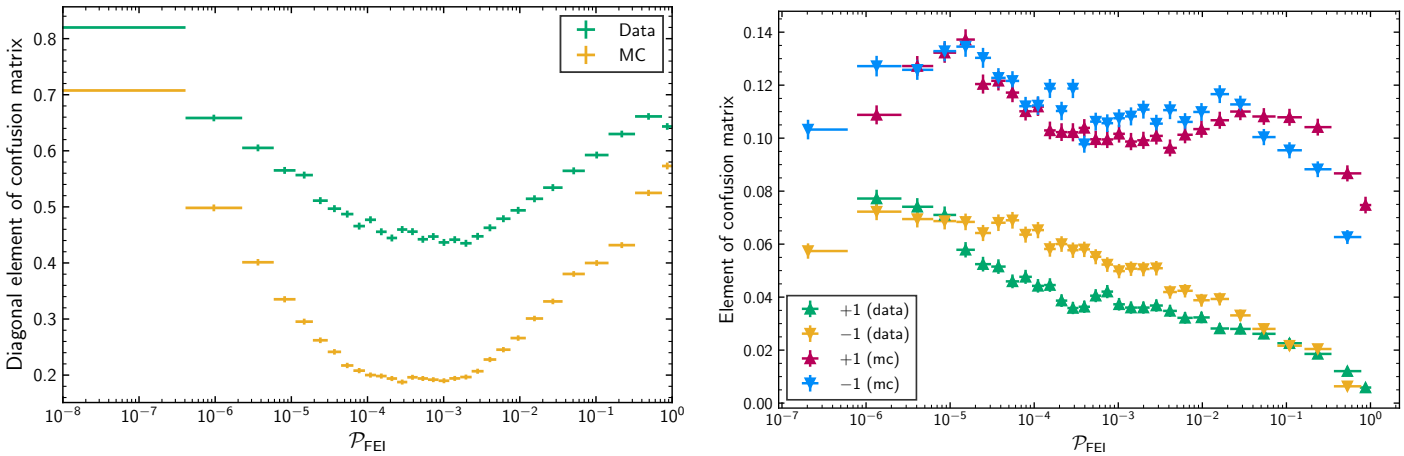
(a) $\mathcal{M}^{\text{Data}}$

(b) $\mathcal{M}^{\text{MC}}$

(c) $\mathcal{M}^{\text{Data}} - (\mathcal{M}^{\text{Data}})^{\intercal}$

(d) $\mathcal{M}^{\text{MC}} - (\mathcal{M}^{\text{MC}})^{\intercal}$

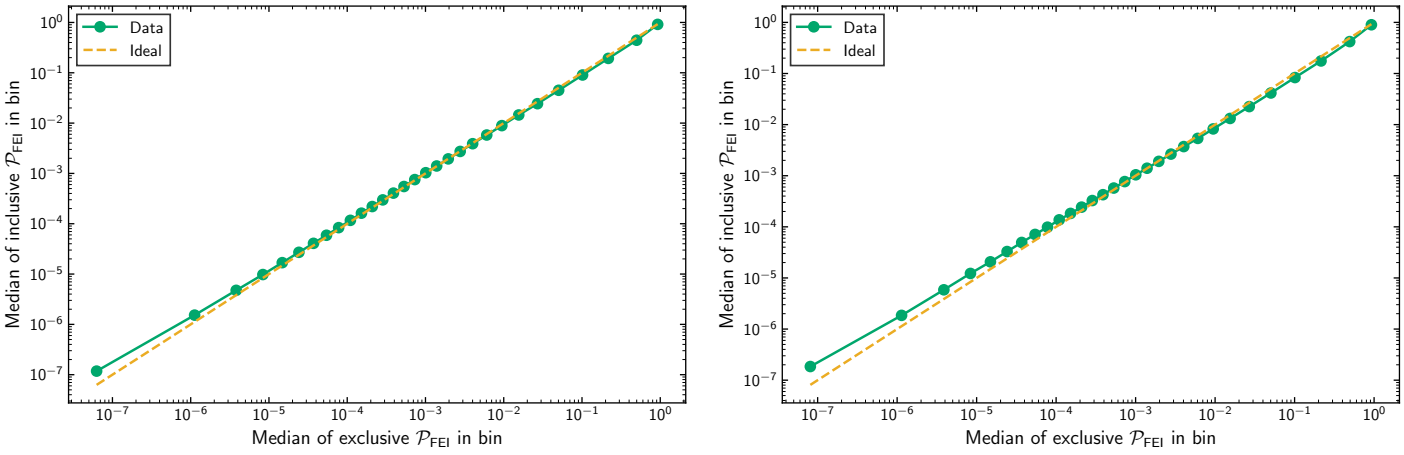
Figure D.17: Migration matrices and their asymmetries for categorization in $\mathcal{P}_{\text{FEI}}$. The lower triangular matrix of the asymmetries $\mathcal{M} - \mathcal{M}^{\intercal}$ is left out because it provides no additional information.



(a) Diagonal elements $\operatorname{diag} \mathcal{M} = \mathcal{M}_{ii}$

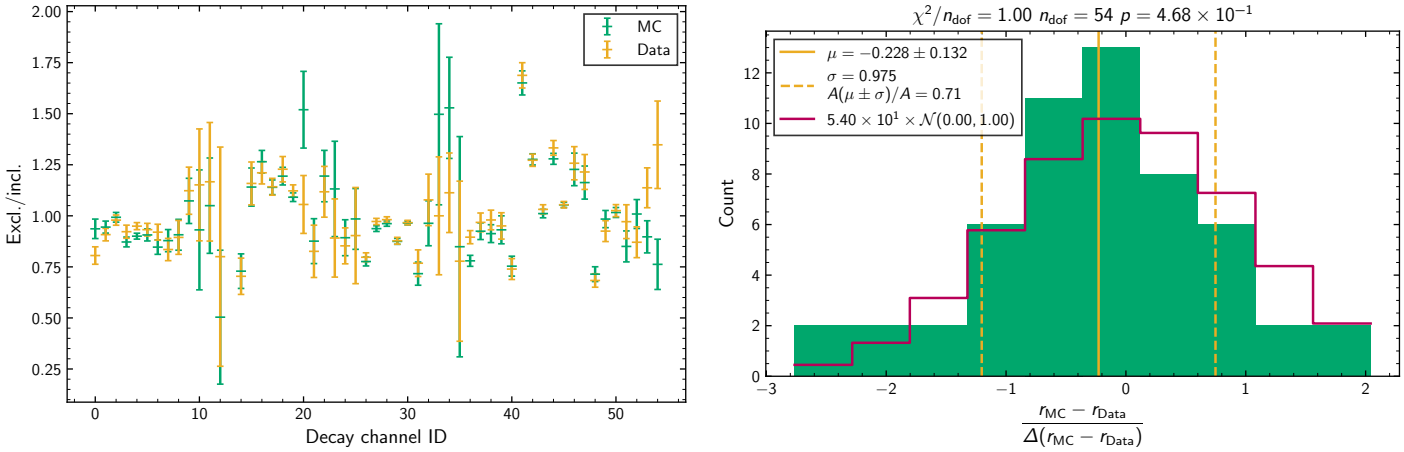
(b) Off-diagonal elements $\mathcal{M}_{i\,i\pm1}$

Figure D.18: Diagonal and off-diagonal elements of the migration matrix for categorization in $\mathcal{P}_{\text{FEI}}$.



(a) Data (b) MC

Figure D.19: Effects of migration on medians of $\mathcal{P}_{\text{FEI}}$: The events in the overlap are binned in $\mathcal{P}_{\text{FEI}}^{\text{incl}}$. For each bin we compare the median of $\mathcal{P}_{\text{FEI}}^{\text{incl}}$ with $\mathcal{P}_{\text{FEI}}^{\text{excl}}$.



(a) r^{Data} and r^{MC}. (b) Pulls between r^{Data} and r^{MC}.

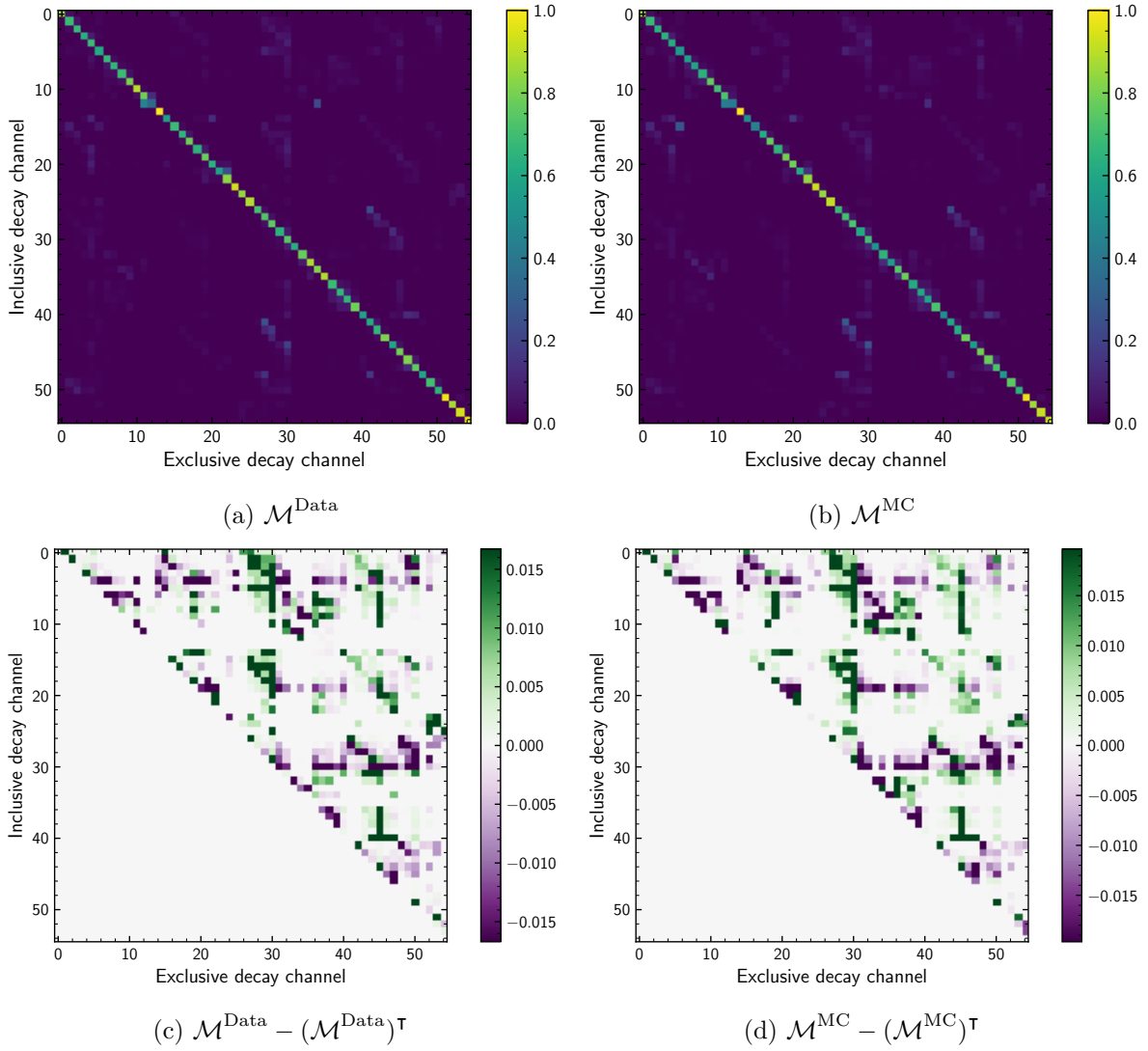
Figure D.20: r^X and pull distributions for categorization in C_{FEI} only.

While the effective migration r_i^X (for the most part) remains within 25 % around unity (Figure D.20a), the corresponding migration matrix has diagonal elements that drop to around 60 %, with MC showing slightly more migration than data (Figure D.22).

A table of the mapping of FEI decay channel ID to reconstruction mode is shown in Table 9.1. The migration matrices and their asymmetries are shown in Figure D.21, diagonal elements of the migration matrices in Figure D.22. Table D.4 lists the ten most frequent inclusive decay channels in the overlap data and the decay channels they most often migrate to.

D.6.5 Migration in the FEI classifier and reconstruction channel

This section extends the information provided in Appendix C.3.4. The diagonal elements of the migration matrix is shown in Figure D.24.



(a) $\mathcal{M}^{\mathrm{Data}}$

(b) $\mathcal{M}^{\mathrm{MC}}$

(c) $\mathcal{M}^{\mathrm{Data}} - (\mathcal{M}^{\mathrm{Data}})^{\intercal}$

(d) $\mathcal{M}^{\mathrm{MC}} - (\mathcal{M}^{\mathrm{MC}})^{\intercal}$

Figure D.21: Migration matrices and their asymmetries for categorization in C_{FEI}. The lower triangular matrix of the asymmetries $\mathcal{M} - \mathcal{M}^{\intercal}$ is left out because it provides no additional information.

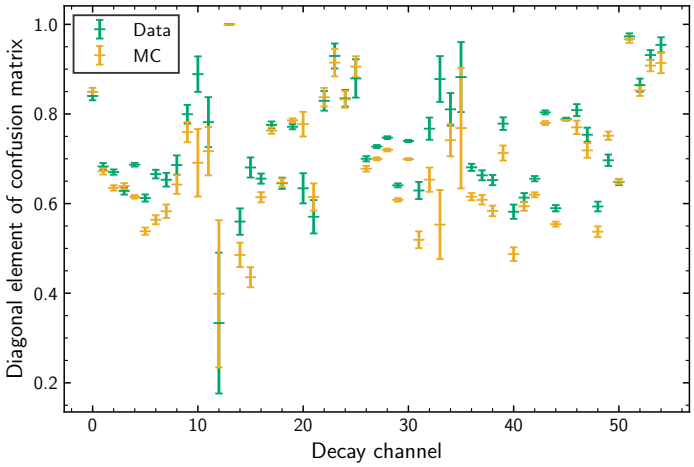


Figure D.22: Diagonal elements of the migration matrix of categorization in C_{FEI}.

Incl. Dch	Identical	Excl. Dch		Excl. Dch		Rest
$\bar{\mathrm{D}}^0\pi^+\pi^+\pi^-\pi^0$	71.9	$\bar{\mathrm{D}}^{*0}\pi^+\pi^+\pi^-\pi^0$	7.7	$\bar{\mathrm{D}}^{*0}\pi^+\pi^+\pi^-$	4.3	16.1
$\bar{\mathrm{D}}^0\pi^+\pi^0\pi^0$	73.3	$\bar{\mathrm{D}}^{*0}\pi^+\pi^0$	7.5	$\bar{\mathrm{D}}^{*0}\pi^+\pi^0\pi^0$	5.3	13.9
$\bar{\mathrm{D}}^{*0}\pi^+\pi^+\pi^-\pi^0$	78.7	$\bar{\mathrm{D}}^0\pi^+\pi^+\pi^-\pi^0$	9.3	$\mathrm{D}^{*-}\pi^+\pi^+\pi^-\pi^0$	2.9	9.1
$\bar{\mathrm{D}}^0\pi^+\pi^0$	71.3	$\bar{\mathrm{D}}^{*0}\pi^+\pi^0$	10.8	$\bar{\mathrm{D}}^0\pi^+\pi^0\pi^0$	4.6	13.2
$\mathrm{D}^-\pi^+\pi^+\pi^-\pi^0$	65.2	$\bar{\mathrm{D}}^0\pi^+\pi^+\pi^-\pi^0$	9.4	$\bar{\mathrm{D}}^{*0}\pi^+\pi^+\pi^-\pi^0$	5.2	20.1
$\bar{\mathrm{D}}^0\pi^+\pi^+\pi^-$	62.4	$\bar{\mathrm{D}}^{*0}\pi^+\pi^+\pi^-$	12.6	$\bar{\mathrm{D}}^0\pi^+\pi^+\pi^-\pi^0$	7.9	17.0
$\bar{\mathrm{D}}^{*0}\pi^+\pi^0\pi^0$	79.1	$\bar{\mathrm{D}}^0\pi^+\pi^0\pi^0$	10.5	$\bar{\mathrm{D}}^{*0}\pi^+\pi^0$	2.7	7.7
$\bar{\mathrm{D}}^{*0}\pi^+\pi^0$	63.7	$\bar{\mathrm{D}}^0\pi^+\pi^0\pi^0$	16.9	$\bar{\mathrm{D}}^0\pi^+\pi^0$	10.0	9.4
$\bar{\mathrm{D}}^{*0}\pi^+\pi^+\pi^-$	57.1	$\bar{\mathrm{D}}^0\pi^+\pi^+\pi^-\pi^0$	22.3	$\bar{\mathrm{D}}^0\pi^+\pi^+\pi^-$	5.0	15.6
$\mathrm{D}^{*-}\pi^+\pi^+\pi^-\pi^0$	77.9	$\bar{\mathrm{D}}^{*0}\pi^+\pi^+\pi^-\pi^0$	5.1	$\bar{\mathrm{D}}^0\pi^+\pi^+\pi^-\pi^0$	5.0	12.0

Table D.4: Elements of the migration matrix for the most frequent decay channels. All numbers are shown in percent. Explanation of columns: *incl. dch* (inclusive FEI reconstruction mode), *identical* (diagonal element of the migration matrix), *excl. dch* (exclusive FEI reconstruction mode), *rest* (remaining percentage that is not on the diagonal and not included in the two largest off-diagonal elements).

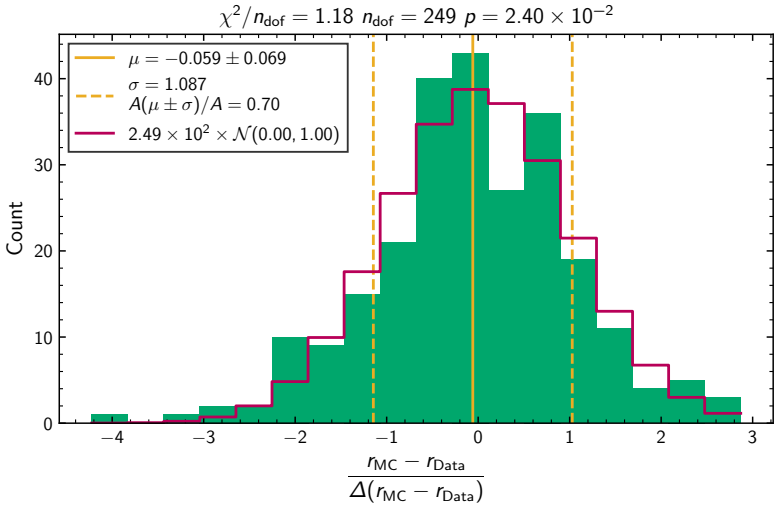


Figure D.23: r^X and pulls for categorization in $(C_{\mathrm{FEI}}, \mathcal{P}_{\mathrm{FEI}})$ categories.

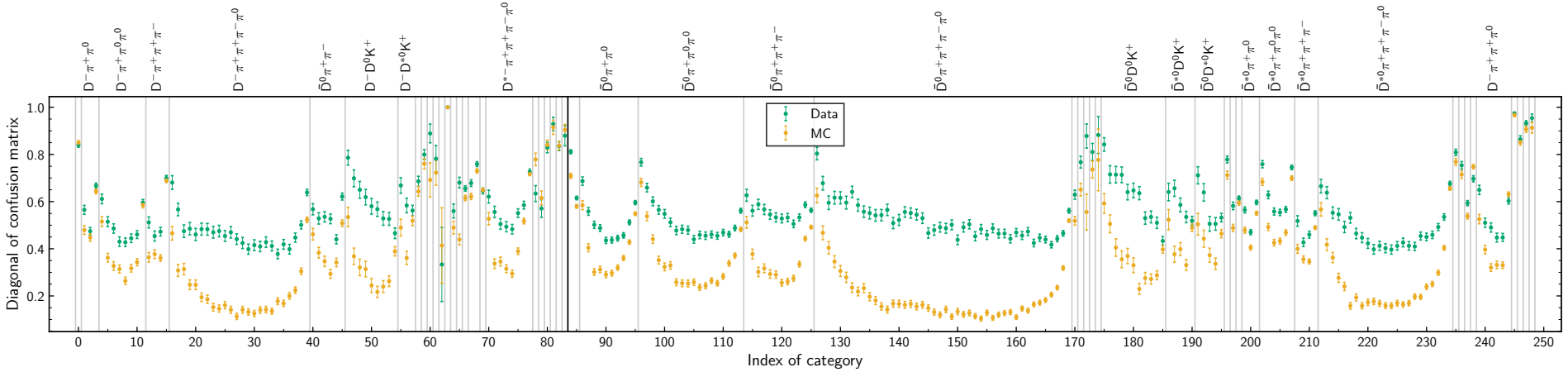


Figure D.24: Diagonal elements of the migration matrix diag $\mathcal{M} = \mathcal{M}_{ii}$ for migration in $(C_{\mathrm{FEI}}, \mathcal{P}_{\mathrm{FEI}})$.

Appendix E

Calibration Results

Below are the numerical results that were described in Chapter 10.

Notes about the interpretation of the columns:

- i is the identifier of the category
- The columns L_1 and L_2 denote the range of the bin in the FEI classifier output $\mathcal{P}_{\text{FEI}}$: $L_1 < \log_{10} \mathcal{P}_{\text{FEI}} < L_2$. If $L_1 = -\infty$, no lower cut on $\mathcal{P}_{\text{FEI}}$ was applied; if $L_2 = 0$, no upper cut on $\mathcal{P}_{\text{FEI}}$ was applied ($0 \leq \mathcal{P}_{\text{FEI}} \leq 1$).
- *CF* denotes the calibration factor $\epsilon_{\text{sig},i}$ which is to be applied to the exclusive dataset in the $|V_{\text{cb}}|$ measurement
- *Unc.*: Uncertainty that is uncorrelated between categories (for example statistical uncertainty).
- *Corr.*: Uncertainty that is assumed to 100 % correlated between categories (for example, branching ratio uncertainties that have no effect on the fit template shapes and enter the calibration factors via the denominator). See Section 9.5 for more information.
- *S*: Strategy with which the calibration factors were extracted:

  P ϵ_i^{sig} was calculated from the primary fit (see Section 9.2.6)

  F ϵ_i^{sig} was calculated from the fallback fit setup as described in Section 9.3.2.

  A ϵ_i^{sig} was calculated with cut-and-counting and not a fit (*average* strategy). See Section 9.3.3.

  Because the calibration factor is calculated as the average of the calibration factors derived for the two different gap models, we always have two fit setups. The first letter denotes the fit setup used with the resonant gap, the second the fit setup used with the non-resonant gap.
- p_1 and p_2: p-values of the goodness of fit test performed for the two fits (see Section 9.2.2)
- *Notes*: Numbers in this column point to notes below this table.

i	Channel	L_1	L_2	CF	Stat	Syst	S	p_1	p_2	Notes
0	$\text{D}^-\pi^+$	$-\infty$	-0.59	1.219	0.045	0.052	PP	0.03	0.01	
1	$\text{D}^-\pi^+$	-0.59	-0.04	1.127	0.035	0.049	PP	0.63	0.40	

Continued on next page

i	Channel	L_1	L_2	CF	Stat	Syst	S	p_1	p_2	Notes
2	$D^-\pi^+$	-0.04	0.00	1.085	0.030	0.050	PP	0.82	0.79	
3	$D^-\pi^+\pi^0$	$-\infty$	-3.34	1.171	0.024	0.052	PP	0.32	0.26	
4	$D^-\pi^+\pi^0$	-3.34	-1.98	1.081	0.021	0.045	PP	0.03	0.02	1
5	$D^-\pi^+\pi^0$	-1.98	0.00	0.972	0.017	0.040	PP	0.14	0.06	
6	$D^-\pi^+\pi^0\pi^0$	$-\infty$	-5.57	1.070	0.018	0.049	PP	0.94	0.89	
7	$D^-\pi^+\pi^0\pi^0$	-5.57	-4.81	1.053	0.018	0.046	PP	0.10	0.07	
8	$D^-\pi^+\pi^0\pi^0$	-4.81	-4.28	1.018	0.018	0.041	PP	0.32	0.18	2
9	$D^-\pi^+\pi^0\pi^0$	-4.28	-3.74	1.044	0.018	0.046	PP	0.39	0.46	
10	$D^-\pi^+\pi^0\pi^0$	-3.74	-3.05	1.012	0.017	0.042	PP	0.22	0.12	3
11	$D^-\pi^+\pi^0\pi^0$	-3.05	0.00	1.028	0.016	0.043	PP	0.74	0.71	
12	$D^-\pi^+\pi^+\pi^-$	$-\infty$	-3.74	1.109	0.018	0.053	PP	0.27	0.42	
13	$D^-\pi^+\pi^+\pi^-$	-3.74	-2.63	0.889	0.015	0.038	PP	0.24	0.15	
14	$D^-\pi^+\pi^+\pi^-$	-2.63	0.00	0.711	0.011	0.029	PP	0.37	0.34	
15	$D^-(\to K^-\pi^-\pi^-)\pi^+\pi^+\pi^-\pi^0$	$-\infty$	-5.74	1.118	0.017	0.048	PP	0.97	0.96	
16	$D^-(\to K^-\pi^-\pi^-)\pi^+\pi^+\pi^-\pi^0$	-5.74	-5.04	1.008	0.017	0.040	PP	0.38	0.40	
17	$D^-(\to K^-\pi^-\pi^-)\pi^+\pi^+\pi^-\pi^0$	-5.04	-4.52	0.925	0.015	0.037	PP	0.87	0.86	
18	$D^-(\to K^-\pi^-\pi^-)\pi^+\pi^+\pi^-\pi^0$	-4.52	-3.99	0.946	0.015	0.039	PP	0.89	0.86	
19	$D^-(\to K^-\pi^-\pi^-)\pi^+\pi^+\pi^-\pi^0$	-3.99	-3.23	0.906	0.014	0.040	PP	0.16	0.20	
20	$D^-(\to K^-\pi^-\pi^-)\pi^+\pi^+\pi^-\pi^0$	-3.23	0.00	0.891	0.014	0.038	PP	0.16	0.25	
21	$D^-(\to$ Other$)\pi^+\pi^+\pi^-\pi^0$	$-\infty$	-6.44	1.060	0.017	0.051	PP	0.45	0.37	
22	$D^-(\to$ Other$)\pi^+\pi^+\pi^-\pi^0$	-6.44	-5.86	1.081	0.018	0.047	PP	0.99	0.99	
23	$D^-(\to$ Other$)\pi^+\pi^+\pi^-\pi^0$	-5.86	-5.48	1.047	0.017	0.044	PP	0.60	0.58	
24	$D^-(\to$ Other$)\pi^+\pi^+\pi^-\pi^0$	-5.48	-5.19	0.997	0.017	0.041	PP	0.05	0.02	
25	$D^-(\to$ Other$)\pi^+\pi^+\pi^-\pi^0$	-5.19	-4.95	0.993	0.017	0.043	PP	0.82	0.85	
26	$D^-(\to$ Other$)\pi^+\pi^+\pi^-\pi^0$	-4.95	-4.72	0.981	0.016	0.042	PP	0.38	0.44	
27	$D^-(\to$ Other$)\pi^+\pi^+\pi^-\pi^0$	-4.72	-4.50	0.946	0.016	0.040	PP	0.75	0.58	
28	$D^-(\to$ Other$)\pi^+\pi^+\pi^-\pi^0$	-4.50	-4.27	0.948	0.016	0.040	PP	0.12	0.17	
29	$D^-(\to$ Other$)\pi^+\pi^+\pi^-\pi^0$	-4.27	-4.00	0.907	0.015	0.039	PP	0.07	0.09	
30	$D^-(\to$ Other$)\pi^+\pi^+\pi^-\pi^0$	-4.00	-3.67	0.913	0.015	0.040	PP	0.06	0.12	
31	$D^-(\to$ Other$)\pi^+\pi^+\pi^-\pi^0$	-3.67	-3.18	0.942	0.015	0.041	PP	0.77	0.61	
32	$D^-(\to$ Other$)\pi^+\pi^+\pi^-\pi^0$	-3.18	0.00	0.887	0.014	0.037	PP	0.31	0.34	
33	$\bar{D}^0\pi^+\pi^-$	$-\infty$	-4.55	1.195	0.022	0.054	PP	0.81	0.79	
34	$\bar{D}^0\pi^+\pi^-$	-4.55	-3.78	1.286	0.025	0.051	PP	0.02	0.02	4
35	$\bar{D}^0\pi^+\pi^-$	-3.78	-3.00	1.217	0.023	0.051	PP	0.40	0.52	
36	$\bar{D}^0\pi^+\pi^-$	-3.00	0.00	1.072	0.019	0.045	PP	0.70	0.61	
37	$D^-D^0K^+$	$-\infty$	-8.20	1.075	0.019	0.044	FF	0.15	0.18	5, 6
38	$D^-D^0K^+$	-8.20	-7.46	1.079	0.016	0.048	FP	0.39	6×10^{-3}	7
39	$D^-D^0K^+$	-7.46	-6.82	1.125	0.019	0.045	PP	0.74	0.69	8
40	$D^-D^0K^+$	-6.82	-6.20	1.141	0.017	0.047	PP	0.40	0.13	
41	$D^-D^0K^+$	-6.20	-5.54	1.130	0.017	0.048	PP	0.85	0.87	
42	$D^-D^0K^+$	-5.54	-4.69	1.096	0.017	0.049	PP	0.61	0.75	
43	$D^-D^0K^+$	-4.69	0.00	1.065	0.015	0.046	PP	0.89	0.87	
44	$D^-D^{*0}K^+$	$-\infty$	-7.35	1.104	0.019	0.050	FF	0.15	0.28	9, 10
45	$D^-D^{*0}K^+$	-7.35	-5.81	1.123	0.019	0.050	FF	0.27	0.30	11, 12
46	$D^-D^{*0}K^+$	-5.81	0.00	1.138	0.019	0.059	FP	0.66	0.92	13
47	$D^{*-}D^0K^+$	$-\infty$	-7.88	1.077	0.024	0.046	PP	0.95	0.79	
48	$D^{*-}D^0K^+$	-7.88	-6.16	1.075	0.024	0.047	FF	0.27	0.22	14, 15

Continued on next page

i	Channel	L_1	L_2	CF	Stat	Syst	S	p_1	p_2	Notes
49	$D^{*-}D^0K^+$	-6.16	0.00	1.014	0.019	0.042	PP	0.61	0.66	
50	$D^{*-}D^{*0}K^+$	$-\infty$	-6.92	1.060	0.029	0.048	PP	0.78	0.81	
51	$D^{*-}D^{*0}K^+$	-6.92	-5.02	1.081	0.029	0.050	PP	0.50	0.50	
52	$D^{*-}D^{*0}K^+$	-5.02	0.00	1.046	0.031	0.045	PP	0.08	0.11	
53	$D^-D^+K^0_S$	$-\infty$	0.00	1.230	0.036	0.055	FF	0.09	0.15	16, 17
54	$D^{*-}D^+K^0_S$	$-\infty$	0.00	1.087	0.037	0.052	PP	0.79	0.81	
55	$D^-D^{*+}K^0_S$	$-\infty$	0.00	1.104	0.074	0.053	PP	0.89	0.87	
56	$D^{*-}D^{*+}K^0_S$	$-\infty$	0.00	1.070	0.087	0.048	AA	NaN	NaN	18, 19, 20, 21
57	$D_s^+D^-$	$-\infty$	-3.06	1.153	0.047	0.058	PP	0.48	0.50	
58	$D_s^+D^-$	-3.06	0.00	1.029	0.037	0.047	PP	0.26	0.22	
59	$D^{*-}\pi^+$	$-\infty$	0.00	1.054	0.034	0.045	PP	0.18	0.18	
60	$D^{*-}\pi^+\pi^0$	$-\infty$	-2.89	1.084	0.030	0.051	PP	0.84	0.91	
61	$D^{*-}\pi^+\pi^0$	-2.89	-1.25	1.068	0.030	0.046	FP	0.70	0.20	22
62	$D^{*-}\pi^+\pi^0$	-1.25	0.00	1.125	0.031	0.045	FF	0.12	0.07	23, 24
63	$D^{*-}\pi^+\pi^0\pi^0$	$-\infty$	-4.28	0.915	0.016	0.044	PP	0.75	0.77	
64	$D^{*-}\pi^+\pi^0\pi^0$	-4.28	-2.97	0.906	0.016	0.041	PP	0.15	0.22	
65	$D^{*-}\pi^+\pi^0\pi^0$	-2.97	0.00	0.850	0.016	0.035	FP	0.87	0.74	25
66	$D^{*-}\pi^+\pi^+\pi^-$	$-\infty$	-2.92	0.895	0.024	0.046	PP	0.13	0.08	
67	$D^{*-}\pi^+\pi^+\pi^-$	-2.92	-1.05	0.665	0.015	0.028	PP	0.59	0.69	
68	$D^{*-}\pi^+\pi^+\pi^-$	-1.05	0.00	0.547	0.015	0.023	FF	0.42	0.37	26, 27
69	$D^{*-}\pi^+\pi^+\pi^-\pi^0$	$-\infty$	-5.00	1.028	0.017	0.053	FF	0.46	0.46	28, 29, 30
70	$D^{*-}\pi^+\pi^+\pi^-\pi^0$	-5.00	-4.35	0.929	0.013	0.042	PP	0.06	0.12	
71	$D^{*-}\pi^+\pi^+\pi^-\pi^0$	-4.35	-3.83	0.879	0.012	0.039	PP	0.40	0.41	
72	$D^{*-}\pi^+\pi^+\pi^-\pi^0$	-3.83	-3.22	0.805	0.011	0.036	PP	0.03	0.07	
73	$D^{*-}\pi^+\pi^+\pi^-\pi^0$	-3.22	-2.33	0.717	0.010	0.030	PP	0.61	0.74	
74	$D^{*-}\pi^+\pi^+\pi^-\pi^0$	-2.33	0.00	0.537	0.008	0.021	PP	0.74	0.77	
75	$D_s^{*+}D^-$	$-\infty$	0.00	0.966	0.034	0.049	PP	7×10^{-3}	0.01	
76	$D_s^+D^{*-}$	$-\infty$	0.00	0.830	0.031	0.034	PP	0.90	0.69	31
77	$D_s^{*+}D^{*-}$	$-\infty$	0.00	0.941	0.031	0.044	PP	0.05	0.04	32
78	$J/\psi K^0_S$	$-\infty$	0.00	1.231	0.089	0.056	FF	0.48	0.52	33, 34
79	$J/\psi K^+\pi^-$	$-\infty$	0.00	1.145	0.041	0.047	PP	0.22	0.20	
80	$J/\psi K^0_S\pi^+\pi^-$	$-\infty$	0.00	0.609	0.041	0.026	FP	0.95	0.07	35
81	$\bar{D}^0\pi^+$	$-\infty$	-0.80	1.186	0.028	0.052	PP	0.18	0.07	
82	$\bar{D}^0\pi^+$	-0.80	-0.05	1.050	0.023	0.045	PP	0.20	0.16	
83	$\bar{D}^0\pi^+$	-0.05	0.00	0.986	0.023	0.042	PP	0.22	0.30	36, 37
84	$\bar{D}^0\pi^+\pi^0$	$-\infty$	-4.17	1.127	0.022	0.048	PP	0.29	0.22	
85	$\bar{D}^0\pi^+\pi^0$	-4.17	-3.63	1.134	0.023	0.049	PP	0.11	0.23	
86	$\bar{D}^0\pi^+\pi^0$	-3.63	-3.21	1.138	0.021	0.048	PP	0.60	0.71	
87	$\bar{D}^0\pi^+\pi^0$	-3.21	-2.79	1.073	0.020	0.044	PP	0.33	0.16	
88	$\bar{D}^0\pi^+\pi^0$	-2.79	-2.32	1.081	0.019	0.045	PP	0.39	0.58	
89	$\bar{D}^0\pi^+\pi^0$	-2.32	-1.76	1.050	0.018	0.044	PP	6×10^{-3}	0.01	
90	$\bar{D}^0\pi^+\pi^0$	-1.76	-1.02	0.884	0.016	0.036	PP	0.21	0.20	
91	$\bar{D}^0\pi^+\pi^0$	-1.02	0.00	0.787	0.013	0.031	PP	8×10^{-3}	0.02	
92	$\bar{D}^0\pi^+\pi^0\pi^0$	$-\infty$	-6.10	1.017	0.018	0.047	PP	0.73	0.83	38
93	$\bar{D}^0\pi^+\pi^0\pi^0$	-6.10	-5.51	1.047	0.018	0.045	PP	0.46	0.46	
94	$\bar{D}^0\pi^+\pi^0\pi^0$	-5.51	-5.13	1.066	0.018	0.044	PP	0.02	0.05	
95	$\bar{D}^0\pi^+\pi^0\pi^0$	-5.13	-4.84	1.039	0.018	0.040	PP	0.10	0.13	39

Continued on next page

i	Channel	L_1	L_2	CF	Stat	Syst	S	p_1	p_2	Notes
96	$\bar{\mathrm{D}}^0\pi^+\pi^0\pi^0$	-4.84	-4.59	0.991	0.017	0.040	PP	0.34	0.40	
97	$\bar{\mathrm{D}}^0\pi^+\pi^0\pi^0$	-4.59	-4.36	0.996	0.018	0.041	PP	0.69	0.64	
98	$\bar{\mathrm{D}}^0\pi^+\pi^0\pi^0$	-4.36	-4.14	1.034	0.018	0.042	PP	0.17	0.22	
99	$\bar{\mathrm{D}}^0\pi^+\pi^0\pi^0$	-4.14	-3.92	1.012	0.017	0.043	PP	0.17	0.29	
100	$\bar{\mathrm{D}}^0\pi^+\pi^0\pi^0$	-3.92	-3.69	0.998	0.017	0.042	PP	0.69	0.60	
101	$\bar{\mathrm{D}}^0\pi^+\pi^0\pi^0$	-3.69	-3.44	0.999	0.016	0.041	PP	0.08	0.10	
102	$\bar{\mathrm{D}}^0\pi^+\pi^0\pi^0$	-3.44	-3.15	1.019	0.016	0.043	PP	0.40	0.42	
103	$\bar{\mathrm{D}}^0\pi^+\pi^0\pi^0$	-3.15	-2.80	0.975	0.016	0.040	PP	0.83	0.79	
104	$\bar{\mathrm{D}}^0\pi^+\pi^0\pi^0$	-2.80	-2.32	0.941	0.016	0.039	PP	0.86	0.95	40
105	$\bar{\mathrm{D}}^0\pi^+\pi^0\pi^0$	-2.32	0.00	0.939	0.015	0.036	PP	0.66	0.69	
106	$\bar{\mathrm{D}}^0\pi^+\pi^+\pi^-$	$-\infty$	-4.74	1.105	0.017	0.056	PP	0.40	0.44	
107	$\bar{\mathrm{D}}^0\pi^+\pi^+\pi^-$	-4.74	-4.15	1.080	0.018	0.051	PP	0.26	0.43	
108	$\bar{\mathrm{D}}^0\pi^+\pi^+\pi^-$	-4.15	-3.76	1.020	0.018	0.043	PP	0.67	0.60	
109	$\bar{\mathrm{D}}^0\pi^+\pi^+\pi^-$	-3.76	-3.44	0.961	0.016	0.041	PP	0.67	0.61	
110	$\bar{\mathrm{D}}^0\pi^+\pi^+\pi^-$	-3.44	-3.12	0.907	0.016	0.038	PP	0.89	0.84	
111	$\bar{\mathrm{D}}^0\pi^+\pi^+\pi^-$	-3.12	-2.77	0.888	0.015	0.036	PP	0.51	0.51	
112	$\bar{\mathrm{D}}^0\pi^+\pi^+\pi^-$	-2.77	-2.33	0.864	0.014	0.036	PP	0.30	0.14	
113	$\bar{\mathrm{D}}^0\pi^+\pi^+\pi^-$	-2.33	-1.58	0.783	0.014	0.030	PP	0.85	0.93	41
114	$\bar{\mathrm{D}}^0\pi^+\pi^+\pi^-$	-1.58	0.00	0.780	0.013	0.029	PP	0.04	0.05	42
115	$\bar{\mathrm{D}}^0(\to \mathrm{K}^-\pi^+\pi^0)\pi^+\pi^+\pi^-\pi^0$	$-\infty$	-5.79	1.016	0.016	0.046	PP	0.57	0.64	
116	$\bar{\mathrm{D}}^0(\to \mathrm{K}^-\pi^+\pi^0)\pi^+\pi^+\pi^-\pi^0$	-5.79	-5.23	0.933	0.015	0.037	PP	0.58	0.57	
117	$\bar{\mathrm{D}}^0(\to \mathrm{K}^-\pi^+\pi^0)\pi^+\pi^+\pi^-\pi^0$	-5.23	-4.86	0.908	0.015	0.035	PP	0.56	0.68	
118	$\bar{\mathrm{D}}^0(\to \mathrm{K}^-\pi^+\pi^0)\pi^+\pi^+\pi^-\pi^0$	-4.86	-4.54	0.880	0.015	0.035	PP	0.43	0.54	
119	$\bar{\mathrm{D}}^0(\to \mathrm{K}^-\pi^+\pi^0)\pi^+\pi^+\pi^-\pi^0$	-4.54	-4.23	0.875	0.014	0.034	PP	0.78	0.73	
120	$\bar{\mathrm{D}}^0(\to \mathrm{K}^-\pi^+\pi^0)\pi^+\pi^+\pi^-\pi^0$	-4.23	-3.90	0.851	0.014	0.034	PP	0.93	0.96	
121	$\bar{\mathrm{D}}^0(\to \mathrm{K}^-\pi^+\pi^0)\pi^+\pi^+\pi^-\pi^0$	-3.90	-3.50	0.876	0.013	0.036	PP	0.35	0.30	
122	$\bar{\mathrm{D}}^0(\to \mathrm{K}^-\pi^+\pi^0)\pi^+\pi^+\pi^-\pi^0$	-3.50	-2.93	0.834	0.013	0.034	PP	0.70	0.51	
123	$\bar{\mathrm{D}}^0(\to \mathrm{K}^-\pi^+\pi^0)\pi^+\pi^+\pi^-\pi^0$	-2.93	0.00	0.735	0.011	0.029	PP	0.07	0.06	
124	$\bar{\mathrm{D}}^0(\to \mathrm{K}^-\pi^+\pi^+\pi^-)\pi^+\pi^+\pi^-\pi^0$	$-\infty$	-5.54	1.124	0.018	0.048	PP	0.16	0.13	
125	$\bar{\mathrm{D}}^0(\to \mathrm{K}^-\pi^+\pi^+\pi^-)\pi^+\pi^+\pi^-\pi^0$	-5.54	-4.95	1.039	0.018	0.042	FF	0.69	0.80	43, 44
126	$\bar{\mathrm{D}}^0(\to \mathrm{K}^-\pi^+\pi^+\pi^-)\pi^+\pi^+\pi^-\pi^0$	-4.95	-4.48	0.994	0.018	0.040	FF	0.60	0.50	45, 46
127	$\bar{\mathrm{D}}^0(\to \mathrm{K}^-\pi^+\pi^+\pi^-)\pi^+\pi^+\pi^-\pi^0$	-4.48	-3.98	1.015	0.016	0.041	PP	0.11	0.10	
128	$\bar{\mathrm{D}}^0(\to \mathrm{K}^-\pi^+\pi^+\pi^-)\pi^+\pi^+\pi^-\pi^0$	-3.98	-3.30	0.979	0.015	0.040	PP	0.71	0.55	
129	$\bar{\mathrm{D}}^0(\to \mathrm{K}^-\pi^+\pi^+\pi^-)\pi^+\pi^+\pi^-\pi^0$	-3.30	0.00	0.844	0.013	0.034	PP	0.14	0.17	
130	$\bar{\mathrm{D}}^0(\to \text{ Other})\pi^+\pi^+\pi^-\pi^0$	$-\infty$	-6.42	1.026	0.017	0.050	PP	0.21	0.19	
131	$\bar{\mathrm{D}}^0(\to \text{ Other})\pi^+\pi^+\pi^-\pi^0$	-6.42	-5.98	1.017	0.017	0.045	PP	0.56	0.58	
132	$\bar{\mathrm{D}}^0(\to \text{ Other})\pi^+\pi^+\pi^-\pi^0$	-5.98	-5.68	1.006	0.017	0.043	PP	0.43	0.45	
133	$\bar{\mathrm{D}}^0(\to \text{ Other})\pi^+\pi^+\pi^-\pi^0$	-5.68	-5.45	0.954	0.016	0.039	PP	0.24	0.28	
134	$\bar{\mathrm{D}}^0(\to \text{ Other})\pi^+\pi^+\pi^-\pi^0$	-5.45	-5.27	0.940	0.015	0.036	PP	0.10	0.06	
135	$\bar{\mathrm{D}}^0(\to \text{ Other})\pi^+\pi^+\pi^-\pi^0$	-5.27	-5.10	0.890	0.015	0.036	PP	0.48	0.61	
136	$\bar{\mathrm{D}}^0(\to \text{ Other})\pi^+\pi^+\pi^-\pi^0$	-5.10	-4.95	0.940	0.015	0.036	PP	0.65	0.62	
137	$\bar{\mathrm{D}}^0(\to \text{ Other})\pi^+\pi^+\pi^-\pi^0$	-4.95	-4.80	0.922	0.015	0.037	PP	0.52	0.47	
138	$\bar{\mathrm{D}}^0(\to \text{ Other})\pi^+\pi^+\pi^-\pi^0$	-4.80	-4.66	0.906	0.015	0.036	PP	0.84	0.76	
139	$\bar{\mathrm{D}}^0(\to \text{ Other})\pi^+\pi^+\pi^-\pi^0$	-4.66	-4.51	0.917	0.015	0.036	PP	0.96	0.98	
140	$\bar{\mathrm{D}}^0(\to \text{ Other})\pi^+\pi^+\pi^-\pi^0$	-4.51	-4.37	0.903	0.015	0.037	PP	0.61	0.76	
141	$\bar{\mathrm{D}}^0(\to \text{ Other})\pi^+\pi^+\pi^-\pi^0$	-4.37	-4.21	0.878	0.015	0.035	PP	0.14	0.15	
142	$\bar{\mathrm{D}}^0(\to \text{ Other})\pi^+\pi^+\pi^-\pi^0$	-4.21	-4.05	0.929	0.014	0.038	PP	0.83	0.83	

Continued on next page

i	Channel	L_1	L_2	CF	Stat	Syst	S	p_1	p_2	Notes
143	$\bar{\mathrm{D}}^0(\to \text{ Other})\pi^+\pi^+\pi^-\pi^0$	-4.05	-3.86	0.888	0.014	0.036	PP	0.06	0.05	
144	$\bar{\mathrm{D}}^0(\to \text{ Other})\pi^+\pi^+\pi^-\pi^0$	-3.86	-3.65	0.894	0.014	0.036	PP	0.31	0.16	
145	$\bar{\mathrm{D}}^0(\to \text{ Other})\pi^+\pi^+\pi^-\pi^0$	-3.65	-3.38	0.838	0.014	0.034	PP	0.41	0.42	
146	$\bar{\mathrm{D}}^0(\to \text{ Other})\pi^+\pi^+\pi^-\pi^0$	-3.38	-3.04	0.849	0.014	0.032	PP	0.36	0.34	47
147	$\bar{\mathrm{D}}^0(\to \text{ Other})\pi^+\pi^+\pi^-\pi^0$	-3.04	-2.53	0.826	0.013	0.034	PP	0.19	0.31	
148	$\bar{\mathrm{D}}^0(\to \text{ Other})\pi^+\pi^+\pi^-\pi^0$	-2.53	0.00	0.685	0.011	0.026	PP	0.88	0.85	
149	$\bar{\mathrm{D}}^0\mathrm{D}^+$	$-\infty$	-6.04	1.000	0.033	0.037	PP	0.05	0.04	
150	$\bar{\mathrm{D}}^0\mathrm{D}^+$	-6.04	-4.11	1.009	0.032	0.045	PP	0.68	0.42	
151	$\bar{\mathrm{D}}^0\mathrm{D}^+$	-4.11	0.00	0.947	0.029	0.044	PP	0.89	0.82	
152	$\bar{\mathrm{D}}^0\mathrm{D}^+\mathrm{K}^0_\mathrm{S}$	$-\infty$	-10.17	1.226	0.028	0.052	FF	0.85	0.73	48, 49
153	$\bar{\mathrm{D}}^0\mathrm{D}^+\mathrm{K}^0_\mathrm{S}$	-10.17	-8.26	1.197	0.029	0.051	FF	0.48	0.31	50, 51
154	$\bar{\mathrm{D}}^0\mathrm{D}^+\mathrm{K}^0_\mathrm{S}$	-8.26	0.00	1.055	0.022	0.044	PP	0.69	0.68	
155	$\bar{\mathrm{D}}^{*0}\mathrm{D}^+\mathrm{K}^0_\mathrm{S}$	$-\infty$	0.00	1.206	0.034	0.051	FF	0.22	0.35	52, 53
156	$\bar{\mathrm{D}}^0\mathrm{D}^{*+}\mathrm{K}^0_\mathrm{S}$	$-\infty$	-8.52	1.180	0.047	0.054	PP	0.45	0.53	
157	$\bar{\mathrm{D}}^0\mathrm{D}^{*+}\mathrm{K}^0_\mathrm{S}$	-8.52	0.00	0.945	0.028	0.041	FF	0.74	0.83	54, 55
158	$\bar{\mathrm{D}}^{*0}\mathrm{D}^{*+}\mathrm{K}^0_\mathrm{S}$	$-\infty$	0.00	1.212	0.086	0.052	FF	0.93	0.94	56, 57
159	$\bar{\mathrm{D}}^0\mathrm{D}^0\mathrm{K}^+$	$-\infty$	-8.43	1.160	0.020	0.045	FF	0.69	0.65	58, 59
160	$\bar{\mathrm{D}}^0\mathrm{D}^0\mathrm{K}^+$	-8.43	-7.84	1.118	0.016	0.044	PP	0.14	0.25	
161	$\bar{\mathrm{D}}^0\mathrm{D}^0\mathrm{K}^+$	-7.84	-7.30	1.107	0.017	0.043	PP	0.48	0.62	
162	$\bar{\mathrm{D}}^0\mathrm{D}^0\mathrm{K}^+$	-7.30	-6.74	1.093	0.017	0.042	PP	0.87	0.83	
163	$\bar{\mathrm{D}}^0\mathrm{D}^0\mathrm{K}^+$	-6.74	-6.20	1.116	0.017	0.042	PP	0.69	0.82	
164	$\bar{\mathrm{D}}^0\mathrm{D}^0\mathrm{K}^+$	-6.20	-5.69	1.051	0.017	0.040	PP	0.80	0.81	
165	$\bar{\mathrm{D}}^0\mathrm{D}^0\mathrm{K}^+$	-5.69	-5.21	1.082	0.018	0.050	FP	0.95	0.95	60
166	$\bar{\mathrm{D}}^0\mathrm{D}^0\mathrm{K}^+$	-5.21	-4.70	1.095	0.017	0.047	PP	0.12	0.14	61, 62
167	$\bar{\mathrm{D}}^0\mathrm{D}^0\mathrm{K}^+$	-4.70	-4.04	1.036	0.015	0.044	PP	0.12	0.24	
168	$\bar{\mathrm{D}}^0\mathrm{D}^0\mathrm{K}^+$	-4.04	0.00	1.032	0.015	0.044	PP	0.06	0.11	
169	$\bar{\mathrm{D}}^{*0}\mathrm{D}^0\mathrm{K}^+$	$-\infty$	-6.63	1.093	0.019	0.046	PP	0.04	0.06	
170	$\bar{\mathrm{D}}^{*0}\mathrm{D}^0\mathrm{K}^+$	-6.63	-5.28	1.151	0.016	0.050	PP	0.34	0.50	
171	$\bar{\mathrm{D}}^{*0}\mathrm{D}^0\mathrm{K}^+$	-5.28	-4.37	1.148	0.017	0.059	FP	6×10^{-3}	0.04	63
172	$\bar{\mathrm{D}}^{*0}\mathrm{D}^0\mathrm{K}^+$	-4.37	0.00	1.058	0.014	0.048	PP	0.07	0.09	
173	$\bar{\mathrm{D}}^0\mathrm{D}^{*0}\mathrm{K}^+$	$-\infty$	-6.50	1.138	0.017	0.048	PP	0.05	0.09	64
174	$\bar{\mathrm{D}}^0\mathrm{D}^{*0}\mathrm{K}^+$	-6.50	-5.45	1.109	0.018	0.048	FF	0.08	0.16	65, 66
175	$\bar{\mathrm{D}}^0\mathrm{D}^{*0}\mathrm{K}^+$	-5.45	-4.50	1.141	0.017	0.050	PP	0.27	0.35	
176	$\bar{\mathrm{D}}^0\mathrm{D}^{*0}\mathrm{K}^+$	-4.50	0.00	1.080	0.017	0.051	FP	0.37	0.86	67, 68
177	$\bar{\mathrm{D}}^{*0}\mathrm{D}^{*0}\mathrm{K}^+$	$-\infty$	-6.39	1.109	0.025	0.049	FF	0.76	0.90	69, 70
178	$\bar{\mathrm{D}}^{*0}\mathrm{D}^{*0}\mathrm{K}^+$	-6.39	-4.79	1.161	0.026	0.057	FP	0.53	0.96	71
179	$\bar{\mathrm{D}}^{*0}\mathrm{D}^{*0}\mathrm{K}^+$	-4.79	0.00	1.127	0.025	0.050	FF	0.20	0.16	72, 73
180	$\mathrm{D}_\mathrm{s}^+\bar{\mathrm{D}}^0$	$-\infty$	-3.60	1.138	0.038	0.055	PP	0.52	0.54	
181	$\mathrm{D}_\mathrm{s}^+\bar{\mathrm{D}}^0$	-3.60	-2.24	0.987	0.035	0.049	PP	0.18	0.18	
182	$\mathrm{D}_\mathrm{s}^+\bar{\mathrm{D}}^0$	-2.24	0.00	0.890	0.027	0.039	PP	0.02	0.03	
183	$\bar{\mathrm{D}}^{*0}\pi^+$	$-\infty$	-1.38	1.067	0.031	0.046	PP	0.87	0.93	
184	$\bar{\mathrm{D}}^{*0}\pi^+$	-1.38	-0.48	1.014	0.032	0.045	PP	0.61	0.50	
185	$\bar{\mathrm{D}}^{*0}\pi^+$	-0.48	0.00	1.053	0.032	0.046	FF	0.03	0.04	74, 75
186	$\bar{\mathrm{D}}^{*0}\pi^+\pi^0$	$-\infty$	-3.19	1.055	0.018	0.046	PP	0.54	0.59	
187	$\bar{\mathrm{D}}^{*0}\pi^+\pi^0$	-3.19	-2.14	0.967	0.016	0.041	PP	0.69	0.51	
188	$\bar{\mathrm{D}}^{*0}\pi^+\pi^0$	-2.14	0.00	0.815	0.013	0.034	PP	1×10^{-3}	2×10^{-3}	76
189	$\bar{\mathrm{D}}^{*0}\pi^+\pi^0\pi^0$	$-\infty$	-4.74	0.918	0.014	0.042	PP	0.54	0.44	

Continued on next page

i	Channel	L_1	L_2	CF	Stat	Syst	S	p_1	p_2	Notes
190	$\bar{\mathrm{D}}^{*0}\pi^+\pi^0\pi^0$	-4.74	-4.02	0.928	0.015	0.038	PP	0.40	0.47	
191	$\bar{\mathrm{D}}^{*0}\pi^+\pi^0\pi^0$	-4.02	-3.46	0.993	0.016	0.042	PP	0.19	0.20	
192	$\bar{\mathrm{D}}^{*0}\pi^+\pi^0\pi^0$	-3.46	-2.81	0.965	0.016	0.042	PP	0.36	0.45	
193	$\bar{\mathrm{D}}^{*0}\pi^+\pi^0\pi^0$	-2.81	0.00	0.957	0.015	0.038	PP	0.12	0.20	77
194	$\bar{\mathrm{D}}^{*0}\pi^+\pi^+\pi^-$	$-\infty$	-3.41	1.005	0.015	0.055	PP	0.29	0.47	
195	$\bar{\mathrm{D}}^{*0}\pi^+\pi^+\pi^-$	-3.41	-2.37	0.753	0.012	0.034	PP	0.51	0.50	
196	$\bar{\mathrm{D}}^{*0}\pi^+\pi^+\pi^-$	-2.37	0.00	0.623	0.009	0.025	PP	0.50	0.55	78
197	$\bar{\mathrm{D}}^{*0}(\to \mathrm{D}^0\pi^0)\pi^+\pi^+\pi^-\pi^0$	$-\infty$	-4.93	1.066	0.017	0.053	PP	0.23	0.23	
198	$\bar{\mathrm{D}}^{*0}(\to \mathrm{D}^0\pi^0)\pi^+\pi^+\pi^-\pi^0$	-4.93	-4.38	0.982	0.016	0.044	PP	0.14	0.13	
199	$\bar{\mathrm{D}}^{*0}(\to \mathrm{D}^0\pi^0)\pi^+\pi^+\pi^-\pi^0$	-4.38	-3.98	0.944	0.015	0.041	PP	0.13	0.11	
200	$\bar{\mathrm{D}}^{*0}(\to \mathrm{D}^0\pi^0)\pi^+\pi^+\pi^-\pi^0$	-3.98	-3.60	0.909	0.014	0.038	PP	0.31	0.48	
201	$\bar{\mathrm{D}}^{*0}(\to \mathrm{D}^0\pi^0)\pi^+\pi^+\pi^-\pi^0$	-3.60	-3.18	0.861	0.013	0.036	PP	0.97	0.98	
202	$\bar{\mathrm{D}}^{*0}(\to \mathrm{D}^0\pi^0)\pi^+\pi^+\pi^-\pi^0$	-3.18	-2.59	0.816	0.012	0.034	PP	0.14	0.08	
203	$\bar{\mathrm{D}}^{*0}(\to \mathrm{D}^0\pi^0)\pi^+\pi^+\pi^-\pi^0$	-2.59	0.00	0.680	0.010	0.026	PP	0.52	0.57	
204	$\bar{\mathrm{D}}^{*0}(\to \mathrm{D}^0\gamma)\pi^+\pi^+\pi^-\pi^0$	$-\infty$	-5.37	0.998	0.015	0.052	PP	0.12	0.18	
205	$\bar{\mathrm{D}}^{*0}(\to \mathrm{D}^0\gamma)\pi^+\pi^+\pi^-\pi^0$	-5.37	-4.85	0.957	0.016	0.044	PP	0.93	0.96	
206	$\bar{\mathrm{D}}^{*0}(\to \mathrm{D}^0\gamma)\pi^+\pi^+\pi^-\pi^0$	-4.85	-4.54	0.961	0.015	0.042	PP	0.88	0.70	
207	$\bar{\mathrm{D}}^{*0}(\to \mathrm{D}^0\gamma)\pi^+\pi^+\pi^-\pi^0$	-4.54	-4.29	0.928	0.015	0.040	PP	0.26	0.24	
208	$\bar{\mathrm{D}}^{*0}(\to \mathrm{D}^0\gamma)\pi^+\pi^+\pi^-\pi^0$	-4.29	-4.07	0.872	0.014	0.037	PP	0.57	0.58	
209	$\bar{\mathrm{D}}^{*0}(\to \mathrm{D}^0\gamma)\pi^+\pi^+\pi^-\pi^0$	-4.07	-3.85	0.864	0.013	0.037	PP	0.43	0.40	
210	$\bar{\mathrm{D}}^{*0}(\to \mathrm{D}^0\gamma)\pi^+\pi^+\pi^-\pi^0$	-3.85	-3.64	0.837	0.013	0.036	PP	0.72	0.80	
211	$\bar{\mathrm{D}}^{*0}(\to \mathrm{D}^0\gamma)\pi^+\pi^+\pi^-\pi^0$	-3.64	-3.40	0.835	0.013	0.036	PP	0.39	0.64	
212	$\bar{\mathrm{D}}^{*0}(\to \mathrm{D}^0\gamma)\pi^+\pi^+\pi^-\pi^0$	-3.40	-3.11	0.801	0.012	0.032	PP	0.13	0.22	79
213	$\bar{\mathrm{D}}^{*0}(\to \mathrm{D}^0\gamma)\pi^+\pi^+\pi^-\pi^0$	-3.11	-2.70	0.788	0.011	0.033	PP	0.90	0.91	
214	$\bar{\mathrm{D}}^{*0}(\to \mathrm{D}^0\gamma)\pi^+\pi^+\pi^-\pi^0$	-2.70	0.00	0.685	0.010	0.028	PP	0.64	0.72	
215	$\mathrm{D}_\mathrm{s}^{*+}\bar{\mathrm{D}}^0$	$-\infty$	-3.15	1.062	0.032	0.057	PP	0.12	0.10	
216	$\mathrm{D}_\mathrm{s}^{*+}\bar{\mathrm{D}}^0$	-3.15	0.00	1.051	0.029	0.047	PP	0.54	0.43	80
217	$\mathrm{D}_\mathrm{s}^{+}\bar{\mathrm{D}}^{*0}$	$-\infty$	-3.11	0.981	0.033	0.045	PP	0.94	0.84	
218	$\mathrm{D}_\mathrm{s}^{+}\bar{\mathrm{D}}^{*0}$	-3.11	0.00	0.978	0.030	0.045	PP	0.64	0.64	
219	$\bar{\mathrm{D}}^0\mathrm{K}^+$	$-\infty$	-2.63	1.133	0.039	0.048	PP	0.58	0.51	
220	$\bar{\mathrm{D}}^0\mathrm{K}^+$	-2.63	-1.41	1.301	0.048	0.054	PP	0.80	0.77	
221	$\bar{\mathrm{D}}^0\mathrm{K}^+$	-1.41	0.00	1.159	0.049	0.049	PP	0.10	0.11	81
222	$\mathrm{D}^-\pi^+\pi^+$	$-\infty$	-2.85	1.119	0.040	0.053	PP	0.17	0.15	
223	$\mathrm{D}^-\pi^+\pi^+$	-2.85	-1.81	1.021	0.037	0.044	PP	9×10^{-3}	4×10^{-3}	
224	$\mathrm{D}^-\pi^+\pi^+$	-1.81	0.00	0.558	0.018	0.023	PP	0.15	0.19	
225	$\mathrm{D}^-\pi^+\pi^+\pi^0$	$-\infty$	-4.59	1.058	0.017	0.048	PP	0.73	0.69	
226	$\mathrm{D}^-\pi^+\pi^+\pi^0$	-4.59	-3.84	0.968	0.016	0.042	PP	0.71	0.70	
227	$\mathrm{D}^-\pi^+\pi^+\pi^0$	-3.84	-3.09	0.905	0.015	0.040	PP	0.65	0.59	82
228	$\mathrm{D}^-\pi^+\pi^+\pi^0$	-3.09	0.00	0.736	0.011	0.029	PP	0.93	0.88	
229	$\mathrm{J}/\psi\mathrm{K}^+$	$-\infty$	0.00	1.127	0.043	0.047	PP	0.45	0.44	
230	$\mathrm{J}/\psi\mathrm{K}^+\pi^+\pi^-$	$-\infty$	0.00	0.785	0.025	0.029	PP	0.26	0.16	
231	$\mathrm{J}/\psi\mathrm{K}^+\pi^0$	$-\infty$	0.00	1.077	0.043	0.049	FF	0.08	0.05	83, 84
232	$\mathrm{J}/\psi\mathrm{K}_\mathrm{S}^0\pi^+$	$-\infty$	0.00	1.147	0.083	0.049	FF	0.09	0.10	85, 86

For more information about the following notes, see Section 9.3.1.

(1) Resonant gap fit: Stat only fit uncertainty 1.68% exceeds full fit uncertainty 1.51%. Replacing the latter with the former.

(2) Resonant gap fit: Stat only fit uncertainty 1.51% exceeds full fit uncertainty 1.36%. Replacing the latter with the former.

(3) Non-resonant gap fit: Stat only fit uncertainty 1.45% exceeds full fit uncertainty 1.29%. Replacing the latter with the former.

(4) Resonant gap fit: Stat only fit uncertainty 1.69% exceeds full fit uncertainty 1.41%. Replacing the latter with the former.

(5) Resonant gap fit: Primary fit failed because: Yield $(-1.7 \pm 1.4) \times 10^3$ for component continuum is below zero.

(6) Non-resonant gap fit: Primary fit failed because: Yield $(-1.3 \pm 1.3) \times 10^3$ for component continuum is below zero.

(7) Resonant gap fit: Primary fit failed because: Yield $(-1.2 \pm 1.2) \times 10^3$ for component continuum is below zero.

(8) Resonant gap fit: Using fallback systematic uncertainty of 1×10^5, because stat. only result is not available

(9) Resonant gap fit: Primary fit failed because: Yield $(-6.3 \pm 2.5) \times 10^3$ for component continuum is below zero.

(10) Non-resonant gap fit: Primary fit failed because: Yield $(-5.5 \pm 2.3) \times 10^3$ for component continuum is below zero.

(11) Resonant gap fit: Primary fit failed because: Yield $(-3.3 \pm 1.4) \times 10^3$ for component continuum is below zero.

(12) Non-resonant gap fit: Primary fit failed because: Yield $(-2.8 \pm 1.4) \times 10^3$ for component continuum is below zero.

(13) Resonant gap fit: Primary fit failed because: Yield $(-1.7 \pm 1.4) \times 10^3$ for component continuum is below zero.

(14) Resonant gap fit: Primary fit failed because: Yield $(-1.0 \pm 0.8) \times 10^3$ for component continuum is below zero.

(15) Non-resonant gap fit: Primary fit failed because: Yield $(-8 \pm 7) \times 10^2$ for component continuum is below zero.

(16) Resonant gap fit: Primary fit failed because: Yield $(-1.5 \pm 0.7) \times 10^3$ for component continuum is below zero.

(17) Non-resonant gap fit: Primary fit failed because: Yield $(-1.4 \pm 0.7) \times 10^3$ for component continuum is below zero.

(18) Resonant gap fit: Primary fit failed because: Yield $(-1.3 \pm 1.1) \times 10^2$ for component fake/secondary is below zero.

(19) Resonant gap fit: Secondary fit failed because: Yield $(-1.0 \pm 0.7) \times 10^2$ for component background is below zero.

(20) Non-resonant gap fit: Primary fit failed because: Yield $(-1.4 \pm 1.2) \times 10^2$ for component fake/secondary is below zero.

(21) Non-resonant gap fit: Secondary fit failed because: Yield $(-1.1 \pm 0.7) \times 10^2$ for component background is below zero.

(22) Resonant gap fit: Primary fit failed because: Yield $(-2.4 \pm 2.3) \times 10^2$ for component continuum is below zero.

(23) Resonant gap fit: Primary fit failed because: Yield $(-3.1 \pm 2.5) \times 10^2$ for component continuum is below zero.

(24) Non-resonant gap fit: Primary fit failed because: Yield $(-2.5 \pm 2.3) \times 10^2$ for component continuum is below zero.

(25) Resonant gap fit: Primary fit failed because: Yield $(-4 \pm 4) \times 10^2$ for component continuum is below zero.

(26) Resonant gap fit: Primary fit failed because: Yield $(-6.3 \pm 2.7) \times 10^2$ for component continuum is below zero.

(27) Non-resonant gap fit: Primary fit failed because: Yield $(-5.5 \pm 1.5) \times 10^2$ for component continuum is below zero.

(28) Resonant gap fit: Primary fit failed because: Yield $(-2.5 \pm 1.3) \times 10^3$ for component continuum is below zero.

(29) Non-resonant gap fit: Stat only fit uncertainty 1.17% exceeds full fit uncertainty 1.06%. Replacing the latter with the former.

(30) Non-resonant gap fit: Primary fit failed because: Yield $(-1.7 \pm 1.2) \times 10^3$ for component continuum is below zero.

(31) Resonant gap fit: Stat only fit uncertainty 3.29% exceeds full fit uncertainty 2.87%. Replacing the latter with the former.

(32) Non-resonant gap fit: Stat only fit uncertainty 2.93% exceeds full fit uncertainty 2.64%. Replacing the latter with the former.

(33) Resonant gap fit: Primary fit failed because: Yield $(-8 \pm 5) \times 10^1$ for component continuum is below zero.

(34) Non-resonant gap fit: Primary fit failed because: Yield $(-7 \pm 5) \times 10^1$ for component continuum is below zero.

(35) Resonant gap fit: Primary fit failed because: Yield $(-6 \pm 4) \times 10^1$ for component continuum is below zero.

(36) Resonant gap fit: Using fallback systematic uncertainty of 6×10^4, because stat. only result is not available

(37) Non-resonant gap fit: Using fallback systematic uncertainty of 7×10^4, because stat. only result is not available

(38) Non-resonant gap fit: Stat only fit uncertainty 1.51% exceeds full fit uncertainty 1.47%. Replacing the latter with the former.

(39) Resonant gap fit: Stat only fit uncertainty 1.48% exceeds full fit uncertainty 0.98%. Replacing the latter with the former.

(40) Non-resonant gap fit: Stat only fit uncertainty 1.49% exceeds full fit uncertainty 1.38%. Replacing the latter with the former.

(41) Resonant gap fit: Stat only fit uncertainty 1.57% exceeds full fit uncertainty 1.36%. Replacing the latter with the former.

(42) Non-resonant gap fit: Stat only fit uncertainty 1.53% exceeds full fit uncertainty 1.52%. Replacing the latter with the former.

(43) Resonant gap fit: Primary fit failed because: Yield $(-2.0 \pm 1.5) \times 10^3$ for component continuum is below zero.

(44) Non-resonant gap fit: Primary fit failed because: Yield $(-1.7 \pm 1.4) \times 10^3$ for component continuum is below zero.

(45) Resonant gap fit: Primary fit failed because: Yield $(-2.3 \pm 1.6) \times 10^3$ for component continuum is below zero.

(46) Non-resonant gap fit: Primary fit failed because: Yield $(-1.6 \pm 1.3) \times 10^3$ for component continuum is below zero.

(47) Resonant gap fit: Stat only fit uncertainty 1.46% exceeds full fit uncertainty 1.27%. Replacing the latter with the former.

(48) Resonant gap fit: Primary fit failed because: Yield $(-2.1 \pm 0.8) \times 10^3$ for component continuum is below zero.

(49) Non-resonant gap fit: Primary fit failed because: Yield $(-2.0 \pm 0.8) \times 10^3$ for component continuum is below zero.

(50) Resonant gap fit: Primary fit failed because: Yield $(-2.5 \pm 1.2) \times 10^3$ for component continuum is below zero.

(51) Non-resonant gap fit: Primary fit failed because: Yield $(-2.2 \pm 1.4) \times 10^3$ for component continuum is below zero.

(52) Resonant gap fit: Primary fit failed because: Yield $(-1.8 \pm 1.0) \times 10^3$ for component continuum is below zero.

(53) Non-resonant gap fit: Primary fit failed because: Yield $(-1.8 \pm 0.9) \times 10^3$ for component continuum is below zero.

(54) Resonant gap fit: Primary fit failed because: Yield $(-6 \pm 4) \times 10^2$ for component continuum is below zero.

(55) Non-resonant gap fit: Primary fit failed because: Yield $(-5 \pm 4) \times 10^2$ for component continuum is below zero.

(56) Resonant gap fit: Primary fit failed because: Yield $(-1.1 \pm 1.0) \times 10^2$ for component continuum is below zero.

(57) Non-resonant gap fit: Primary fit failed because: Yield $(-9 \pm 9) \times 10^1$ for component continuum is below zero.

(58) Resonant gap fit: Primary fit failed because: Yield $(-1.6 \pm 1.2) \times 10^3$ for component continuum is below zero.

(59) Non-resonant gap fit: Primary fit failed because: Yield $(-1.2 \pm 1.2) \times 10^3$ for component continuum is below zero.

(60) Resonant gap fit: Primary fit failed because: Yield $(-1.7 \pm 1.4) \times 10^3$ for component continuum is below zero.

(61) Resonant gap fit: Using fallback systematic uncertainty of 1×10^5, because stat. only result is not available

(62) Non-resonant gap fit: Stat only fit uncertainty 1.29% exceeds full fit uncertainty 1.12%. Replacing the latter with the former.

(63) Resonant gap fit: Primary fit failed because: Yield $(-2.3 \pm 1.4) \times 10^3$ for component continuum is below zero.

(64) Resonant gap fit: Using fallback systematic uncertainty of 2×10^5, because stat. only result is not available

(65) Resonant gap fit: Primary fit failed because: Yield $(-5.5 \pm 1.9) \times 10^3$ for component continuum is below zero.

(66) Non-resonant gap fit: Primary fit failed because: Yield $(-4.7 \pm 1.7) \times 10^3$ for component continuum is below zero.

(67) Resonant gap fit: Primary fit failed because: Yield $(-1.5 \pm 1.4) \times 10^3$ for component continuum is below zero.

(68) Non-resonant gap fit: Using fallback systematic uncertainty of 2×10^5, because stat. only result is not available

(69) Resonant gap fit: Primary fit failed because: Yield $(-1.9 \pm 0.8) \times 10^3$ for component continuum is below zero.

(70) Non-resonant gap fit: Primary fit failed because: Yield $(-1.6 \pm 0.9) \times 10^3$ for component continuum is below zero.

(71) Resonant gap fit: Primary fit failed because: Yield $(-1.1 \pm 0.9) \times 10^3$ for component continuum is below zero.

(72) Resonant gap fit: Primary fit failed because: Yield $(-3.8 \pm 1.1) \times 10^3$ for component continuum is below zero.

(73) Non-resonant gap fit: Primary fit failed because: Yield $(-3.5 \pm 1.0) \times 10^3$ for component continuum is below zero.

(74) Resonant gap fit: Primary fit failed because: Yield $(-2.8 \pm 1.8) \times 10^2$ for component continuum is below zero.

(75) Non-resonant gap fit: Primary fit failed because: Yield $(-2.2 \pm 1.8) \times 10^{2}$ for component continuum is below zero.

(76) Resonant gap fit: Using fallback systematic uncertainty of 1×10^{5}, because stat. only result is not available

(77) Non-resonant gap fit: Stat only fit uncertainty 1.37% exceeds full fit uncertainty 1.17%. Replacing the latter with the former.

(78) Non-resonant gap fit: Using fallback systematic uncertainty of 1×10^{5}, because stat. only result is not available

(79) Resonant gap fit: Stat only fit uncertainty 1.27% exceeds full fit uncertainty 1.12%. Replacing the latter with the former.

(80) Non-resonant gap fit: Stat only fit uncertainty 2.33% exceeds full fit uncertainty 1.68%. Replacing the latter with the former.

(81) Non-resonant gap fit: Stat only fit uncertainty 3.82% exceeds full fit uncertainty 2.88%. Replacing the latter with the former.

(82) Non-resonant gap fit: Stat only fit uncertainty 1.37% exceeds full fit uncertainty 1.30%. Replacing the latter with the former.

(83) Resonant gap fit: Primary fit failed because: Yield $(-4.0 \pm 1.8) \times 10^{2}$ for component continuum is below zero.

(84) Non-resonant gap fit: Primary fit failed because: Yield $(-3.6 \pm 1.7) \times 10^{2}$ for component continuum is below zero.

(85) Resonant gap fit: Primary fit failed because: Yield $(-9 \pm 6) \times 10^{1}$ for component continuum is below zero.

(86) Non-resonant gap fit: Primary fit failed because: Yield $(-9 \pm 6) \times 10^{1}$ for component continuum is below zero.

Acronyms and symbols

`b2bii` Belle to Belle II data conversion package of basf2.

ACC Aerogel Cherenkov Counters.

API Application Programming Interface.

basf2 Belle II analysis software framework.

BAU Baryon asymmetry of the Universe.

BCS Best Candidate Selection.

BDT Boosted Decision Tree, a machine learning model.

BGL Form factor parameterization named after Boyd, Grinstein, and Lebed.

BP Benchmark Point.

C.L. Confidence Limit.

CDC Central Drift Chamber.

CF Consistent flavor: Reconstructed flavor of the tag side B meson matches that of the signal side B meson.

$\boldsymbol{C_{\mathrm{FEI}}}$ Reconstruction mode/decay channel chosen by the FEI for the $\mathrm{B_{tag}}$. This can either mean a specification only up to the level of the first daughter particles, or include reconstruction modes for (some of) the daughter particles as well.

CKM Cabibbo–Kobayashi–Maskawa.

CLN Form factor parameterization named after Caprini, Lellouch, and Neubert.

CMS Center of Mass system.

CSF Correct signal flavor: The $\mathrm{B_{sig}}$ flavor is correct.

CTF Correct tag flavor: The $\mathrm{B_{tag}}$ flavor reconstructed by the FEI is correct.

$\boldsymbol{\Delta E_{\mathrm{B_{tag}}}}$ Energy difference between the tag B meson and the CMS momentum of the beam, see Equation (4.1.1).

ECL Electromagnetic Calorimeter.

EFC Extreme Forward Calorimeter.

EFT Effective Field Theory.

EMMS Extended Matrix Mixing Strategy, a strategy to separately calibrate ITF and CTF.

FEI Full Event Interpretation, a tagging approach developed for the Belle II experiment.

FF Form Factor.

FR Full Reconstruction, a tagging approach used at the Belle experiment.

FSP Final State Particle.

HEP High Energy Physics.

HQET Heavy Quark Effective Theory.

HQS Heavy Quark Symmetry.

HSF High Energy Physics Software Foundation.

IF Inconsistent flavor: Reconstructed flavor of the tag side B meson does not match that of the signal side B meson.

ISF Incorrect signal flavor: The $\mathrm{B}_{\mathrm{sig}}$ flavor is incorrect.

ISGW2 Form factor model for $\mathrm{B} \to \mathrm{D}^{**}\ell\nu_\ell$ decays named after Isgur, Scora, Grinstein, and Wise.

ITF Incorrect tag flavor: The $\mathrm{B}_{\mathrm{tag}}$ flavor reconstructed by the FEI is incorrect.

KLM $\mathrm{K}_{\mathrm{L}}^0$ and Muon Detector.

$\boldsymbol{\mathcal{L}_{\mathrm{CS}}}$ Classifier output of the continuum suppression classifier. A number between 0.0 (likely a continuum event) and 1.0 (likely a $\mathrm{B\overline{B}}$ event).

LHC Large Hadron Collider.

LLSW Form factor model for $\mathrm{B} \to \mathrm{D}^{**}\ell\nu_\ell$ decays named after Leibovich, Ligeti, Stewart, Wise.

$\boldsymbol{m_{\mathrm{bc}}^{\mathrm{tag}}}$ Beam constrained mass of the tag B meson, see Equation (4.1.1).

MC Monte Carlo.

MMS Matrix Mixing Strategy, a strategy to separately calibrate ITF and CTF.

$\boldsymbol{N_{\mathrm{cat}}}$ Total number of categories (note that categories indices start at zero).

$|\vec{p}_\ell^{\,*}|$ Lepton momentum in the B rest frame.

NP New Physics (as opposed to the physics of the SM).

NRG Non-resonant gap model, see Section 7.3.4.

OPE Operator Product Expansion.

PDF Probability Density Function.

$\boldsymbol{\mathcal{P}_{\mathrm{FEI}}}$ Classifier output of the FEI (signal probability).

PID Particle Identity.

QCD Quantum Chromodynamics.

QED Quantum Electrodynamics.

RG Resonant gap model, see Section 7.3.4.

ROE Rest Of Event. For a particular selection of reconstructed particles, the corresponding ROE is the complementary selection that includes all unassociated particles.

SF Flavor of the B_{sig}, if reconstructed.

SM Standard Model of particle physics.

SUSY Supersymmetry.

SVD Silicon Vertex Detector.

SVD1 Silicon Vertex Detector, first configuration.

SVD2 Silicon Vertex Detector, second configuration.

TF Flavor of the B_{tag} (i.e., flavor as reported by the FEI).

TOF Time Of Flight system.

V0 Detector signature of a neutral particle decaying into two charged tracks, forming a V-shape.

List of Tables

List of Figures

Bibliography

[1] S. Weinberg. *The Quantum Theory of Fields. Vol. 1: Foundations*. Cambridge University Press, 2005. ISBN: 978-0-521-67053-1.

[2] S. Weinberg. *The Quantum Theory of Fields. Vol. 2: Modern Applications*. Cambridge University Press, 2013. ISBN: 978-0-521-67054-8.

[3] S. Weinberg. "The Making of the Standard Model". In: *The European Physical Journal C* 34.1 (May 2004), pp. 5–13. ISSN: 1434-6044, 1434-6052. DOI: 10.1140/epjc/s2004-01761-1. arXiv: hep-ph/0401010.

[4] A. Pich. "The Standard Model of Electroweak Interactions". Feb. 1, 2005. arXiv: hep-ph/0502010. URL: http://arxiv.org/abs/hep-ph/0502010 (visited on 01/08/2022).

[5] W. Buchmuller and C. Ludeling. "Field Theory and Standard Model". In: *High-Energy Physics. Proceedings, European School, Kitzbuehel, Austria, August 21-September, 2005*. 2006. URL: https://arxiv.org/abs/hep-ph/0609174.

[6] M. E. Peskin and D. V. Schroeder. *An Introduction to Quantum Field Theory*. Reading, USA: Addison-Wesley, 1995. ISBN: 978-0-201-50397-5. URL: http://www.slac.stanford.edu/%20mpeskin/QFT.html.

[7] *Standard Model of Elementary Particles*. URL: https://commons.wikimedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg (visited on 01/08/2022).

[8] E. Drexler. *Elementary Particle Interactions in the Standard Model*. Apr. 17, 2014. URL: https://commons.wikimedia.org/wiki/File:Elementary_particle_interactions_in_the_Standard_Model.png (visited on 01/08/2022).

[9] P. W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons". In: *Phys. Rev. Lett.* 13 (1964), pp. 508–509. DOI: 10.1103/PhysRevLett.13.508.

[10] F. Englert and R. Brout. "Broken Symmetry and the Mass of Gauge Vector Mesons". In: *Phys. Rev. Lett.* 13 (1964), pp. 321–323. DOI: 10.1103/PhysRevLett.13.321.

[11] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. "Global Conservation Laws and Massless Particles". In: *Phys. Rev. Lett.* 13 (1964), pp. 585–587. DOI: 10.1103/PhysRevLett.13.585.

[12] Y. Nambu. "Quasi-Particles and Gauge Invariance in the Theory of Superconductivity". In: *Physical Review* 117.3 (Feb. 1, 1960), pp. 648–663. ISSN: 0031-899X. DOI: 10.1103/PhysRev.117.648.

[13] J. Goldstone. "Field Theories with 'Superconductor' Solutions". In: *Il Nuovo Cimento* 19.1 (Jan. 1961), pp. 154–164. ISSN: 0029-6341, 1827-6121. DOI: 10.1007/BF02812722.

[14] J. Goldstone, A. Salam, and S. Weinberg. "Broken Symmetries". In: *Physical Review* 127.3 (Aug. 1, 1962), pp. 965–970. ISSN: 0031-899X. DOI: 10.1103/PhysRev.127.965.

[15] B. C. Hall. *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*. Vol. 222. Graduate Texts in Mathematics. Cham: Springer International Publishing, 2015. ISBN: 978-3-319-13466-6. DOI: 10.1007/978-3-319-13467-3.

[16] N. Cabibbo. "Unitary Symmetry and Leptonic Decays". In: *Physical Review Letters* 10.12 (June 15, 1963), pp. 531–533. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.10.531.

[17] M. Kobayashi and T. Maskawa. "CP-Violation in the Renormalizable Theory of Weak Interaction". In: *Progress of Theoretical Physics* 49.2 (Feb. 1973), pp. 652–657. ISSN: 0033-068X. DOI: 10.1143/PTP.49.652.

[18] B. Pontecorvo. "Mesonium and Anti-Mesonium". In: *Soviet Physics–JETP [translation of Zhurnal Eksperimentalnoi i Teoreticheskoi Fiziki]* 6 (1957), p. 429.

[19] Y. Fukuda et al. "Evidence for Oscillation of Atmospheric Neutrinos". In: *Physical Review Letters* 81.8 (Aug. 24, 1998), pp. 1562–1567. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.81.1562.

[20] Q. R. Ahmad et al. "Measurement of the Rate of $\nu_e + d \rightarrow p + p + e^-$ Interactions Produced by ^8B Solar Neutrinos at the Sudbury Neutrino Observatory". In: *Physical Review Letters* 87.7 (July 25, 2001), p. 071301. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.87.071301.

[21] I. Esteban, M. C. Gonzalez-Garcia, M. Maltoni, T. Schwetz, and A. Zhou. "The Fate of Hints: Updated Global Analysis of Three-Flavor Neutrino Oscillations". In: *Journal of High Energy Physics* 2020.9 (Sept. 2020), p. 178. ISSN: 1029-8479. DOI: 10.1007/JHEP09(2020)178. arXiv: 2007.14792.

[22] Z. Maki, M. Nakagawa, and S. Sakata. "Remarks on the Unified Model of Elementary Particles". In: *Progress of Theoretical Physics* 28.5 (Nov. 1962), pp. 870–880. ISSN: 0033-068X. DOI: 10.1143/PTP.28.870.

[23] S. Descotes-Genon and P. Koppenburg. "The CKM Parameters". In: *Annual Review of Nuclear and Particle Science* 67.1 (Oct. 12, 2017), pp. 97–127. ISSN: 0163-8998, 1545-4134. DOI: 10.1146/annurev-nucl-101916-123109. arXiv: 1702.08834.

[24] J. Charles, A. Höcker, H. Lacker, S. Laplace, F. R. Le Diberder, J. Malclés, J. Ocariz, M. Pivk, and L. Roos. "CP Violation and the CKM Matrix: Assessing the Impact of the Asymmetric B Factories". In: *The European Physical Journal C* 41.1 (May 2005), pp. 1–131. ISSN: 1434-6044, 1434-6052. DOI: 10.1140/epjc/s2005-02169-1.

[25] Particle Data Group et al. "Review of Particle Physics: Vcb and Vub". In: *Progress of Theoretical and Experimental Physics* 2020.8 (Aug. 14, 2020), p. 083C01. ISSN: 2050-3911. DOI: 10.1093/ptep/ptaa104.

[26] L.-L. Chau and W.-Y. Keung. "Comments on the Parametrization of the Kobayashi-Maskawa Matrix". In: *Physical Review Letters* 53.19 (Nov. 5, 1984), pp. 1802–1805. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.53.1802.

[27] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay. "Evidence for the 2π Decay of the K_2^0 Meson". In: *Physical Review Letters* 13.4 (July 27, 1964), pp. 138–140. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.13.138.

[28] L. Wolfenstein. "Parametrization of the Kobayashi-Maskawa Matrix". In: *Physical Review Letters* 51.21 (Nov. 21, 1983), pp. 1945–1947. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.51.1945.

[29] A. J. Buras, M. E. Lautenbacher, and G. Ostermaier. "Waiting for the top quark mass, $K^+ \rightarrow \pi^+ \nu\bar{\nu}$, B_s^0-$\bar{B}_s^0$ mixing, and CP asymmetries in B-decays". In: *Physical Review D* 50.5 (Sept. 1, 1994), pp. 3433–3446. ISSN: 0556-2821. DOI: 10.1103/PhysRevD.50.3433.

[30] C. Jarlskog. "Commutator of the Quark Mass Matrices in the Standard Electroweak Model and a Measure of Maximal CP Nonconservation". In: *Physical Review Letters* 55.10 (Sept. 2, 1985), pp. 1039–1042. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.55.1039.

[31] The ATLAS Collaboration. "Observation of a New Particle in the Search for the Standard Model Higgs Boson with the ATLAS Detector at the LHC". In: *Physics Letters B* 716.1 (Sept. 2012), pp. 1–29. ISSN: 03702693. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214.

[32] The CMS Collaboration. "Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC". In: *Physics Letters B* 716.1 (Sept. 2012), pp. 30–61. ISSN: 03702693. DOI: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235.

[33] R. Mann. *An Introduction to Particle Physics and the Standard Model*. 0th ed. CRC Press, Nov. 18, 2009. ISBN: 978-1-4200-8300-2. DOI: 10.1201/9781420083002.

[34] C. Rovelli. "QUANTUM GRAVITY". In: *Philosophy of Physics*. Elsevier, 2007, pp. 1287–1329. ISBN: 978-0-444-51560-5. DOI: 10.1016/B978-044451560-5/50015-4.

[35] A. Grinbaum. "Which Fine-Tuning Arguments Are Fine?" In: *Foundations of Physics* 42.5 (May 2012), pp. 615–631. ISSN: 0015-9018, 1572-9516. DOI: 10.1007/s10701-012-9629-9. arXiv: 0903.4055.

[36] A. Hook. "TASI Lectures on the Strong CP Problem and Axions". July 20, 2021. arXiv: 1812.02669 [hep-ph]. URL: http://arxiv.org/abs/1812.02669 (visited on 01/18/2022).

[37] R. D. Peccei and H. R. Quinn. "CP Conservation in the Presence of Pseudoparticles". In: *Physical Review Letters* 38.25 (June 20, 1977), pp. 1440–1443. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.38.1440.

[38] S. P. Martin. "A Supersymmetry Primer". July 1998. DOI: 10.1142/9789812839657_0001. arXiv: hep-ph/9709356.

[39] S. J. Gates Jr. "Sticking with SUSY". In: *Physics World* 27.10 (Oct. 2014), pp. 32–37. ISSN: 0953-8585, 2058-7058. DOI: 10.1088/2058-7058/27/10/33.

[40] G. Bertone and D. Hooper. "A History of Dark Matter". In: *Reviews of Modern Physics* 90.4 (Oct. 15, 2018), p. 045002. ISSN: 0034-6861, 1539-0756. DOI: 10.1103/RevModPhys.90.045002. arXiv: 1605.04909.

[41] F. Zwicky. "Republication of: The Redshift of Extragalactic Nebulae". In: *General Relativity and Gravitation* 41.1 (Jan. 2009), pp. 207–224. ISSN: 0001-7701, 1572-9532. DOI: 10.1007/s10714-008-0707-4.

[42] J. D. Simon. "The Faintest Dwarf Galaxies". In: *Annual Review of Astronomy and Astrophysics* 57.1 (Aug. 18, 2019), pp. 375–415. ISSN: 0066-4146, 1545-4282. DOI: 10.1146/annurev-astro-091918-104453. arXiv: 1901.05465.

[43] P. Salucci. "The Distribution of Dark Matter in Galaxies". In: *The Astronomy and Astrophysics Review* 27.1 (Dec. 2019), p. 2. ISSN: 0935-4956, 1432-0754. DOI: 10.1007/s00159-018-0113-1. arXiv: 1811.08843.

[44] S. W. Allen, A. E. Evrard, and A. B. Mantz. "Cosmological Parameters from Observations of Galaxy Clusters". In: *Annual Review of Astronomy and Astrophysics* 49.1 (Sept. 22, 2011), pp. 409–470. ISSN: 0066-4146, 1545-4282. DOI: 10.1146/annurev-astro-081710-102514. arXiv: 1103.4829.

[45] Planck Collaboration et al. "Planck 2018 Results. VI. Cosmological Parameters". In: *Astronomy & Astrophysics* 641 (Sept. 2020), A6. ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361/201833910. arXiv: 1807.06209.

[46] J. Lesgourgues and S. Pastor. "Massive Neutrinos and Cosmology". In: *Physics Reports* 429.6 (July 2006), pp. 307–379. ISSN: 03701573. DOI: 10.1016/j.physrep.2006.04.001. arXiv: astro-ph/0603494.

[47] T. Clifton, P. G. Ferreira, A. Padilla, and C. Skordis. "Modified Gravity and Cosmology". In: *Physics Reports* 513.1-3 (Mar. 2012), pp. 1–189. ISSN: 03701573. DOI: 10.1016/j.physrep.2012.01.001. arXiv: 1106.2476.

[48] P. G. Ferreira. "Cosmological Tests of Gravity". In: *Annual Review of Astronomy and Astrophysics* 57.1 (Aug. 18, 2019), pp. 335–374. ISSN: 0066-4146, 1545-4282. DOI: 10.1146/annurev-astro-091918-104423. arXiv: 1902.10503.

[49] K. Pardo and D. N. Spergel. "What Is the Price of Abandoning Dark Matter? Cosmological Constraints on Alternative Gravity Theories". In: *Physical Review Letters* 125.21 (Nov. 16, 2020), p. 211101. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.125.211101. arXiv: 2007.00555.

[50] J. M. Ezquiaga and M. Zumalacárregui. "Dark Energy after GW170817: Dead Ends and the Road Ahead". In: *Physical Review Letters* 119.25 (Dec. 18, 2017), p. 251304. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.119.251304. arXiv: 1710.05901.

[51] T. Baker, E. Bellini, P. G. Ferreira, M. Lagos, J. Noller, and I. Sawicki. "Strong Constraints on Cosmological Gravity from GW170817 and GRB 170817A". In: *Physical Review Letters* 119.25 (Dec. 18, 2017), p. 251301. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.119.251301. arXiv: 1710.06394.

[52] P. Dirac. "The Quantum Theory of the Electron. Part II". In: *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 118.779 (Mar. 1928), pp. 351–361. ISSN: 0950-1207, 2053-9150. DOI: 10.1098/rspa.1928.0056.

[53] B. Abi et al. "Measurement of the Positive Muon Anomalous Magnetic Moment to 0.46 Ppm". In: *Physical Review Letters* 126.14 (Apr. 7, 2021), p. 141801. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.126.141801.

[54] J. P. Miller, R. Eduardo de, B. L. Roberts, and D. Stöckinger. "Muon $g-2$: Experiment and Theory". In: *Annual Review of Nuclear and Particle Science* 62.1 (Nov. 23, 2012), pp. 237–264. ISSN: 0163-8998, 1545-4134. DOI: 10.1146/annurev-nucl-031312-120340.

[55] J. Schwinger. "On Quantum-Electrodynamics and the Magnetic Moment of the Electron". In: *Physical Review* 73.4 (Feb. 15, 1948), pp. 416–417. ISSN: 0031-899X. DOI: 10.1103/PhysRev.73.416.

[56] L. Canetti, M. Drewes, and M. Shaposhnikov. "Matter and Antimatter in the Universe". In: *New Journal of Physics* 14.9 (Sept. 17, 2012), p. 095012. ISSN: 1367-2630. DOI: 10.1088/1367-2630/14/9/095012. arXiv: 1204.4186.

[57] A. D. Sakharov. "Violation of CP invariance, C asymmetry, and baryon asymmetry of the universe". In: *Soviet Physics Uspekhi* 34.5 (May 31, 1991), pp. 392–393. ISSN: 0038-5670. DOI: 10.1070/PU1991v034n05ABEH002497.

[58] N. S. Manton. "Topology in the Weinberg-Salam Theory". In: *Physical Review D* 28.8 (Oct. 15, 1983), pp. 2019–2026. ISSN: 0556-2821. DOI: 10.1103/PhysRevD.28.2019.

[59] F. R. Klinkhamer and N. S. Manton. "A Saddle-Point Solution in the Weinberg-Salam Theory". In: *Physical Review D* 30.10 (Nov. 15, 1984), pp. 2212–2220. ISSN: 0556-2821. DOI: 10.1103/PhysRevD.30.2212.

[60] G. Elor, J. Harz, S. Ipek, and B. Shakya. "Snowmass 2021 Letter of Interest: The Baryon Asymmetry Of The Universe". In: 2021.

[61] P. Arnold, D. Son, and L. G. Yaffe. "The hot baryon violation rate is $O(\alpha_W^5 T^4)$". In: *Physical Review D* 55.10 (May 15, 1997), pp. 6264–6273. ISSN: 0556-2821, 1089-4918. DOI: 10.1103/PhysRevD.55.6264. arXiv: hep-ph/9609481.

[62] P. Arnold and L. McLerran. "Sphalerons, Small Fluctuations, and Baryon-Number Violation in Electroweak Theory". In: *Physical Review D* 36.2 (July 15, 1987), pp. 581–595. ISSN: 0556-2821. DOI: 10.1103/PhysRevD.36.581.

[63] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson. "Experimental Test of Parity Conservation in Beta Decay". In: *Physical Review* 105.4 (Feb. 15, 1957), pp. 1413–1415. ISSN: 0031-899X. DOI: 10.1103/PhysRev.105.1413.

[64] LHCb collaboration et al. "Test of lepton universality using $B^+ \to K^+\ell^+\ell^-$ decays". In: *Physical Review Letters* 113.15 (Oct. 6, 2014), p. 151601. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.113.151601. arXiv: 1406.6482.

[65] LHCb collaboration et al. "Test of lepton universality with $B^0 \to K^{*0}\ell^+\ell^-$ decays". In: *Journal of High Energy Physics* 2017.8 (Aug. 2017), p. 55. ISSN: 1029-8479. DOI: 10.1007/JHEP08(2017)055. arXiv: 1705.05802.

[66] LHCb Collaboration et al. "Search for lepton-universality violation in $B^+ \to K^+\ell^+\ell^-$ decays". In: *Physical Review Letters* 122.19 (May 13, 2019), p. 191801. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.122.191801. arXiv: 1903.09252.

[67] LHCb collaboration et al. "Test of Lepton Universality in Beauty-Quark Decays". Mar. 22, 2021. arXiv: 2103.11769 [hep-ex]. URL: http://arxiv.org/abs/2103.11769 (visited on 04/07/2022).

[68] LHCb collaboration et al. "Measurement of form-factor independent observables in the decay $B^0 \to K^{*0}\mu^+\mu^-$". In: *Physical Review Letters* 111.19 (Nov. 4, 2013), p. 191801. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.111.191801. arXiv: 1308.1707.

[69] LHCb collaboration et al. "Angular analysis of the $B^0 \to K^{*0}\mu^+\mu^-$ decay using $3\,\mathrm{fb}^{-1}$ of integrated luminosity". In: *Journal of High Energy Physics* 2016.2 (Feb. 2016), p. 104. ISSN: 1029-8479. DOI: 10.1007/JHEP02(2016)104. arXiv: 1512.04442.

[70] HFLAV Collaboration. *Average of $R(D)$ and $R(D^*)$ - Spring 2021 update*. URL: https://hflav-eos.web.cern.ch/hflav-eos/semi/spring21/html/RDsDsstar/RDRDs.html (visited on 02/06/2022).

[71] The BaBar Collaboration. "Evidence for an excess of $B \to D^{(*)}\tau\nu$ decays". In: *Physical Review Letters* 109.10 (Sept. 6, 2012), p. 101802. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.109.101802. arXiv: 1205.5442.

[72] The BABAR Collaboration, J. P. Lees, et al. "Measurement of an Excess of $B \to D^{(*)}\tau\nu$ Decays and Implications for Charged Higgs Bosons". In: *Physical Review D* 88.7 (Oct. 31, 2013), p. 072012. ISSN: 1550-7998, 1550-2368. DOI: 10.1103/PhysRevD.88.072012. arXiv: 1303.0571.

[73] Belle Collaboration 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: *Physical Review D* 92.7 (Oct. 26, 2015), p. 072014. ISSN: 1550-7998, 1550-2368. DOI: 10.1103/PhysRevD.92.072014. arXiv: 1507.03233.

[74] R. Aaij et al. "Measurement of the ratio of branching fractions $\mathcal{B}(\overline{B}^0 \to D^{*+}\tau^-\overline{\nu}_\tau)/\mathcal{B}(\overline{B}^0 \to D^{*+}\mu^-\overline{\nu}_\mu)$". In: *Physical Review Letters* 115.11 (Sept. 9, 2015), p. 111803. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.115.111803. arXiv: 1506.08614.

[75] S. Hirose et al. "Measurement of the τ lepton polarization and $R(D^*)$ in the decay $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$". In: *Physical Review Letters* 118.21 (May 26, 2017), p. 211801. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.118.211801. arXiv: 1612.00529.

[76] 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: *Physical Review D* 97.1 (Jan. 10, 2018), p. 012004. ISSN: 2470-0010, 2470-0029. DOI: 10.1103/PhysRevD.97.012004. arXiv: 1709.00129.

[77] 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.

[78] LHCb collaboration et al. "Test of Lepton Flavor Universality by the measurement of the $B^0 \to D^{*-}\tau^+\nu_\tau$ branching fraction using three-prong τ decays". In: *Physical Review D* 97.7 (Apr. 25, 2018), p. 072013. ISSN: 2470-0010, 2470-0029. DOI: 10.1103/PhysRevD.97.072013. arXiv: 1711.02505.

[79] The Belle Collaboration et al. "Measurement of $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ with a semileptonic tagging method". In: *Physical Review Letters* 124.16 (Apr. 24, 2020), p. 161803. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.124.161803. arXiv: 1910.05864.

[80] D. Bigi and P. Gambino. "Revisiting $B \to D\ell\nu$". In: *Physical Review D* 94.9 (Nov. 11, 2016), p. 094008. ISSN: 2470-0010, 2470-0029. DOI: 10.1103/PhysRevD.94.094008. arXiv: 1606.08030.

[81] P. Gambino, M. Jung, and S. Schacht. "The V_{cb} puzzle: an update". In: *Physics Letters B* 795 (Aug. 2019), pp. 386–390. ISSN: 03702693. DOI: 10.1016/j.physletb.2019.06.039. arXiv: 1905.08209.

[82] M. Bordone, M. Jung, and D. van Dyk. "Theory determination of $\bar{B} \to D^{(*)}\ell^-\bar{\nu}$ form factors at $\mathcal{O}(1/m_c^2)$". In: *The European Physical Journal C* 80.2 (Feb. 2020), p. 74. ISSN: 1434-6044, 1434-6052. DOI: 10.1140/epjc/s10052-020-7616-4. arXiv: 1908.09398.

[83] F. Metzner. "Preparation of a Measurement of $\mathcal{R}(D^{(*)})$ with Leptonic τ and Hadronic FEI Tag at the Belle Experiment". Karlsruhe: Karlsruhe Institute of Technology (KIT), Feb. 2022.

[84] F. Metzner. *$R(D^*)$ with Leptonic τ and Analysis of $B \to D^{(*)}\ell\nu$ Decays with Hadronic FEI.* Belle Note 1616.

[85] A. V. Manohar and M. B. Wise. "Inclusive Semileptonic B and Polarized Λ_b Decays from QCD". In: *Physical Review D* 49.3 (Feb. 1, 1994), pp. 1310–1329. ISSN: 0556-2821. DOI: 10.1103/PhysRevD.49.1310. arXiv: hep-ph/9308246.

[86] I. I. Bigi, M. Shifman, N. G. Uraltsev, and A. Vainshtein. "QCD Predictions for Lepton Spectra in Inclusive Heavy Flavour Decays". In: *Physical Review Letters* 71.4 (July 26, 1993), pp. 496–499. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.71.496. arXiv: hep-ph/9304225.

[87] O. Buchmueller and H. Flaecher. "Fit to Moments of Inclusive $B \to X_c\ell\nu_\ell$ and $B \to X_s\gamma$ Decay Distributions using Heavy Quark Expansions in the Kinetic Scheme". In: *Physical Review D* 73.7 (Apr. 18, 2006), p. 073008. ISSN: 1550-7998, 1550-2368. DOI: 10.1103/PhysRevD.73.073008. arXiv: hep-ph/0507253.

[88] M. Fael, T. Mannel, and K. K. Vos. "Vcb Determination from Inclusive b → c Decays: An Alternative Method". In: *Journal of High Energy Physics* 2019.2 (Feb. 2019), p. 177. ISSN: 1029-8479. DOI: 10.1007/JHEP02(2019)177.

[89] R. van Tonder et al. "Measurements of q^2 moments of inclusive $B \to X_c\ell\nu_\ell$ decays with hadronic tagging". In: *Physical Review D* 104.11 (Dec. 30, 2021), p. 112011. ISSN: 2470-0010, 2470-0029. DOI: 10.1103/PhysRevD.104.112011.

[90] Particle Data Group et al. "Review of Particle Physics: Introduction". In: *Progress of Theoretical and Experimental Physics* 2020.8 (Aug. 14, 2020). ISSN: 2050-3911. DOI: 10.1093/ptep/ptaa104.

[91] D. Buskulic et al. "Measurements of $|V_{cb}|$, form factors and branching fractions in the decays $\bar{B}^0 \to D^{*+}\ell^-\bar{\nu}_e ll$ and $\bar{B}^0 \to D^+\ell^-\bar{\nu}_\ell$". In: *Physics Letters B* 395.3-4 (Mar. 1997), pp. 373–387. ISSN: 03702693. DOI: 10.1016/S0370-2693(97)00071-3.

[92] G. Abbiendi et al. "Measurement of $|V_{cb}|$ using $\bar{B}^0 \to D^{*+}\ell^-\bar{\nu}_\ell$ decays". In: *Physics Letters B* 482.1-3 (June 2000), pp. 15–30. ISSN: 03702693. DOI: 10.1016/S0370-2693(00)00457-3.

[93] P. Abreu et al. "Measurement of $|V_{cb}|$ from the decay process $\bar{B}^0 \to D^{*+}\ell^-\bar{\nu}_\ell$". In: *Physics Letters B* 510.1-4 (June 2001), pp. 55–74. ISSN: 03702693. DOI: 10.1016/S0370-2693(01)00569-X.

[94] N. E. Adam et al. "Determination of the $\bar{B} \to D^*\ell\nu_\ell$ decay width and $|V_{cb}|$". In: *Physical Review D* 67.3 (Feb. 27, 2003), p. 032001. ISSN: 0556-2821, 1089-4918. DOI: 10.1103/PhysRevD.67.032001.

[95] The DELPHI Collaboration. "Measurement of $|V_{cb}|$ using the semileptonic decay $\overline{\mathrm{B}^0_\mathrm{d}} \to \mathrm{D}^*\ell^-\bar{\nu}_\ell$". In: *The European Physical Journal C* 33.2 (Mar. 2004), pp. 213–232. ISSN: 1434-6044, 1434-6052. DOI: 10.1140/epjc/s2004-01598-6.

[96] B. Aubert et al. "Measurements of the semileptonic decays $\bar{B} \to D\ell\bar{\nu}_\ell$ and $\bar{B} \to D^*\ell\bar{\nu}_\ell$ using a global fit to $DX\ell\bar{\nu}_\ell$ final states". In: *Physical Review D* 79.1 (Jan. 15, 2009), p. 012002. ISSN: 1550-7998, 1550-2368. DOI: 10.1103/PhysRevD.79.012002.

[97] E. Waheed et al. "Measurement of the CKM Matrix Element $|V_{cb}|$ from $B^0 \to D^{*-}\ell^+\nu_\ell$ at Belle". In: *Physical Review D* (Apr. 19, 2019). DOI: 10.1103/PhysRevD.100.052007. arXiv: 1809.03290.

[98] Particle Data Group et al. "Review of Particle Physics". In: *Progress of Theoretical and Experimental Physics* 2020.8 (Aug. 14, 2020), p. 083C01. ISSN: 2050-3911. DOI: 10.1093/ptep/ptaa104.

[99] Y. Amhis et al. "Averages of b-hadron, c-hadron, and τ-lepton properties as of 2018: Heavy Flavor Averaging Group (HFLAV)". In: *The European Physical Journal C* 81.3 (Mar. 2021), p. 226. ISSN: 1434-6044, 1434-6052. DOI: 10.1140/epjc/s10052-020-8156-7.

[100] M. Neubert. "Heavy-Quark Symmetry". In: *Physics Reports* 245.5-6 (Sept. 1994), pp. 259–395. ISSN: 03701573. DOI: 10.1016/0370-1573(94)90091-4.

[101] A. Sirlin. "Large m_W, m_Z behaviour of the $\mathcal{O}(\alpha)$ corrections to semileptonic processes mediated by W". In: *Nuclear Physics B* 196.1 (Mar. 1982), pp. 83–92. ISSN: 05503213. DOI: 10.1016/0550-3213(82)90303-0.

[102] F. U. Bernlochner, M. F. Sevilla, D. J. Robinson, and G. Wormser. "Semitauonic b-hadron decays: A lepton flavor universality laboratory". In: *Reviews of Modern Physics* 94.1 (Feb. 4, 2022), p. 015003. ISSN: 0034-6861, 1539-0756. DOI: 10.1103/RevModPhys.94.015003. arXiv: 2101.08326.

[103] 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: *Physical Review D* 89.11 (June 19, 2014), p. 114504. ISSN: 1550-7998, 1550-2368. DOI: 10.1103/PhysRevD.89.114504. arXiv: 1403.0635.

[104] A. V. Avilés-Casco, C. DeTar, D. Du, A. El-Khadra, A. Kronfeld, J. Laiho, and R. S. Van de Water. "$B \to D^*\ell\nu$ at non-zero recoil". In: *EPJ Web of Conferences* 175 (2018). Ed. by M. Della Morte, P. Fritzsch, E. Gámiz Sánchez, and C. Pena Ruano, p. 13003. ISSN: 2100-014X. DOI: 10.1051/epjconf/201817513003.

[105] A. Vaquero, C. DeTar, A. X. El-Khadra, A. S. Kronfeld, J. Laiho, and R. S. Van de Water. "$B \to D^*\ell\nu$ at non-zero recoil". June 7, 2019. arXiv: 1906.01019 [hep-ex, physics:hep-lat]. URL: http://arxiv.org/abs/1906.01019 (visited on 03/31/2022).

[106] D. Ferlewicz, P. Urquijo, and E. Waheed. "Revisiting fits to $B^0 \to D^{*-}\ell^+\nu_\ell$ to measure $|V_{cb}|$ with novel methods and preliminary LQCD data at nonzero recoil". In: *Physical Review D* 103.7 (Apr. 29, 2021), p. 073005. ISSN: 2470-0010, 2470-0029. DOI: 10.1103/PhysRevD.103.073005.

[107] S. Aoki et al. “FLAG Review 2019”. In: *The European Physical Journal C* 80.2 (Feb. 2020), p. 113. ISSN: 1434-6044, 1434-6052. DOI: 10.1140/epjc/s10052-019-7354-7. arXiv: 1902.08191.

[108] J. Harrison, C. Davies, and M. Wingate. “Lattice QCD calculation of the $B_{(s)} \to D^*_{(s)}\ell\nu$ form factors at zero recoil and implications for $|V_{cb}|$”. In: *Physical Review D* 97.5 (Mar. 9, 2018), p. 054502. ISSN: 2470-0010, 2470-0029. DOI: 10.1103/PhysRevD.97.054502. arXiv: 1711.11013.

[109] 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: *Physical Review D* 95.11 (June 8, 2017), p. 115008. ISSN: 2470-0010, 2470-0029. DOI: 10.1103/PhysRevD.95.115008. arXiv: 1703.05330.

[110] N. Isgur, D. Scora, B. Grinstein, and M. B. Wise. “Semileptonic B and D Decays in the Quark Model”. In: *Physical Review D* 39.3 (Feb. 1, 1989), pp. 799–818. ISSN: 0556-2821. DOI: 10.1103/PhysRevD.39.799.

[111] N. Isgur and M. B. Wise. “Weak Decays of Heavy Mesons in the Static Quark Approximation”. In: *Physics Letters B* 232.1 (Nov. 1989), pp. 113–117. ISSN: 03702693. DOI: 10.1016/0370-2693(89)90566-2.

[112] C. G. Boyd, B. Grinstein, and R. F. Lebed. “Model-Independent Determinations of $B \to D\ell\nu_\ell$, $D^*\ell\nu_\ell$ Form Factors”. Aug. 1, 1995. arXiv: hep-ph/9508211. URL: http://arxiv.org/abs/hep-ph/9508211 (visited on 02/27/2022).

[113] C. G. Boyd, B. Grinstein, and R. F. Lebed. “Precision Corrections to Dispersive Bounds on Form Factors”. In: *Physical Review D* 56.11 (Dec. 1, 1997), pp. 6895–6911. ISSN: 0556-2821, 1089-4918. DOI: 10.1103/PhysRevD.56.6895.

[114] D. Bigi, P. Gambino, and S. Schacht. “A fresh look at the determination of $|V_{cb}|$ from $B \to D^*\ell nu_\ell$”. In: *Physics Letters B* 769 (June 2017), pp. 441–445. ISSN: 03702693. DOI: 10.1016/j.physletb.2017.04.022. arXiv: 1703.06124.

[115] S. G. Krantz. *Handbook of Complex Variables*. Boston, MA: Birkhäuser Boston, 1999. ISBN: 978-1-4612-1588-2. DOI: 10.1007/978-1-4612-1588-2.

[116] I. Caprini, L. Lellouch, and M. Neubert. “Dispersive bounds on the shape of $\bar{B} \to D^{(*)}\ell\bar{\nu}_\ell$ form factors”. In: *Nucl. Phys.* B530 (1998), pp. 153–181. DOI: 10.1016/S0550-3213(98)00350-2.

[117] E. Eichten and B. Hill. “An Effective Field Theory for the Calculation of Matrix Elements Involving Heavy Quarks”. In: *Physics Letters B* 234.4 (Jan. 1990), pp. 511–516. ISSN: 03702693. DOI: 10.1016/0370-2693(90)92049-O.

[118] H. Georgi. “An Effective Field Theory for Heavy Quarks at Low Energies”. In: *Physics Letters B* 240.3-4 (Apr. 1990), pp. 447–450. ISSN: 03702693. DOI: 10.1016/0370-2693(90)91128-X.

[119] N. Isgur and M. B. Wise. “Weak Transition Form Factors between Heavy Mesons”. In: *Physics Letters B* 237.3-4 (Mar. 1990), pp. 527–530. ISSN: 03702693. DOI: 10.1016/0370-2693(90)91219-2.

[120] B. Grinstein and A. Kobach. “Model-independent extraction of $|V_{cb}|$ from $B \to D^*\ell\nu$”. In: *Physics Letters B* 771 (Aug. 2017), pp. 359–364. ISSN: 03702693. DOI: 10.1016/j.physletb.2017.05.078.

[121] The Belle Collaboration 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”. Feb. 14, 2017. arXiv: 1702.01521 [hep-ex]. URL: http://arxiv.org/abs/1702.01521 (visited on 03/17/2021).

[122] F. U. Bernlochner, Z. Ligeti, and D. J. Robinson. "$N = 5, 6, 7, 8$: Nested hypothesis tests and truncation dependence of $|V_{cb}|$". In: *Physical Review D* 100.1 (July 23, 2019), p. 013005. ISSN: 2470-0010, 2470-0029. DOI: 10.1103/PhysRevD.100.013005. arXiv: 1902.09553.

[123] M. Duraisamy, P. Sharma, and A. Datta. "Azimuthal $B \to D^* \tau^- \bar{\nu}_\tau$ angular distribution with tensor operators". In: *Physical Review D* 90.7 (Oct. 10, 2014), p. 074013. ISSN: 1550-7998, 1550-2368. DOI: 10.1103/PhysRevD.90.074013.

[124] K. Lieret. "Construction of Angular Observables Sensitive to New Physics in $B \to D^* \tau \nu$ Decays and Measurement of Differential Cross Sections of $B \to D^* \ell \nu$ decays with Hadronic Tagging at Belle". MA thesis. Munich: Ludwig Maximilian University of Munich, June 9, 2018. URL: https://docs.belle2.org/record/2532.

[125] J. Aebischer, T. Kuhr, and K. Lieret. "Clustering of $\overline{B} \to D^{(*)} \tau^- \overline{\nu}_\tau$ kinematic distributions with ClusterKinG". In: *Journal of High Energy Physics* 2020.4 (Apr. 2020), p. 7. ISSN: 1029-8479. DOI: 10.1007/JHEP04(2020)007.

[126] A. Carvalho, M. Dall'Osso, T. Dorigo, F. Goertz, C. A. Gottardo, and M. Tosi. "Higgs Pair Production: Choosing Benchmarks with Cluster Analysis". In: *JHEP* 04 (2016), p. 126. DOI: 10.1007/JHEP04(2016)126. arXiv: 1507.02245 [hep-ph].

[127] A. Carvalho, M. Dall'Osso, P. d. C. Manzano, T. Dorigo, F. Goertz, M. Gouzevich, and M. Tosi. "Analytical Parametrization and Shape Classification of Anomalous HH Production in the EFT Approach". July 19, 2016. arXiv: 1608.06578 [hep-ex, physics:hep-ph]. URL: http://arxiv.org/abs/1608.06578 (visited on 03/20/2022).

[128] A. Carvalho, F. Goertz, K. Mimasu, M. Gouzevitch, and A. Aggarwal. "On the Reinterpretation of Non-Resonant Searches for Higgs Boson Pairs". In: *Journal of High Energy Physics* 2021.2 (Feb. 2021), p. 49. ISSN: 1029-8479. DOI: 10.1007/JHEP02(2021)049.

[129] V. Khachatryan et al. "Search for Two Higgs Bosons in Final States Containing Two Photons and Two Bottom Quarks in Proton-Proton Collisions at 8 TeV". In: *Physical Review* D94.CMS-HIG-13-032, CERN-EP-2016-050 (5 2016), p. 052012. DOI: 10.1103/PhysRevD.94.052012. arXiv: 1603.06896 [hep-ex].

[130] B. Grzadkowski, M. Iskrzynski, M. Misiak, and J. Rosiek. "Dimension-Six Terms in the Standard Model Lagrangian". In: *JHEP* 10.IFT-9-2010, TTP10-35 (2010), p. 085. DOI: 10.1007/JHEP10(2010)085. arXiv: 1008.4884 [hep-ph].

[131] E. E. Jenkins, A. V. Manohar, and M. Trott. "Renormalization Group Evolution of the Standard Model Dimension Six Operators I: Formalism and Lambda Dependence". In: *JHEP* 10 (2013), p. 087. DOI: 10.1007/JHEP10(2013)087. arXiv: 1308.2627 [hep-ph].

[132] E. E. Jenkins, A. V. Manohar, and M. Trott. "Renormalization Group Evolution of the Standard Model Dimension Six Operators II: Yukawa Dependence". In: *JHEP* 01 (2014), p. 035. DOI: 10.1007/JHEP01(2014)035. arXiv: 1310.4838 [hep-ph].

[133] R. Alonso, E. E. Jenkins, A. V. Manohar, and M. Trott. "Renormalization Group Evolution of the Standard Model Dimension Six Operators III: Gauge Coupling Dependence and Phenomenology". In: *JHEP* 04.CERN-PH-TH-2013-305 (2014), p. 159. DOI: 10.1007/JHEP04(2014)159. arXiv: 1312.2014 [hep-ph].

[134] A. Celis, J. Fuentes-Martín, A. Vicente, and J. Virto. "DsixTools: The Standard Model Effective Field Theory Toolkit". In: *The European Physical Journal C* 77.6 (June 2017), p. 405. ISSN: 1434-6044, 1434-6052. DOI: 10.1140/epjc/s10052-017-4967-6.

[135] J. Aebischer, A. Crivellin, M. Fael, and C. Greub. "Matching of gauge invariant dimension-six operators for $b \to s$ and $b \to c$ transitions". In: *JHEP* 05.CERN-PH-TH-2015-278 (2016), p. 037. DOI: 10.1007/JHEP05(2016)037. arXiv: 1512.02830 [hep-ph].

[136] E. E. Jenkins, A. V. Manohar, and P. Stoffer. "Low-Energy Effective Field Theory below the Electroweak Scale: Operators and Matching". In: *JHEP* 03 (2018), p. 016. DOI: 10.1007/JHEP03(2018)016. arXiv: 1709.04486 [hep-ph].

[137] W. Dekens and P. Stoffer. "Low-Energy Effective Field Theory below the Electroweak Scale: Matching at One Loop". In: *JHEP* 10 (2019), p. 197. DOI: 10.1007/JHEP10(2019)197. arXiv: 1908.05295 [hep-ph].

[138] J. Aebischer, M. Fael, C. Greub, and J. Virto. "B Physics beyond the Standard Model at One Loop: Complete Renormalization Group Evolution below the Electroweak Scale". In: *JHEP* 09 (2017), p. 158. DOI: 10.1007/JHEP09(2017)158. arXiv: 1704.06639 [hep-ph].

[139] E. E. Jenkins, A. V. Manohar, and P. Stoffer. "Low-Energy Effective Field Theory below the Electroweak Scale: Anomalous Dimensions". In: *JHEP* 01 (2018), p. 084. DOI: 10.1007/JHEP01(2018)084. arXiv: 1711.05270 [hep-ph].

[140] J. Fuentes-Martín, P. Ruiz-Femenía, A. Vicente, and J. Virto. "DsixTools 2.0: The Effective Field Theory Toolkit". In: *The European Physical Journal C* 81.2 (Feb. 2021), p. 167. ISSN: 1434-6044, 1434-6052. DOI: 10.1140/epjc/s10052-020-08778-y.

[141] J. C. Criado. "MatchingTools: A Python Library for Symbolic Effective Field Theory Calculations". In: *Computer Physics Communications* 227 (2018), pp. 42–50. DOI: 10.1016/j.cpc.2018.02.016. arXiv: 1710.06445 [hep-ph].

[142] J. Aebischer, J. Kumar, and D. M. Straub. "Wilson: A Python Package for the Running and Matching of Wilson Coefficients above and below the Electroweak Scale". In: *The European Physical Journal C* 78.12 (Dec. 2018), p. 1026. ISSN: 1434-6044, 1434-6052. DOI: 10.1140/epjc/s10052-018-6492-7.

[143] J. Aebischer et al. "WCxf: An Exchange Format for Wilson Coefficients beyond the Standard Model". In: *Computer Physics Communications* 232 (2018), pp. 71–83. DOI: 10.1016/j.cpc.2018.05.022. arXiv: 1712.05298 [hep-ph].

[144] J. Aebischer, J. Kumar, P. Stangl, and D. M. Straub. "A Global Likelihood for Precision Constraints and Flavour Anomalies". In: *The European Physical Journal C* 79.6 (June 2019), p. 509. ISSN: 1434-6044, 1434-6052. DOI: 10.1140/epjc/s10052-019-6977-z.

[145] J. C. Criado. "BasisGen: Automatic Generation of Operator Bases". In: *The European Physical Journal C* 79.3 (Mar. 2019), p. 256. ISSN: 1434-6044, 1434-6052. DOI: 10.1140/epjc/s10052-019-6769-5.

[146] B. Gripaios and D. Sutherland. "DEFT: A Program for Operators in EFT". In: *JHEP* 01 (2019), p. 128. DOI: 10.1007/JHEP01(2019)128. arXiv: 1807.07546 [hep-ph].

[147] I. Brivio et al. "Computing Tools for the SMEFT". In: *Computing Tools for the SMEFT*. Ed. by J. Aebischer, M. Fael, A. Lenz, M. Spannowsky, and J. Virto. 2019. arXiv: 1910.11003 [hep-ph].

[148] D. M. Straub. "Flavio: A Python Package for Flavour and Precision Phenomenology in the Standard Model and Beyond". Oct. 18, 2018. arXiv: 1810.08132 [hep-ex, physics:hep-ph]. URL: http://arxiv.org/abs/1810.08132 (visited on 03/19/2022).

[149] W. Porod, F. Staub, and A. Vicente. "A Flavor Kit for BSM Models". In: *The European Physical Journal C* 74.8 (Aug. 2014), p. 2992. ISSN: 1434-6044, 1434-6052. DOI: 10.1140/epjc/s10052-014-2992-2.

[150] F. Mahmoudi. "SuperIso: A Program for calculating the isospin asymmetry of $B \to K^* \gamma$ in the MSSM". In: *Computer Physics Communications* 178 (2008), pp. 745–754. DOI: 10.1016/j.cpc.2007.12.006. arXiv: 0710.2067 [hep-ph].

[151] W. Porod. "SPheno, a program for calculating supersymmetric spectra, SUSY particle decays and SUSY particle production at e^+e^- colliders". In: *Computer Physics Communications* 153.ZU-TH-01-03 (2003), pp. 275–315. DOI: 10.1016/S0010-4655(03)00222-4. arXiv: hep-ph/0301101.

[152] W. Porod and F. Staub. "SPheno 3.1: Extensions Including Flavour, CP-phases and Models beyond the MSSM". In: *Computer Physics Communications* 183 (2012), pp. 2458–2469. DOI: 10.1016/j.cpc.2012.05.021. arXiv: 1104.1573 [hep-ph].

[153] J. A. Evans and D. Shih. "FormFlavor Manual". May 31, 2016. arXiv: 1606.00003 [hep-ph]. URL: http://arxiv.org/abs/1606.00003 (visited on 03/19/2022).

[154] A. Dedes, M. Paraskevas, J. Rosiek, K. Suxho, and L. Trifyllis. "SmeftFR – Feynman Rules Generator for the Standard Model Effective Field Theory". In: *Computer Physics Communications* 247 (Feb. 2020), p. 106931. ISSN: 00104655. DOI: 10.1016/j.cpc.2019.106931.

[155] I. Brivio, Y. Jiang, and M. Trott. "The SMEFTsim Package, Theory and Tools". In: *JHEP* 12 (2017), p. 070. DOI: 10.1007/JHEP12(2017)070. arXiv: 1709.06492 [hep-ph].

[156] W. Altmannshofer, P. S. Bhupal Dev, and A. Soni. "$R(D^{(*)})$ anomaly: A possible hint for natural supersymmetry with R-parity violation". In: *Physical Review* D96.9 (2017), p. 095010. DOI: 10.1103/PhysRevD.96.095010. arXiv: 1704.06659 [hep-ph].

[157] C. Murgui, A. Peñuelas, M. Jung, and A. Pich. "Global fit to $b \to c\tau\nu$ transitions". In: *Journal of High Energy Physics* 2019.9 (Sept. 2019), p. 103. ISSN: 1029-8479. DOI: 10.1007/JHEP09(2019)103.

[158] R.-X. Shi, L.-S. Geng, B. Grinstein, S. Jäger, and J. M. Camalich. "Revisiting the new-physics interpretation of the $b \to c\tau\nu_\tau$ data". In: *Journal of High Energy Physics* 2019.12 (Dec. 2019), p. 65. ISSN: 1029-8479. DOI: 10.1007/JHEP12(2019)065.

[159] D. Becirevic, M. Fedele, I. Nisandzic, and A. Tayduganov. "Lepton Flavor Universality tests through angular observables of $\overline{B} \to D^{(*)}\ell\overline{\nu}$ decay modes". July 13, 2019. arXiv: 1907.02257 [hep-ex, physics:hep-ph]. URL: http://arxiv.org/abs/1907.02257 (visited on 03/19/2022).

[160] J. D. Gómez, N. Quintero, and E. Rojas. "Charged current $b \to c\tau\bar{\nu}_\tau$ anomalies in a general W' boson scenario". In: *Physical Review* D100.9 (2019), p. 093003. DOI: 10.1103/PhysRevD.100.093003. arXiv: 1907.08357 [hep-ph].

[161] M. Algueró, B. Capdevila, A. Crivellin, S. Descotes-Genon, P. Masjuan, J. Matias, and J. Virto. "Emerging Patterns of New Physics with and without Lepton Flavour Universal Contributions". In: *The European Physical Journal C* 79.8 (Aug. 2019), p. 714. ISSN: 1434-6044, 1434-6052. DOI: 10.1140/epjc/s10052-019-7216-3.

[162] M. Ciuchini, A. M. Coutinho, M. Fedele, E. Franco, A. Paul, L. Silvestrini, and M. Valli. "New physics in $b \to s\ell\ell$ confronts new data on lepton universality". In: *The European Physical Journal C* 79.8 (Aug. 2019), p. 719. ISSN: 1434-6044, 1434-6052. DOI: 10.1140/epjc/s10052-019-7210-9.

[163] A. Datta, J. Kumar, and D. London. "The B anomalies and new physics in $b \to se^+e^-$". In: *Physics Letters* B797.UdeM-GPP-TH-19-270 (2019), p. 134858. DOI: 10.1016/j.physletb.2019.134858. arXiv: 1903.10086 [hep-ph].

[164] J. Aebischer, W. Altmannshofer, D. Guadagnoli, M. Reboud, P. Stangl, and D. M. Straub. "B-decay discrepancies after Moriond 2019". In: *The European Physical Journal C* 80.3 (Mar. 2020), p. 252. ISSN: 1434-6044, 1434-6052. DOI: 10.1140/epjc/s10052-020-7817-x.

[165] F. U. Bernlochner, S. Duell, Z. Ligeti, M. Papucci, and D. J. Robinson. "Das Ist Der HAMMER: Consistent New Physics Interpretations of Semileptonic Decays". In: *The European Physical Journal C* 80.9 (Sept. 2020), p. 883. ISSN: 1434-6044, 1434-6052. DOI: 10.1140/epjc/s10052-020-8304-0.

[166] S. Duell, F. Bernlochner, Z. Ligeti, M. Papucci, and D. Robinson. "HAMMER: Reweighting Tool for Simulated Data Samples". In: *Proceedings, 38th International Conference on High Energy Physics (ICHEP 2016): Chicago, IL, USA, August 3-10, 2016*. Vol. ICHEP2016. PoS. 2017, p. 1074. DOI: 10.22323/1.282.1074.

[167] F. J. Massey. "The Kolmogorov-Smirnov Test for Goodness of Fit". In: *Journal of the American Statistical Association* 46.253 (Mar. 1951), pp. 68–78. ISSN: 0162-1459, 1537-274X. DOI: 10.1080/01621459.1951.10500769.

[168] A. N. Pettitt. "A Two-Sample Anderson-Darling Rank Statistic". In: *Biometrika* 63.1 (Apr. 1976), pp. 161–168. ISSN: 0006-3444. DOI: 10.1093/biomet/63.1.161. eprint: http://oup.prod.sis.lan/biomet/article-pdf/63/1/161/6689229/63-1-161.pdf.

[169] S. S. Wilks. "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses". In: *The Annals of Mathematical Statistics* 9.1 (Mar. 1938), pp. 60–62. ISSN: 0003-4851. DOI: 10.1214/aoms/1177732360.

[170] J. A. Hartigan. *Clustering Algorithms*. Wiley Series in Probability and Mathematical Statistics. New York, NY: Wiley, 1975. URL: http://cds.cern.ch/record/105051.

[171] J. MacQueen. "Some Methods for Classification and Analysis of Multivariate Observations". In: *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics*. Berkeley, Calif.: University of California Press, 1967, pp. 281–297. URL: https://projecteuclid.org/euclid.bsmsp/1200512992.

[172] J. E. Gentle, L. Kaufman, and P. J. Rousseuw. "Finding Groups in Data: An Introduction to Cluster Analysis." In: *Biometrics* 47.2 (June 1991), p. 788. ISSN: 0006341X. DOI: 10.2307/2532178. JSTOR: 2532178.

[173] K. Lieret and A. Celis. *ClusterKinG: Clustering of Kinematic Graphs*. Jan. 3, 2022. URL: https://github.com/clusterking/clusterking (visited on 03/18/2022).

[174] K. Lieret. *Physics Distributions for the ClusterKinG Package*. Jan. 3, 2022. URL: https://github.com/clusterking/clusterking_physics (visited on 03/18/2022).

[175] W. McKinney. "Data Structures for Statistical Computing in Python". In: Python in Science Conference. Austin, Texas, 2010, pp. 56–61. DOI: 10.25080/Majora-92bf1922-00a.

[176] E. Jones, T. Oliphant, P. Peterson, et al. *SciPy: Open Source Scientific Tools for Python*. 2001. URL: http://www.scipy.org/.

[177] K. Lieret. *ClusterKinG Documentation*. URL: https://clusterking.readthedocs.io/en/latest/ (visited on 03/27/2022).

[178] A. Celis, M. Jung, X.-Q. Li, and A. Pich. "Scalar contributions to $b \to c(u)\tau\nu_\tau$ transitions". In: *Physics Letters* B771 (2017), pp. 168–179. DOI: 10.1016/j.physletb.2017.05.037. arXiv: 1612.07757 [hep-ph].

[179] M. Blanke, A. Crivellin, S. de Boer, T. Kitahara, M. Moscati, U. Nierste, and I. Nišandžić. "Impact of polarization observables and $B_c \to \tau\nu_\tau$ on new physics explanations of the $b \to c\tau\nu$ anomaly". In: *Physical Review* D99.PSI-PR-18-16; TTP-18-42, PSI-PR–18–16, TTP–18–42 (7 2019), p. 075006. DOI: 10.1103/PhysRevD.99.075006. arXiv: 1811.09603 [hep-ph].

[180] S. Bhattacharya, S. Nandi, and S. K. Patra. "$b \to c\tau\nu_\tau$ Decays: a catalogue to compare, constrain, and correlate new physics effects". In: *The European Physical Journal C* 79.3 (Mar. 2019), p. 268. ISSN: 1434-6044, 1434-6052. DOI: 10.1140/epjc/s10052-019-6767-7.

[181] D. Bečirević, I. Doršner, S. Fajfer, N. Košnik, D. A. Faroughy, and O. Sumensari. "Scalar Leptoquarks from Grand Unified Theories to Accommodate the B-physics Anomalies". In: *Physical Review* D98.LPT-Orsay-18-76, LPT-ORSAY-18-76 (5 2018), p. 055003. DOI: 10.1103/PhysRevD.98.055003. arXiv: 1806.05689 [hep-ph].

[182] B. Bhattacharya, A. Datta, S. Kamali, and D. London. "CP violation in $\overline{B}^0 \to D^{*+}\mu^-\overline{\nu}_\mu$". In: *JHEP* 05.UdeM-GPP-TH-19-269, MITP/19-016 (2019), p. 191. DOI: 10.1007/JHEP05(2019)191. arXiv: 1903.02567 [hep-ph].

[183] K. Adamczyk. "Semitauonic B Decays at Belle/Belle II". In: *10th International Workshop on the CKM Unitarity Triangle (CKM 2018) Heidelberg, Germany, September 17-21, 2018*. 2019. arXiv: 1901.06380 [hep-ex].

[184] M. Jung and D. M. Straub. "Constraining new physics in $b \to c\ell\nu$ transitions". In: *JHEP* 01 (2019), p. 009. DOI: 10.1007/JHEP01(2019)009. arXiv: 1801.01112 [hep-ph].

[185] R. Alonso, A. Kobach, and J. Martin Camalich. "New physics in the kinematic distributions of $\overline{B}\to D^{(*)}\tau^-(\to\ell^-\overline{\nu}_\ell\nu_\tau)\overline{\nu}_\tau$". In: *Physical Review* D94.9 (2016), p. 094021. DOI: 10.1103/PhysRevD.94.094021. arXiv: 1602.07671 [hep-ph].

[186] P. Zyla et al. "Review of Particle Physics". In: *PTEP* 2020.8 (2020), p. 083C01. DOI: 10.1093/ptep/ptaa104.

[187] 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.

[188] K. Abe et al. "Observation of Mixing-Induced CP Violation in the Neutral B Meson System". In: *Physical Review D* 66.3 (Aug. 29, 2002), p. 032007. ISSN: 0556-2821, 1089-4918. DOI: 10.1103/PhysRevD.66.032007.

[189] J. Brodzicka et al. "Physics Achievements from the Belle Experiment". In: *Progress of Theoretical and Experimental Physics* 2012.1 (Dec. 29, 2012), pp. 4D001–. ISSN: 2050-3911. DOI: 10.1093/ptep/pts072.

[190] S.-K. Choi et al. "Observation of a Resonancelike Structure in the $\pi^\pm\psi'$ Mass Distribution in Exclusive $B \to K\pi^\pm\psi'$ Decays". In: *Physical Review Letters* 100.14 (Apr. 8, 2008), p. 142001. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.100.142001.

[191] D. Besson and T. Skwarnicki. "Upsilon Spectroscopy: Transitions in the Bottomonium System". In: *Annual Review of Nuclear and Particle Science* 43.1 (Dec. 1993), pp. 333–378. ISSN: 0163-8998, 1545-4134. DOI: 10.1146/annurev.ns.43.120193.002001.

[192] J. Hasenbusch. "Analysis of inclusive semileptonic B meson decays with tau lepton final states at the Belle experiment". Bonn: Rheinische Friedrich-Wilhelms-Universität Bonn, Dec. 19, 2017. URL: https://hdl.handle.net/20.500.11811/7578.

[193] T. Abe et al. "Belle II Technical Design Report". Nov. 1, 2010. arXiv: 1011.0352 [hep-ex, physics:physics]. URL: http://arxiv.org/abs/1011.0352 (visited on 03/19/2022).

[194] K. Akai, K. Furukawa, and H. Koiso. "SuperKEKB Collider". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 907 (Nov. 2018), pp. 188–199. ISSN: 01689002. DOI: 10.1016/j.nima.2018.08.017.

[195] Y. Ohnishi et al. "Accelerator Design at SuperKEKB". In: *Progress of Theoretical and Experimental Physics* 2013.3 (Mar. 26, 2013), 3A011–. ISSN: 2050-3911. DOI: 10.1093/ptep/pts083.

[196] E. Kou et al. "The Belle II Physics Book". In: *Progress of Theoretical and Experimental Physics* 2019.12 (Dec. 1, 2019), p. 123C01. ISSN: 2050-3911. DOI: 10.1093/ptep/ptz106.

[197] T. Kuhr, C. Pulvermacher, M. Ritter, T. Hauth, and N. Braun. "The Belle II Core Software". In: *Computing and Software for Big Science* 3.1 (Dec. 2019), p. 1. ISSN: 2510-2036, 2510-2044. DOI: 10.1007/s41781-018-0017-9. arXiv: 1809.04299.

[198] T. Hauth. "Belle II Track Reconstruction and Results from First Collisions". In: *EPJ Web of Conferences* 214 (2019). Ed. by A. Forti, L. Betev, M. Litmaath, O. Smirnova, and P. Hristov, p. 02032. ISSN: 2100-014X. DOI: 10.1051/epjconf/201921402032.

[199] *Belle II Luminosity - Belle II - DESY Confluence*. URL: https://confluence.desy.de/display/BI/Belle+II+Luminosity (visited on 03/01/2022).

[200] I. Abe et al. “The KEKB Injector Linac”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 499.1 (Feb. 2003), pp. 167–190. ISSN: 01689002. DOI: 10.1016/S0168-9002(02)01787-4.

[201] S. Kurokawa and E. Kikutani. “Overview of the KEKB Accelerators”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 499.1 (Feb. 2003), pp. 1–7. ISSN: 01689002. DOI: 10.1016/S0168-9002(02)01771-0.

[202] Y. Funakoshi and M. Tanaka. “Operation Scheme and Statistics of KEKB”. In: IPAC 2011. Ed. by C. Petit-Jean-Genaz. 2011. URL: https://accelconf.web.cern.ch/ipac2011/papers/thpz022.pdf.

[203] M. Chalmers. *KEK Reclaims Luminosity Record*. CERN Courier. June 30, 2020. URL: https://cerncourier.com/a/kek-reclaims-luminosity-record/ (visited on 11/13/2021).

[204] M. Chalmers. *SuperKEKB Raises the Bar*. CERN Courier. Aug. 22, 2021. URL: https://cerncourier.com/a/superkekb-raises-the-bar/ (visited on 11/13/2021).

[205] A. Abashian et al. “Muon Identification in the Belle Experiment at KEKB”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 491.1-2 (Sept. 2002), pp. 69–82. ISSN: 01689002. DOI: 10.1016/S0168-9002(02)01164-6.

[206] A. Abashian et al. “The Belle Detector”. In: *Nucl. Instrum. Meth.* A479 (2002), pp. 117–232. DOI: 10.1016/S0168-9002(01)02013-7.

[207] R. Abe et al. “The New Beampipe for the Belle Experiment”. In: *Instrumentation. Proceedings, 10th International Conference, VCI 10, Vienna, Austria, February 16-21, 2004* A535 (2004), pp. 558–561. DOI: 10.1016/j.nima.2004.07.260.

[208] Z. Natkaniec et al. “Status of the Belle Silicon Vertex Detector”. In: *Proceedings, 13th International Workshop on Vertex Detectors (Vertex 2004): Menaggio, Como, Italy, September 13-18, 2004* A560 (2006), pp. 1–4. DOI: 10.1016/j.nima.2005.11.228.

[209] Y. Ushiroda. “Belle Silicon Vertex Detectors”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 511.1-2 (Sept. 2003), pp. 6–10. ISSN: 01689002. DOI: 10.1016/S0168-9002(03)01739-X.

[210] R. Wigmans. “Calorimeters”. In: *Handbook of Particle Detection and Imaging*. Ed. by C. Grupen and I. Buvat. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 497–517. ISBN: 978-3-642-13270-4. DOI: 10.1007/978-3-642-13271-1_20.

[211] M. Krammer and W. Mitaroff. “Tracking Detectors”. In: *Handbook of Particle Detection and Imaging*. Ed. by C. Grupen and I. Buvat. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 265–295. ISBN: 978-3-642-13270-4. DOI: 10.1007/978-3-642-13271-1_12.

[212] O. Frost. “A Local Tracking Algorithm for the Central Drift Chamber of Belle II”. Diploma Thesis. Karlsruhe: Karlsruhe Institute of Technology (KIT), 2013. URL: https://publish.etp.kit.edu/record/21906.

[213] M. Akatsu et al. “Cathode Image Readout in the BELLE Central Drift Chamber”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 454.2-3 (Nov. 2000), pp. 322–333. ISSN: 01689002. DOI: 10.1016/S0168-9002(00)00487-3.

[214] H. Hirano et al. “A High-Resolution Cylindrical Drift Chamber for the KEK B-factory”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 455.2 (2000), pp. 294–304. ISSN: 0168-9002. DOI: 10.1016/S0168-9002(00)00513-1.

[215] S. Uno et al. “Study of a drift chamber filled with a helium-ethane mixture”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 330.1 (1993), pp. 55–63. ISSN: 0168-9002. DOI: 10.1016/0168-9002(93)91304-6.

[216] J. Engelfried. “Particle Identification”. In: *Handbook of Particle Detection and Imaging*. Ed. by C. Grupen and I. Buvat. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 125–137. ISBN: 978-3-642-13270-4. DOI: 10.1007/978-3-642-13271-1_6.

[217] K. Emi et al. “Study of a Measurement and the Gas-Gain Saturation by a Prototype Drift Chamber for the BELLE-CDC”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 379.2 (Sept. 1996), pp. 225–231. ISSN: 01689002. DOI: 10.1016/0168-9002(96)00716-4.

[218] M. Tomoto et al. “Z-Trigger System with the BELLE Central Drift Chamber”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 447.3 (June 2000), pp. 416–423. ISSN: 01689002. DOI: 10.1016/S0168-9002(99)01075-X.

[219] B. Ratcliff and J. Schwiening. “Cherenkov Counters”. In: *Handbook of Particle Detection and Imaging*. Ed. by C. Grupen and I. Buvat. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 453–471. ISBN: 978-3-642-13270-4. DOI: 10.1007/978-3-642-13271-1_18.

[220] T. Iijima et al. “Aerogel Cherenkov Counter for the BELLE Detector”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 453.1-2 (Oct. 2000), pp. 321–325. ISSN: 01689002. DOI: 10.1016/S0168-9002(00)00652-5.

[221] H. Kichimi et al. “The BELLE TOF System”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 453.1-2 (Oct. 2000), pp. 315–320. ISSN: 01689002. DOI: 10.1016/S0168-9002(00)00651-3.

[222] H. Kichimi and G. Varner. *Design of the TOF Front End Electronics*. Belle Note 138 (internal). Sept. 18, 1996. URL: https://belle.kek.jp/secured/belle_note/gn138/btof_elec_note_vsn2.ps.gz.

[223] T. Hebbeker and K. Hoepfner. “Muon Spectrometers”. In: *Handbook of Particle Detection and Imaging*. Ed. by C. Grupen and I. Buvat. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 473–496. ISBN: 978-3-642-13270-4. DOI: 10.1007/978-3-642-13271-1_19.

[224] A. Abashian et al. “The K_L/μ detector subsystem for the BELLE experiment at the KEK B-factory”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 449.1-2 (July 2000), pp. 112–124. ISSN: 01689002. DOI: 10.1016/S0168-9002(99)01383-2.

[225] K. Hanagaki, H. Kakuno, H. Ikeda, T. Iijima, and T. Tsukamoto. “Electron Identification in Belle”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 485.3 (June 2002), pp. 490–503. ISSN: 01689002. DOI: 10.1016/S0168-9002(01)02113-1.

[226] M. Barrett, T. Hara, M. Hern, D. Kalita, P. Kettunen, and P. Shingade. “Belle II RAW Data Management: The Online-Offline Data Transfer System”. ICHEP 2020 (Prague). July 30, 2020. URL: https://indico.cern.ch/event/868940/contributions/3814705/attachments/2082197/3497494/.

[227] *HEP Software Foundation: Training Working Group*. URL: https://hepsoftwarefoundation.org/workinggroups/training.html (visited on 01/04/2022).

[228] *HEP Software Foundation*. URL: https://hepsoftwarefoundation.org/ (visited on 02/10/2022).

[229] M. Ritter, T. Kuhr, M. Starič, and Belle II Software Group. "High Level Interface to Conditions Data at Belle II". In: *Journal of Physics: Conference Series* 898 (Oct. 2017), p. 042046. ISSN: 1742-6588, 1742-6596. DOI: 10.1088/1742-6596/898/4/042046.

[230] L. Wood, T. Elsethagen, M. Schram, and E. Stephan. "Conditions Database for the Belle II Experiment". In: *Journal of Physics: Conference Series* 898 (Oct. 2017), p. 042060. ISSN: 1742-6588, 1742-6596. DOI: 10.1088/1742-6596/898/4/042060.

[231] *Documentation of the Belle II Software*. URL: https://software.belle2.org/ (visited on 03/25/2022).

[232] *Belle II Collaboration: GitHub Organization*. GitHub. URL: https://github.com/belle2 (visited on 03/25/2022).

[233] C. Pulvermacher. "Analysis Software and Full Event Interpretation for the Belle II Experiment". In: (2015). DOI: 10.5445/IR/1000050704.

[234] D. Liventsev. "Study of $B \to D^{**}\ell\nu$ with full reconstruction tagging". In: *Physical Review D* 77.9 (May 12, 2008), p. 091503. ISSN: 1550-7998, 1550-2368. DOI: 10.1103/PhysRevD.77.091503. arXiv: 0711.3252.

[235] Belle Collaboration and I. Bizjak. "Determination of $|V_{ub}|$ from Measurements of the Inclusive Charmless Semileptonic Partial Rates of B Mesons using Full Reconstruction Tags". In: *Physical Review Letters* 95.24 (Dec. 8, 2005), p. 241801. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.95.241801. arXiv: hep-ex/0505088.

[236] M. Feindt. "A Neural Bayesian Estimator for Conditional Probability Densities". Feb. 18, 2004. arXiv: physics/0402093. URL: http://arxiv.org/abs/physics/0402093 (visited on 03/15/2022).

[237] M. Feindt and U. Kerzel. "The NeuroBayes Neural Network Package". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 559.1 (Apr. 2006), pp. 190–194. ISSN: 01689002. DOI: 10.1016/j.nima.2005.11.166.

[238] 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: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 654.1 (Oct. 2011), pp. 432–440. ISSN: 01689002. DOI: 10.1016/j.nima.2011.06.008.

[239] T. Keck et al. "The Full Event Interpretation: An Exclusive Tagging Algorithm for the Belle II Experiment". In: *Computing and Software for Big Science* 3.1 (Dec. 2019), p. 6. ISSN: 2510-2036, 2510-2044. DOI: 10.1007/s41781-019-0021-8.

[240] T. Keck. "Machine Learning Algorithms for the Belle II Experiment and Their Validation on Belle Data". Karlsruhe Institute of Technology (KIT), 2017. URL: https://publish.etp.kit.edu/record/21404.

[241] T. Keck. "FastBDT: A Speed-Optimized Multivariate Classification Algorithm for the Belle II Experiment". In: *Computing and Software for Big Science* 1.1 (Nov. 2017), p. 2. ISSN: 2510-2036, 2510-2044. DOI: 10.1007/s41781-017-0002-8.

[242] T. Keck. *FastBDT*. Mar. 4, 2022. URL: https://github.com/thomaskeck/FastBDT (visited on 03/16/2022).

[243] M. Gelb et al. "B2BII: Data Conversion from Belle to Belle II". In: *Computing and Software for Big Science* 2.1 (Nov. 2018), p. 9. ISSN: 2510-2036, 2510-2044. DOI: 10.1007/s41781-018-0016-x.

[244] *CherryPy — A Minimalist Python Web Framework*. URL: https://cherrypy.dev/ (visited on 03/26/2022).

[245] *Buildbot: The Continuous Integration Framework*. URL: https://buildbot.net/ (visited on 03/26/2022).

[246] R. Itoh. "BASF - BELLE AnalysiS Framework". In: 9th International Conference on Computing in High-Energy Physics (CHEP 1997). 1997. URL: https://www.ifh.de/CHEP97/paper/244.ps.

[247] K. Lieret. "A New Software Training Model at Belle II". ACAT 2021 (Daejeon, South Korea). Dec. 2021. URL: https://indico.cern.ch/event/855454/contributions/4605008/ (visited on 02/08/2022).

[248] K. Lieret et al. *A New Software Training Model at Belle II*. Feb. 2022. URL: https://docs.belle2.org/record/2946.

[249] *The Carpentries*. The Carpentries. URL: https://carpentries.org/index.html (visited on 02/10/2022).

[250] *Software Carpentry*. Software Carpentry. URL: http://software-carpentry.org/ (visited on 01/04/2022).

[251] E. Becker and F. Michonneau. *The Carpentries Curriculum Development Handbook*. URL: https://carpentries.github.io/curriculum-development/ (visited on 01/04/2022).

[252] G. Wilson. "Software Carpentry: Lessons Learned". In: *F1000Research* 3 (Jan. 28, 2016), p. 62. ISSN: 2046-1402. DOI: 10.12688/f1000research.3-62.v2.

[253] S. Malik et al. "Software Training in HEP". In: *Computing and Software for Big Science* 5.1 (Dec. 2021), p. 22. ISSN: 2510-2036, 2510-2044. DOI: 10.1007/s41781-021-00069-9. arXiv: 2103.00659.

[254] *LHCb Starterkit*. URL: https://lhcb.github.io/starterkit/ (visited on 01/04/2022).

[255] G. Brandl. *Sphinx: Python Documentation Generator*. URL: https://www.sphinx-doc.org/ (visited on 01/04/2022).

[256] *reStructuredText*. Oct. 22, 2021. URL: https://docutils.sourceforge.io/rst.html (visited on 02/07/2022).

[257] *Rebrandly: Custom URL Shortener, Branded Link Management, API*. URL: https://www.rebrandly.com/ (visited on 02/10/2022).

[258] K. Lieret. *Training & Documentation Hackathon*. URL: https://indico.belle2.org/event/3796/ (visited on 02/10/2022).

[259] Belle II Documentation, Training and Software Outreach group and K. Lieret. *6th Belle II Starter Kit Workshop*. URL: https://indico.belle2.org/event/2853/ (visited on 02/09/2022).

[260] Belle II Documentation, Training and Software Outreach group and K. Lieret. *7th Belle II Starter Kit Workshop*. URL: https://indico.belle2.org/event/3795/ (visited on 02/09/2022).

[261] *Indico: Robust and Scalable Event Management*. URL: https://getindico.io/ (visited on 02/09/2022).

[262] Atlassian. *Confluence: Your Remote-Friendly Team Workspace*. URL: https://www.atlassian.com/software/confluence (visited on 02/09/2022).

[263] *Gather: A Better Way to Meet Online*. URL: https://www.gather.town/ (visited on 02/09/2022).

[264] *Gatheround: Crack the Culture Code*. URL: https://gatheround.com/ (visited on 02/09/2022).

[265] Belle II Software Training and Documentation group. *Belle 2 Training Material*. URL: training.belle2.org/.

[266] F. Gaede. "HEP Software Foundation and White Paper". 10th Annual Meeting of the Helmholtz Alliance "Physics at the Terascale". URL: https://indico.desy.de/event/15730/contributions/19906/ (visited on 02/10/2022).

[267] J. Albrecht et al. "A Roadmap for HEP Software and Computing R&D for the 2020s". In: *Computing and Software for Big Science* 3.1 (Dec. 2019), p. 7. ISSN: 2510-2036, 2510-2044. DOI: 10.1007/s41781-018-0018-8. arXiv: 1712.06982.

[268] HEP Software Foundation et al. "HEP Software Foundation Community White Paper Working Group - Training, Staffing and Careers". Jan. 17, 2019. arXiv: 1807.02875 [hep-ex, physics:physics]. URL: http://arxiv.org/abs/1807.02875 (visited on 02/10/2022).

[269] *HEP Software Foundation Training Material*. GitHub. URL: https://github.com/hsf-training (visited on 02/10/2022).

[270] S. Ponce et al. *C++ Course*. HEP Software Foundation Training Material, Jan. 19, 2022. URL: https://github.com/hsf-training/cpluspluscourse (visited on 02/10/2022).

[271] T. Tantau, V. Miletić, and J. Wright. *beamer – A LaTeX class for producing presentations and slides*. URL: https://ctan.org/pkg/beamer (visited on 02/10/2022).

[272] G. Fidalgo-Rodríguez. *Python-for-STEM-Teachers-Workshop*. HEP Software Foundation Training Material, Aug. 17, 2020. URL: https://github.com/hsf-training/hsf-k12-python-for-stem-teachers (visited on 02/10/2022).

[273] *Google Colaboratory*. URL: https://colab.research.google.com/ (visited on 02/10/2022).

[274] *GitHub Pages*. URL: https://pages.github.com/ (visited on 02/10/2022).

[275] *Jekyll: Simple, Blog-Aware, Static Sites*. URL: https://jekyllrb.com/ (visited on 02/10/2022).

[276] HSF Training WG. *Template and Theming for Lesson Modules of the HSF Training Center*. HEP Software Foundation Training Material, July 30, 2021. URL: https://github.com/hsf-training/hsf-styles (visited on 02/10/2022).

[277] The Carpentries. *Styles for The Carpentries Lessons*. The Carpentries, Feb. 10, 2022. URL: https://github.com/carpentries/styles (visited on 02/10/2022).

[278] J. Gruber and A. Swartz. *Daring Fireball: Markdown*. URL: https://daringfireball.net/projects/markdown/ (visited on 02/10/2022).

[279] *HSF Software Training Center*. URL: https://hepsoftwarefoundation.org/training/curriculum.html (visited on 02/10/2022).

[280] HSF Training WG. *So I Want to Host/Request/Organize a Training Event...* URL: https://hepsoftwarefoundation.org/training/howto-event.html (visited on 02/11/2022).

[281] HSF Training WG and IRIS-HEP. *Virtual Pipelines Training*. URL: https://indico.cern.ch/event/904759/ (visited on 02/10/2022).

[282] HSF Training WG and IRIS-HEP. *GitHub CI/CD Training*. URL: https://indico.cern.ch/event/1001128/ (visited on 02/10/2022).

[283] HSF Training WG and IRIS-HEP. *Virtual Docker Training*. URL: https://indico.cern.ch/event/934651/ (visited on 02/10/2022).

[284] HSF Training WG, IRIS-HEP, and University of Manchester. *ML + GPU Training*. URL: https://indico.cern.ch/event/958112/ (visited on 02/10/2022).

[285] HSF Training WG and SIDIS. *1st HEP C++ Course and Hands-on Training*. URL: https://indico.cern.ch/event/946584/ (visited on 02/10/2022).

[286] *Software Institute for Data-Intensive Sciences*. URL: https://sidis.web.cern.ch/ (visited on 02/10/2022).

[287] HSF Training WG and IRIS-HEP. *Data Analysis for STEM Teachers*. URL: https://indico.cern.ch/event/927162/ (visited on 02/10/2022).

[288] *Institute for Research and Innovation in Software for High Energy Physics*. URL: http://iris-hep.org/ (visited on 02/10/2022).

[289] *Framework for Integrated Research Software Training*. Framework for Integrated Research Software Training (FIRST-HEP). URL: https://first-hep.org/ (visited on 02/10/2022).

[290] K. Lieret. "Community Building". HSF WLCG Virtual Workshop. Nov. 19, 2020. URL: https://indico.cern.ch/event/941278/contributions/4084356/ (visited on 02/11/2022).

[291] *The HSF Training Community*. URL: https://hepsoftwarefoundation.org/training/community.html (visited on 02/11/2022).

[292] *Recognize All Contributors*. All Contributors. URL: https://allcontributors.org (visited on 02/11/2022).

[293] K. Lieret. *Software Paradigms and Programming Patterns: Material*. Version lmu-patterns-2021. Zenodo, Jan. 11, 2021. DOI: 10.5281/ZENODO.4432319.

[294] *Inverted CERN School of Computing 2020*. Indico. URL: https://indico.cern.ch/event/853710/ (visited on 03/27/2022).

[295] HSF Training WG and K. Lieret. *HSF Training Hackathon*. URL: https://indico.cern.ch/event/975487/ (visited on 02/10/2022).

[296] K. Lieret and H. T. WG. *HSF IRIS-HEP Training Challenge 2*. Indico. URL: https://indico.cern.ch/event/1088551/ (visited on 03/27/2022).

[297] S. Mönig/Falke, F. Bernlocher, J. Hasenbusch, M. Rüger, and J. Dingfelder. *Determination of the CKM matrix element $|V_{cb}|$ and search for new physics using semileptonic $B \to D^*\ell\nu$ decays with hadronic tagging*. Belle Note 1419 (Internal). 2016. URL: http://belle.kek.jp/secured/belle_note/gn1419/.

[298] C. Bobeth, M. Bordone, N. Gubernari, M. Jung, and D. van Dyk. "Lepton-flavour non-universality of $\bar{B} \to D^*\ell\bar{\nu}$ angular distributions in and beyond the Standard Model". Apr. 5, 2021. arXiv: 2104.02094 [hep-ex, physics:hep-ph]. URL: http://arxiv.org/abs/2104.02094 (visited on 06/14/2021).

[299] F. Bernlocher, T. Kuhr, K. Lieret, F. Metzner, and M. Prim. *V_{cb} with hadronically tagged $B \to D^*\ell\nu$ Decays*. Belle Note 1615 (Internal). URL: https://belle.kek.jp/secured/belle_note/gn1615/.

[300] S. Ryu. *Study of π^0 efficiency using $\tau^- \to \pi^-\pi^0\nu_\tau$*. Belle Note 1224 (internal). Apr. 20, 2012. URL: https://belle.kek.jp/secured/belle_note/gn1224/pi0eff.pdf.

[301] A. Höcker and V. Kartvelishvili. "SVD Approach to Data Unfolding". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 372.3 (Apr. 1996), pp. 469–481. ISSN: 01689002. DOI: 10.1016/0168-9002(95)01478-0.

[302] G. Cowan. "A Survey of Unfolding Methods for Particle Physics". In: *Advanced Statistical Techniques in Particle Physics*. Conference on Advanced Statistical Techniques in Particle Physics. Ed. by M. R. Whalley and L. Lyons. Vol. 0203181. Durham, UK, 2002, pp. 248–257.

[303] V. Blobel. "Unfolding Methods in Particle Physics". In: *PHYSTAT 2011*. Geneva: CERN, 2011, pp. 240–251. DOI: 10.5170/CERN-2011-006.240.

[304] R. Glattauer and C. Schwanda. *Measurement of the decay $B \to D\ell\nu$ in fully reconstructed events and determination of the CKM matrix element V_{cb}*. Belle Note 1345 (Internal). URL: https://belle.kek.jp/secured/belle_note/gn1345/BN1345_v7.pdf.

[305] F. Meier and A. Vossen. *Measurement of $B \to D^{**}\ell\nu_\ell$ branching fractions using Belle data and hadronic Full Event Interpretation*. Belle Note 1569. URL: https://belle.kek.jp/secured/belle_note/gn1569/BN1569_v2.2.pdf.

[306] M. Rüger. "Korrektur von b-tagging Effizienzen für eine $B \longrightarrow D^*\ell\nu$-Analyse am Belle Experiment". 2016.

[307] Belle II Collaboration et al. "A calibration of the Belle II hadronic tag-side reconstruction algorithm with $B \to X\ell\nu$ decays". Aug. 17, 2020. arXiv: 2008.06096 [hep-ex]. URL: http://arxiv.org/abs/2008.06096 (visited on 12/02/2021).

[308] J. Schwab. "Calibration of the Full Event Interpretation for the Belle and the Belle II Experiment". MA thesis. Karlsruhe Institute of Technology (KIT), 2017. URL: https://publish.etp.kit.edu/record/21422.

[309] D. J. Lange. "The EvtGen Particle Decay Simulation Package". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 462.1-2 (Apr. 2001), pp. 152–155. ISSN: 01689002. DOI: 10.1016/S0168-9002(01)00089-4.

[310] T. Sjöstrand, P. Edén, C. Friberg, L. Lönnblad, G. Miu, S. Mrenna, and E. Norrbin. "High-Energy-Physics Event Generation with Pythia 6.1". In: *Computer Physics Communications* 135.2 (Apr. 2001), pp. 238–259. ISSN: 00104655. DOI: 10.1016/S0010-4655(00)00236-8.

[311] E. Barberio and Z. Was. "PHOTOS - a Universal Monte Carlo for QED Radiative Corrections: Version 2.0". In: *Computer Physics Communications* 79.2 (Apr. 1994), pp. 291–308. ISSN: 00104655. DOI: 10.1016/0010-4655(94)90074-4.

[312] R. Brun, F. Bruyant, F. Carminati, S. Giani, M. Maire, A. McPherson, G. Patrick, and L. Urban. "GEANT: Detector Description and Simulation Tool". In: (2008). DOI: 10.17181/CERN.MUHF.DMJ1.

[313] R. Brun, F. Bruyant, M. Maire, A. C. McPherson, and P. Zanarini. *GEANT3*. 1987. URL: https://cds.cern.ch/record/1119728/export/hx?ln=en.

[314] F. U. Bernlochner, Z. Ligeti, and S. Turczyk. "A Proposal to Solve Some Puzzles in Semileptonic B Decays". In: *Physical Review D* 85.9 (May 30, 2012), p. 094033. ISSN: 1550-7998, 1550-2368. DOI: 10.1103/PhysRevD.85.094033. arXiv: 1202.1834.

[315] K. Schulten. "Notes on Quantum Mechanics". Apr. 18, 2000. URL: https://www.ks.uiuc.edu/Services/Class/PHYS480/qm_PDF/QM_Book.pdf (visited on 03/03/2022).

[316] D. H. Perkins. *Introduction to High Energy Physics*. 4th ed. Cambridge University Press, Apr. 13, 2000. ISBN: 978-0-511-80904-0. DOI: 10.1017/CBO9780511809040.

[317] J. Hartung, G. Knapp, and B. K. Sinha. *Statistical Meta-Analysis with Applications*. Wiley Series in Probability and Statistics. Hoboken, N.J: Wiley, 2008. 247 pp. ISBN: 978-0-470-29089-7.

[318] S. Duell and F. Bernlochner. *Correction of* BF($B \to X_c\ell\nu_\ell$) *in generic MC - Belle II - DESY Confluence*. URL: https://confluence.desy.de/pages/viewpage.action?pageId=202394372 (visited on 02/03/2022).

[319] LHCb Collaboration et al. "Measurements of the Branching fractions for $B_{(s)} \to D_{(s)}\pi\pi\pi$ and $\Lambda_b^0 \to \Lambda_c^+\pi\pi\pi$". In: *Physical Review D* 84.9 (Nov. 2, 2011), p. 092001. ISSN: 1550-7998, 1550-2368. DOI: 10.1103/PhysRevD.84.092001. arXiv: 1109.6831.

[320] The DELPHI Collaboration. "Determination of Heavy Quark Non-Perturbative Parametersfrom Spectral Moments in Semileptonic B Decays". In: *The European Physical Journal C* 45.1 (Jan. 2006), pp. 35–59. ISSN: 1434-6044, 1434-6052. DOI: 10.1140/epjc/s2005-02406-7.

[321] D. Liventsev et al. "Study of $B \to D^{**}\ell\nu_\ell$ with full reconstruction tagging". In: *Physical Review D* 77.9 (May 12, 2008), p. 091503. ISSN: 1550-7998, 1550-2368. DOI: 10.1103/PhysRevD.77.091503.

[322] B. Aubert et al. "Measurement of the Branching Fractions of $B^- \to D^{**}\ell^-\bar{\nu}_\ell$ Decays in Events Tagged by a Fully Reconstructed B Meson". In: *Physical Review Letters* 101.26 (Dec. 23, 2008), p. 261802. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.101.261802.

[323] J. Stypula et al. "Evidence for $B \to D_s^+ K^-\ell^-\bar{\nu}_\ell$ and search for $B \to D_s^{*+} K^-\ell^-\bar{\nu}_\ell$". In: *Physical Review D* 86.7 (Oct. 23, 2012), p. 072007. ISSN: 1550-7998, 1550-2368. DOI: 10.1103/PhysRevD.86.072007.

[324] T. B. Collaboration. "Observation of $\overline{B} \to D^{(*)}\pi^+\pi^-\ell^-\bar{\nu}$ decays in e^+e^- collisions at the $\Upsilon(4S)$ resonance". In: *Physical Review Letters* 116.4 (Jan. 26, 2016), p. 041801. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.116.041801. arXiv: 1507.08303.

[325] R. A. Briere and E. R. Oxford. *Update of Charm Decay Tables*. BELLE2-NOTE-PH-2020-008 (Internal). Mar. 2020. URL: https://docs.belle2.org/record/1889.

[326] L. Hinz. *Lepton ID Efficiency Correction and Systematic Error*. Belle Note (Internal) 954. July 2006, p. 28. URL: https://belle.kek.jp/secured/belle_note/gn954/.

[327] P. Urquijo. *Semi-Inclusive Semileptonic B Decays*. Belle Note (Internal) 1251. Aug. 31, 2012, p. 103. URL: https://belle.kek.jp/secured/belle_note/gn1251/bn1251_v1.4.pdf.

[328] M. Prim. *b2-hive/eFFORT v0.1.0*. Version v0.1.0. Zenodo, July 29, 2020. DOI: 10.5281/ZENODO.3965699.

[329] D. Scora and N. Isgur. "Semileptonic Meson Decays in the Quark Model: An Update". In: *Physical Review D* 52.5 (Sept. 1, 1995), pp. 2783–2812. ISSN: 0556-2821. DOI: 10.1103/PhysRevD.52.2783.

[330] A. K. Leibovich, Z. Ligeti, I. W. Stewart, and M. B. Wise. "Semileptonic B Decays to Excited Charmed Mesons". In: *Physical Review D* 57.1 (Jan. 1, 1998), pp. 308–330. ISSN: 0556-2821, 1089-4918. DOI: 10.1103/PhysRevD.57.308.

[331] M. Welsch. "Measurement of the branching fraction of $B \to D^{**}(\to D^{(*)}\pi)\ell\nu$ in the context of a $R(D^{(*)})$ analysis at the Belle experiment". Karlsruhe Institute of Technology (KIT), 2019. URL: https://publish.etp.kit.edu/record/21909.

[332] W. Dungel. *Systematic investigation of the reconstruction efficiency of low momentum $\pi^\pm$ and π^0*. Belle Note 1176 (Internal). Jan. 13, 2011. URL: https://belle.kek.jp/secured/belle_note/gn1176/bn1176.pdf.

[333] E. White. *Determination of K_S Efficiency and Systematic Uncertainty*. Belle Note 1207 (internal). July 19, 2011. URL: https://belle.kek.jp/secured/belle_note/gn1207/bn1207_v2.ps.

[334] S. Nishida. *Study of Kaon and Pion Identification Using Inclusive D^* samples*. Belle Note 779 (internal). Jan. 17, 2005.

[335] M. Staric. *Track Smearing in MC*. Sept. 19, 2012. URL: https://belle.kek.jp/secured/wiki/doku.php?id=physics:charm:tracksmearing (visited on 03/15/2022).

[336] M. Staric. "Track Smearing in MC". Belle Charm Group Meeting (Internal). Apr. 21, 2011. URL: https://kds.kek.jp/event/6971/contributions/142292/attachments/111232/131990/track_smear.pdf.

[337] H. Nakano, A. Ishikawa, and K. Sumisawa. *Ks Selection with NeuroBayes and nisKsFinder Class*. Belle Note 1253 (internal).

[338] M. Röhrken. "Time-Dependent CP Violation Measurements in Neutral b Meson to Double-Charm Decays at the Japanese Belle Experiment". Karlsruhe Institute of Technology (KIT), 2012. URL: https://publish.etp.kit.edu/record/20864.

[339] E. Farhi. "Quantum Chromodynamics Test for Jets". In: *Physical Review Letters* 39.25 (Dec. 19, 1977), pp. 1587–1588. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.39.1587.

[340] D. M. Asner et al. "Search for Exclusive Charmless Hadronic B Decays". In: *Physical Review D* 53.3 (Feb. 1, 1996), pp. 1039–1050. ISSN: 0556-2821, 1089-4918. DOI: 10.1103/PhysRevD.53.1039.

[341] G. C. Fox and S. Wolfram. "Event shapes in e^+e^- annihilation". In: *Nuclear Physics B* 149.3 (Mar. 1979), pp. 413–496. ISSN: 05503213. DOI: 10.1016/0550-3213(79)90003-8.

[342] G. C. Fox and S. Wolfram. "Observables for the Analysis of Event Shapes in e^+e^- Annihilation and Other Processes". In: *Physical Review Letters* 41.23 (Dec. 4, 1978), pp. 1581–1585. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.41.1581.

[343] F. Metzner and M. Welsch. *TemplateFitter*. Dec. 7, 2021. URL: https://github.com/FelixMetzner/TemplateFitter (visited on 04/03/2022).

[344] F. James and M. Roos. "Minuit - a System for Function Minimization and Analysis of the Parameter Errors and Correlations". In: *Computer Physics Communications* 10.6 (Dec. 1975), pp. 343–367. ISSN: 00104655. DOI: 10.1016/0010-4655(75)90039-9.

[345] B. Bhuyan. *High p_T Tracking Efficiency Using Partially Reconstructed D^* Decays*. Belle Note (Internal) 1165 v1. URL: https://belle.kek.jp/secured/belle_note/gn1165/BN1165_v1.pdf.

[346] J. Gratrex, M. Hopfer, and R. Zwicky. "Generalised helicity formalism, higher moments and the $B{\to}K_{J_K}(\to K\pi)\overline{\ell}_1\ell_2$ angular distributions". In: *Physical Review* D93.5 (2016), p. 054008. DOI: 10.1103/PhysRevD.93.054008. arXiv: 1506.03970 [hep-ph].

[347] D. Bečirević, S. Fajfer, I. Nišandžić, and A. Tayduganov. "Angular distributions of $\bar{B} \to D^{(*)}\ell\bar{\nu}_\ell$ decays and search of new physics". In: *Nuclear Physics B* 946 (Sept. 2019), p. 114707. ISSN: 05503213. DOI: 10.1016/j.nuclphysb.2019.114707.

Thank you!

This work would not have been possible without the support of so many of you.

I would like to express my gratitude to Prof. Dr. Thomas Kuhr for his continuous support over the last four years. Thank you for introducing me to the exciting field of flavor physics and letting me explore so many different topics at once! Knowing that your door is always open has been a great comfort.

Prof. Dr. Florian Bernlochner, thank you for sharing your vast physics knowledge and intuition. It has been inspiring to see you working on both theoretical and experimental frontiers and combining both disciplines.

I would like to thank Prof. Dr. Andreas Burkert, Prof. Dr. Wolfgang Dünnweber, Prof. Dr. Otmar Biebel, and Prof. Dr. Daniel Gruen for joining the examination board on short notice.

Dr. Felix Metzner, thank you for so generously sharing your code with me and providing support at literally any time of the week. Dr. Markus Prim, thank you for working on the $|V_{\mathrm{cb}}|$ analysis with me and for answering my many questions.

I want to thank Dr. Alejandro Celis for proposing the initial idea of the clustering studies and Dr. Jason Aebischer for helping to select promising observables and to interpret the results.

Dr. James Kahn, I look back to many happy days spent in the office together. The Dresden hackathon is also a fond memory! Thank you for helping me brush up the abstract of this thesis!

Dr. Ilya Komarov, thank you for introducing me to the Belle II training group and guiding me initially. I also remember having a lot of fun at BNL together.

Dr. Martin Ritter, thank you for supporting my activities with the Belle II training group. We all miss your humorous takes on the latest news from our experiment.

I would like to thank Prof. Dr. Sudhir Malik, Dr. Samuel Meehan, and Dr. Graeme A Stewart for welcoming me to the HSF training group. I look forward to joining forces again soon!

I also want to express my gratitude to all Belle II training group members and the HSF training group for their fantastic work!

I thank Dr. Thomas Lueck for many valuable comments about the tagging calibration.

Dr. Thomas Hauth, thank you for introducing me to the validation package.

I want to thank Alvaro de Diego, Dr. Christiane Lorenz, Dr. Hannah Schrenk, Isabel Fischer, Dr. James Kahn, Dr. Leander Stecker, Lena Polke, Marcel Schaub, Martin Weber, Nathalie Eberlein, Dr. Nikolai Hartmann, Pascal Schmolz, Dr. Pas-

cal Stucky, Philip Lüghausen, Dr. Raynette van Tonder, Stephan Kürten, and Yaroslav Kulii for proofreading parts of this thesis on very short notice. The bulk of this text was completed on an extremely tight timeline, and you really boosted my confidence in the submission.

Lastly, none of this would have been possible without the most generous support of my family and friends. You know who you are, and I am much indebted to you all. Thank you for always having my back! To many more happy years together!