
Accelerating Discovery:
Model-Agnostic Likelihoods for the
Reinterpretation of Particle Physics
Results and their Application to the
Belle II $B^+ \rightarrow K^+ \nu \bar{\nu}$ Measurement

Lorenz Gärtner



Munich, 4.11.2025

**Accelerating Discovery:
Model-Agnostic Likelihoods for the
Reinterpretation of Particle Physics
Results and their Application to the
Belle II $B^+ \rightarrow K^+ \nu \bar{\nu}$ Measurement**

Lorenz Gärtner

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

vorgelegt von
Lorenz Gärtner
aus Wien, Österreich

München, den 4.11.2025

Erstgutachter: Prof. Dr. Thomas Kuhr
Zweitgutachter: Prof. Dr. Otmar Biebel
Tag der mündlichen Prüfung: 17.12.2025

Zusammenfassung

Experimentelle Ergebnisse in der Hochenergiephysik sind naturgemäß hypothesenabhängig, da Kollisionsanalysen darauf ausgelegt sind, bestimmte theoretische Rahmenbedingungen zu überprüfen. Die Vielfalt theoretischer Modelle führt zu einer Diskrepanz zwischen dem Tempo der Theorieentwicklung und der experimentellen Validierung. Dies macht eine Neuinterpretation bestehender Ergebnisse im Hinblick auf alternative Hypothesen erforderlich. Bestehende Strategien stoßen an ihre Grenzen: Eine vollständige Neuanalyse erfordert enorme Rechenressourcen und experimentelles Fachwissen, während vereinfachte Methoden durch die Vernachlässigung kinematischer Variationen und systematischer Unsicherheiten an Genauigkeit einbüßen.

Diese Arbeit entwickelt, validiert und wendet eine neuartige Methodik zur Neuinterpretation an, die statistische Genauigkeit bewahrt und gleichzeitig den Rechenaufwand drastisch reduziert, wodurch schnelle Rückschlüsse auf alternative theoretische Parameter durch leicht verteilbare Likelihoods ermöglicht werden.

Die Methode gewichtet die gemeinsame Verteilung von Rekonstruktions- und kinematischen Variablen entsprechend dem Verhältnis der theoretischen Vorhersagen von Alternativ- zu Nullhypothesen neu. Ausgehend von einer histogrammbasierten Wahrscheinlichkeit werden so aktualisierte Vorlagen für jedes theoretische Modell erstellt, wobei Korrelationen und systematische Unsicherheiten erhalten bleiben. Die Methodik wird anhand von Beispielen validiert und auf die Belle II $B^+ \rightarrow K^+ \nu \bar{\nu}$ Messung (2.7σ -Abweichung von dem Standardmodell) angewendet. Es werden zwei Analysen vorgestellt: eine Neuinterpretation der Weak Effective Theory, die Wilson-Koeffizienten einschränkt, und eine Suche nach leichter neuer Physik über $B^+ \rightarrow K^+ X$ Zweiteilchenzerfälle.

Die Neuinterpretation der Weak Effective Theory liefert die ersten direkten experimentellen Einschränkungen für die Wilson-Koeffizienten $b \rightarrow s \nu \bar{\nu}$. Der dominante Vektorkoeffizient $|C_{VL} + C_{VR}|$ erreicht seinen best-fit Wert bei 11.3 (Standard Model: 6.6) mit 95%-Konfidenzintervallen von $[1.9, 16.2]$, $[0.0, 15.4]$ und $[0.0, 11.2]$ für den Vektor-, Skalar- und Tensorsektor. Die Analyse der leichten neuen Physik ergibt eine best-fit Resonanzmasse von $m_X = 2.1$ GeV und einem Verzweigungsverhältnis von $\mathcal{B}(B^+ \rightarrow K^+ X) = 9.2 \cdot 10^{-6}$, was zu einer Signifikanz von 3.0σ gegenüber der Standard Model-Vorhersage führt. Die 95%-Konfidenzintervalle betragen $[1.9, 2.7]$ GeV bzw. $[3.4, 14.0] \cdot 10^{-6}$. Die Arbeit präsentiert auch den aktuellen Stand der kombinierten Vierkanal-Analyse von Belle II für $B \rightarrow K^{(*)} \nu \bar{\nu}$.

Diese Arbeit schafft einen neuartigen Rahmen für Neuinterpretationen. Die Veröffentlichung der daraus resultierenden modellunabhängigen Likelihoods schafft einen wichtigen Präzedenzfall für offene Wissenschaft in der experimentellen Teilchenphysik. Die Verwendung der Methodik von LHCb zeigt ihre weitreichende Bedeutung. Durch die Ermöglichung einer schnellen theoretischen Validierung demokratisiert dieser Ansatz den Zugang zu experimentellen Daten, erhöht die wissenschaftliche Rentabilität von Investitionen und fördert die Findable, Accessible, Interoperable and Reusable (FAIR) Datenprinzipien in der Hochenergiephysik.

Abstract

Experimental results in high-energy physics are inherently hypothesis-dependent, as collider analyses are designed to test specific theoretical frameworks. The proliferation of theoretical models creates a mismatch between the pace of theory development and experimental validation, motivating the reinterpretation of existing results in terms of alternative hypotheses. Existing strategies face limitations: full reanalysis requires enormous computational resources and experimental expertise, while simplified methods sacrifice accuracy by neglecting kinematic variations and systematic uncertainties.

This thesis develops, validates, and applies a novel reinterpretation method that preserves statistical rigour while dramatically reducing computational requirements, enabling rapid inference on alternative theoretical parameters through distributable likelihoods.

The method reweights the joint distribution of reconstruction and kinematic variables according to the ratio of theoretical predictions of alternative to null hypotheses. Starting from a histogram-based likelihood, this produces updated templates for any theoretical model while preserving correlations and systematic uncertainties. The method is validated through examples and applied to the Belle II $B^+ \rightarrow K^+ \nu \bar{\nu}$ measurement (2.7σ tension with the Standard Model). Two analyses are presented: a Weak Effective Theory reinterpretation constraining Wilson coefficients, and a light new physics search via $B^+ \rightarrow K^+ X$ two-body decays.

The Weak Effective Theory reinterpretation provides the first direct experimental constraints on $b \rightarrow s \nu \bar{\nu}$ Wilson coefficients. The dominant vector coefficient $|C_{\text{VL}} + C_{\text{VR}}|$ peaks at 11.3 (Standard Model: 6.6) with 95% credible intervals of [1.9, 16.2], [0.0, 15.4], and [0.0, 11.2] for vector, scalar, and tensor sectors. The light new physics analysis yields a best-fit resonance mass of $m_X = 2.1$ GeV and branching ratio of $\mathcal{B}(B^+ \rightarrow K^+ X) = 9.2 \cdot 10^{-6}$, leading to a 3.0σ significance over the Standard Model prediction. The 95% credible intervals are [1.9, 2.7] GeV and $[3.4, 14.0] \cdot 10^{-6}$, respectively. The thesis also presents the current status on Belle II's combined four-channel $B \rightarrow K^{(*)} \nu \bar{\nu}$ analysis.

This work establishes a novel reinterpretation framework. The public release of the resulting model-agnostic likelihoods establishes an important precedent for open science in experimental particle physics. The method's adoption by LHCb demonstrates its broader impact. By facilitating rapid theoretical validation, this approach democratizes experimental data access, enhances scientific return on investments, and advances Findable, Accessible, Interoperable and Reusable (FAIR) Data principles in high-energy physics.

Contents

| | |
|--|------------|
| Zusammenfassung | I |
| Abstract | III |
| 1 Introduction | 1 |
| 2 Theoretical background | 5 |
| 2.1 The Standard Model of particle physics | 6 |
| 2.1.1 The Standard Model Lagrangian | 8 |
| 2.1.2 Spontaneous symmetry breaking | 10 |
| 2.1.3 Flavour changing neutral currents | 14 |
| 2.2 Effective field theories | 15 |
| 2.2.1 Weak Effective Theory | 16 |
| 2.3 Hadronic matrix elements | 17 |
| 2.4 Theory predictions for $B \rightarrow K^{(*)}\nu\bar{\nu}$ | 20 |
| 2.4.1 Observables for $B \rightarrow K\nu\bar{\nu}$ | 20 |
| 2.4.2 Observables for $B \rightarrow K^*\nu\bar{\nu}$ | 21 |
| 2.4.3 The EOS software package | 23 |
| 3 The Belle II experiment | 27 |
| 3.1 B factories | 28 |
| 3.2 Accelerator: SuperKEKB | 28 |
| 3.3 The Belle II detector | 29 |
| 3.3.1 Pixel vertex detector | 30 |
| 3.3.2 Silicon vertex detector | 32 |
| 3.3.3 Central drift chamber | 32 |
| 3.3.4 Particle identification system | 33 |
| 3.3.4.1 Barrel: Time-of-propagation counter | 33 |
| 3.3.4.2 End-cap: Aerogel Ring-Imaging Cherenkov detector | 33 |
| 3.3.5 Electromagnetic calorimeter | 33 |
| 3.3.6 K_0^L and μ detection | 34 |
| 3.3.7 Trigger | 34 |
| 3.4 Computing and software | 35 |
| 3.4.1 Reconstruction | 35 |

| | | |
|----------|---|-----------|
| 3.4.2 | Simulation | 36 |
| 4 | Binned statistical models in particle physics | 39 |
| 4.1 | Probability and random variables | 40 |
| 4.1.1 | Interpretations of probability | 41 |
| 4.1.1.1 | Frequentist probability | 41 |
| 4.1.1.2 | Subjective probability | 41 |
| 4.1.1.3 | The common ground | 42 |
| 4.2 | The nature of statistical inference | 42 |
| 4.3 | Counting experiments | 43 |
| 4.3.1 | Binned likelihoods | 44 |
| 4.4 | HistFactory | 45 |
| 4.4.1 | Modifier types | 47 |
| 4.4.2 | Bayesian HistFactory | 49 |
| 4.4.3 | HistFactory in Python: pyhf | 49 |
| 4.4.4 | Extended features for HistFactory and pyhf | 51 |
| 4.4.4.1 | Arbitrarily correlated uncertainties | 51 |
| 4.4.4.2 | Custom modifiers | 53 |
| 5 | Statistical inference in particle physics | 57 |
| 5.1 | Parameter estimation | 57 |
| 5.1.1 | Method of maximum likelihood | 58 |
| 5.1.1.1 | Asymptotic properties of maximum likelihood estimators | 58 |
| 5.1.2 | Maximum a posteriori estimate | 59 |
| 5.1.3 | Equivalence of maximum likelihood and maximum a posteriori estimators | 59 |
| 5.2 | Model comparison | 60 |
| 5.2.1 | Point-wise model comparison: hypothesis testing | 60 |
| 5.2.1.1 | Weighted P -values | 61 |
| 5.2.1.2 | Functional form of the test statistic | 63 |
| 5.2.1.3 | Distribution of the test statistic | 63 |
| 5.2.1.4 | Nested hypotheses | 64 |
| 5.2.1.5 | Significance | 65 |
| 5.2.2 | Global model comparison: Bayes factors | 65 |
| 5.3 | Goodness-of-fit | 67 |
| 5.4 | Limit setting | 68 |
| 5.4.1 | Credible intervals | 68 |
| 5.4.2 | Confidence intervals | 69 |
| 6 | Ongoing reinterpretation efforts in particle physics | 73 |
| 6.1 | Data preservation | 73 |
| 6.2 | Analysis preservation | 74 |
| 6.3 | Statistical models and their preservation | 75 |
| 6.4 | Types of reinterpretation | 76 |

| | | |
|----------|--|------------|
| 6.5 | Public tools for reinterpretation | 76 |
| 6.6 | Challenges and future goals | 77 |
| 7 | Model-agnostic likelihoods for the reinterpretation of particle physics results | 79 |
| 7.1 | Reinterpretation through reweighting | 80 |
| 7.1.1 | Discrete reweighting | 81 |
| 7.1.2 | Limitation | 82 |
| 7.2 | Implementation of the reweighting method | 82 |
| 7.3 | Application of the reweighting method to toy examples | 82 |
| 7.3.1 | $B \rightarrow K\nu\bar{\nu}$ | 83 |
| 7.3.1.1 | Weak Effective Theory parametrization | 83 |
| 7.3.1.2 | Datasets | 84 |
| 7.3.1.3 | Full statistical model | 87 |
| 7.3.1.4 | Reinterpretation results | 89 |
| 7.3.2 | Combination of $B \rightarrow K\nu\bar{\nu}$ and $B \rightarrow K^*\nu\bar{\nu}$ | 90 |
| 7.3.2.1 | $B \rightarrow K^*\nu\bar{\nu}$ Weak Effective Theory parametrization | 90 |
| 7.3.2.2 | Datasets | 92 |
| 7.3.2.3 | Full statistical model | 94 |
| 7.3.2.4 | Reinterpretation results | 94 |
| 7.3.3 | The necessity for reinterpretation | 97 |
| 7.4 | Discussion and significance of the method | 97 |
| 8 | A model-agnostic likelihood for the Belle II $B^+ \rightarrow K^+\nu\bar{\nu}$ measurement | 101 |
| 8.1 | The Belle II $B^+ \rightarrow K^+\nu\bar{\nu}$ analysis | 103 |
| 8.1.1 | Data and simulated samples | 103 |
| 8.1.2 | Background suppression | 103 |
| 8.1.3 | Signal region definition | 104 |
| 8.1.4 | Signal extraction result | 105 |
| 8.2 | Joint number densities for the Belle II $B^+ \rightarrow K^+\nu\bar{\nu}$ analysis | 107 |
| 9 | A Weak Effective Theory reinterpretation of the Belle II $B^+ \rightarrow K^+\nu\bar{\nu}$ measurement | 111 |
| 9.1 | Weak Effective Theory as alternative hypothesis | 112 |
| 9.2 | Marginal posterior and credible intervals | 112 |
| 9.2.1 | Statistical model and prior choices | 113 |
| 9.2.2 | Posterior sampling and marginalization | 113 |
| 9.2.3 | Credible intervals and posterior mode | 114 |
| 9.2.4 | Physical interpretation of the posterior | 114 |
| 9.2.5 | Prior sensitivity | 117 |
| 9.2.6 | Naive reinterpretation | 118 |
| 9.3 | Goodness-of-fit | 119 |
| 9.4 | Model comparison | 122 |
| 9.4.1 | Bayes factors | 122 |

| | | |
|-----------|--|------------|
| 9.4.2 | Frequentist hypothesis test | 123 |
| 9.5 | Discussion | 124 |
| 10 | A light new physics reinterpretation of the Belle II $B^+ \rightarrow K^+ \nu \bar{\nu}$ measurement | 127 |
| 10.1 | The $B^+ \rightarrow K^+ X$ model as alternative hypothesis | 128 |
| 10.2 | Marginal posterior and credible intervals | 131 |
| 10.2.1 | Statistical model and prior choices | 131 |
| 10.2.2 | Posterior sampling and marginalization | 131 |
| 10.2.3 | Credible intervals and posterior mode | 132 |
| 10.2.4 | Physical interpretation of the posterior | 133 |
| 10.2.5 | Prior sensitivity | 135 |
| 10.3 | Frequentist mass scan | 135 |
| 10.3.1 | Injection study validation | 138 |
| 10.4 | Goodness-of-fit | 138 |
| 10.5 | Model comparison | 140 |
| 10.5.1 | Bayes factors | 140 |
| 10.5.2 | Frequentist hypothesis test | 141 |
| 10.6 | Discussion | 142 |
| 11 | A combined analysis of $B \rightarrow K^{(*)} \nu \bar{\nu}$ decays at Belle II | 145 |
| 11.1 | Theory predictions | 146 |
| 11.2 | Analysis outline | 147 |
| 11.3 | Background suppression | 149 |
| 11.4 | Signal region definition | 149 |
| 11.5 | Signal extraction | 150 |
| 11.5.1 | Isospin average | 152 |
| 11.5.2 | 4-channel correlation | 152 |
| 11.6 | Discussion | 153 |
| 12 | Conclusion | 165 |
| | List of Publications | 169 |
| A | Theoretical background | 171 |
| A.1 | Group generators | 171 |
| A.1.1 | U(1) generators | 171 |
| A.1.2 | SU(2) generators | 171 |
| A.1.3 | SU(3) generators | 172 |
| A.2 | Dirac matrices | 173 |
| A.3 | $B \rightarrow K^*$ form factors | 174 |
| A.4 | $B \rightarrow K^* \nu \bar{\nu}$ hadronic form factor parameters | 175 |
| B | Statistical inference in particle physics | 179 |
| B.1 | Probability density functions | 179 |

| | | |
|----------|--|------------|
| B.1.1 | Cumulative distribution functions | 180 |
| B.1.2 | Transformation properties | 180 |
| B.1.3 | Expectation values | 181 |
| B.2 | Conjugate priors | 181 |
| B.3 | Proof of the Cramér-Rao bound | 183 |
| B.4 | Fisher information | 184 |
| B.5 | Proof of consistency of maximum likelihood estimators | 184 |
| B.6 | Proof of asymptotic normality of the maximum likelihood estimator | 185 |
| B.7 | Proof of the Neyman-Pearson lemma | 186 |
| B.8 | Likelihood ratio test statistic for normally distributed data | 187 |
| B.9 | Convergence of random variables | 188 |
| B.10 | Posterior validation | 190 |
| B.10.1 | Potential scale reduction factor | 190 |
| B.10.2 | Effective sample size | 190 |
| B.10.3 | Visual diagnostics | 190 |
| C | Model-agnostic likelihoods for the reinterpretation of particle physics results | 193 |
| C.1 | Individual contributions | 193 |
| C.2 | Optimizing the kinematic binning | 194 |
| C.3 | A recipe for application of this reinterpretation method | 194 |
| D | A model-agnostic likelihood for the Belle II $B^+ \rightarrow K^+ \nu \bar{\nu}$ measurement | 195 |
| D.1 | Individual contributions | 195 |
| D.2 | Final-state radiation | 195 |
| D.3 | Kinematic q^2 binning | 196 |
| D.4 | Impact of removing unphysical q^2 values in the hadronic tagging analysis | 197 |
| E | A Weak Effective Theory reinterpretation of the Belle II $B^+ \rightarrow K^+ \nu \bar{\nu}$ measurement | 199 |
| E.1 | Belle II HEPData reinterpretation inventory | 199 |
| E.2 | Weak Effective Theory posterior validation | 199 |
| E.3 | Weak Effective Theory posteriors for the inclusive and hadronic tagging analyses | 201 |
| E.3.1 | Marginal posterior for the inclusive tagging analysis | 201 |
| E.3.2 | Marginal posterior for hadronic tagging analysis | 203 |
| E.4 | Goodness-of-fit | 203 |
| F | A light new physics reinterpretation of the Belle II $B^+ \rightarrow K^+ \nu \bar{\nu}$ measurement | 207 |
| F.1 | $B^+ \rightarrow K^+ X$ posterior validation | 207 |
| F.2 | Signal injection study | 207 |
| G | A combined analysis of $B \rightarrow K^{(*)} \nu \bar{\nu}$ decays at Belle II | 215 |
| G.1 | Individual contributions | 215 |

CONTENTS

| | |
|--|------------|
| List of Acronyms | 217 |
| Disclaimer on the use of language tools | 229 |
| Acknowledgements | 231 |

Chapter 1

Introduction

How do we make progress in science? The foundation of scientific inquiry rests on a systematic approach that has guided discovery for centuries. According to the Oxford Dictionary, the scientific method is defined as [1]

“The scientific method is now commonly represented as ideally comprising some or all of (a) systematic observation, measurement, and experimentation, (b) induction and the formulation of hypotheses, (c) the making of deductions from the hypotheses, (d) the experimental testing of the deductions, and (if necessary) (e) the modification of the hypotheses [...].”

A crucial aspect of this definition is that hypotheses and their consequent deductions (predictions) must precede experimental testing. The implications of this ordering are profound: experimental results are inherently hypothesis-dependent, since experiments are deliberately designed to test specific theoretical frameworks.

High-Energy Physics (**HEP**) exemplifies this principle. Enormous investments in particle accelerators, sophisticated detectors, and complex data analysis pipelines are all guided by theoretical predictions we seek to validate. The patterns we search for in high-dimensional datasets, the statistical methods we employ, and even the very definition of a “signal” are all determined by our chosen theoretical hypothesis. Consequently, experimental results are meaningful only within the context of the specific theoretical framework under investigation.

This hypothesis-dependence, however, is often overlooked when interpreting experimental results, creating a critical methodological problem. When new theoretical models emerge with different predictions, their validity may be assessed against experiments that were never designed to test them. Such assessments risk corrupting the scientific method by modifying or excluding hypotheses on inadequate empirical grounds.

Addressing this methodological challenge is the central goal of this work.

The Standard Model (**SM**) of particle physics represents the currently accepted paradigm in **HEP** — a gauge quantum field theory that has withstood empirical tests with extraor-

dinary precision. Despite its remarkable success, extracting concrete predictions from this complex theoretical framework remains computationally challenging. Effective Field Theories (EFTs) address this challenge by approximating the full theory up to a specific energy scale, systematically incorporating effects from both SM and potential Beyond the Standard Model (BSM) physics while maintaining computational tractability. The theoretical foundations underlying these approaches are discussed in Chapter 2.

The search for BSM physics often focuses on rare particle decays, where BSM contributions become relatively more prominent compared to SM processes, making deviations easier to detect. A particularly compelling example from flavour physics, which forms the central focus of this work, involves $B \rightarrow K^{(*)}\nu\bar{\nu}$ decays studied at Belle II. These decay channels have gained significant attention following the Belle II collaboration’s recent $B^+ \rightarrow K^+\nu\bar{\nu}$ measurement, which exhibits a 2.7σ tension with SM predictions [2]. The Belle II experiment is described in Chapter 3.

The complexity of particle collider data analysis necessitates sophisticated statistical models and inference techniques to ensure reliable results with high statistical power. However, this very complexity means that not all theoretical hypotheses can be directly tested against experimental data within practical time and computational constraints. Statistical frameworks underlying these challenges are discussed in Chapters 4 and 5.

These practical limitations motivate a powerful alternative approach: *reinterpretation* — the process of interpreting experimental results obtained under one theoretical hypothesis in terms of alternative hypotheses with different predictions. Rather than conducting entirely new measurements for each theoretical model, reinterpretation leverages existing experimental results to constrain or validate new physics scenarios. While the concept itself is not novel, existing methodological approaches have various limitations, which are reviewed in Chapter 6.

The novel reinterpretation methodology presented in this work addresses key requirements: accuracy, generality, computational efficiency, and accessibility. The approach enables fast, precise reinterpretations while enabling the preservation and distribution of full statistical models. Both the method and its validation through examples are presented in Chapter 7.

The practical application of this methodology is demonstrated using the aforementioned Belle II $B^+ \rightarrow K^+\nu\bar{\nu}$ measurement [2], originally interpreted within the SM framework. The analysis is presented in Chapter 8, followed by reinterpretations in terms of the Weak Effective Theory (WET) [3–5] in Chapter 9. The WET encompasses new physics models such as leptoquarks [6–8] and heavy Z' bosons [9–12]. A reinterpretation in terms of a two-body light new physics model is presented in Chapter 10. This encompasses new physics models such as axions [10, 13–15], axion-like particles [10, 16–18] or other dark-sector mediators [19–22].

The reproducibility and distributability of statistical models represent more than scientific best practices — they enable powerful combined analyses that can dramatically reduce uncertainties. Building on this capability, the Belle II collaboration is developing

a combined analysis of four different $B \rightarrow K^{(*)}\nu\bar{\nu}$ decay channels, whose current status is described in [Chapter 11](#). This combined analysis will significantly enhance sensitivity to [BSM](#) effects and [WET](#) parameters, establishing an excellent foundation for future reinterpretation studies at Belle II and beyond.

CHAPTER 1. INTRODUCTION

Chapter 2

Theoretical background

¹ Theoretical frameworks provide the mathematical foundation for interpreting experimental observations and testing our understanding of fundamental physics. The **SM** of particle physics represents our most successful description of the fundamental particles and their interactions, excluding gravity. Despite its remarkable experimental validation across energy scales from the eV to the TeV regime [10], the **SM** cannot account for several observed phenomena, including dark matter, dark energy, neutrino masses, and the baryon asymmetry of the universe. These limitations provide compelling evidence for **BSM** physics.

Flavour Changing Neutral Current (**FCNC**) processes offer a particularly sensitive probe of **BSM** physics. In the **SM**, such transitions are forbidden at tree level and can only proceed through strongly suppressed loop-induced effects, resulting in high sensitivity to potential **BSM** contributions [23–26]. Among these, the $b \rightarrow s\nu\bar{\nu}$ transition stands out as theoretically clean due to the absence of photon exchanges, which break factorization in charged lepton modes like $b \rightarrow sl^+\ell^-$. The rare decays $B \rightarrow K^{(*)}\nu\bar{\nu}$, mediated by $b \rightarrow s\nu\bar{\nu}$, thus provide an ideal laboratory for precision tests of the **SM** and searches for **BSM** signatures.

EFTs provide the optimal theoretical framework for analysing such processes. By systematically separating high-energy **BSM** physics from low-energy **SM** dynamics, **EFTs** enable model-independent parametrization of new physics effects through Wilson coefficients of higher-dimensional operators. This approach proves particularly powerful when the **BSM** scale lies well above current experimental reach, as it captures the effects of entire classes of **BSM** models through a finite set of effective operators.

For $b \rightarrow s\nu\bar{\nu}$ transitions, the **WET** provides the appropriate low-energy description below the electroweak scale. The theoretical prediction for $B \rightarrow K^{(*)}\nu\bar{\nu}$ decays requires factorizing short-distance physics (encoded in Wilson coefficients) from long-distance, non-

¹Natural units with $\hbar = c = 1$ are used throughout this work. The Minkowski metric tensor is defined as $g^{\mu\nu} = \text{diag}(1, -1, -1, -1)$. Repeated indices are summed over unless otherwise specified. The Minkowski inner product of two four-vectors p^μ and q^μ is defined as $p \cdot q = p^\mu q_\mu = p^0 q^0 - \vec{p} \cdot \vec{q}$.

perturbative Quantum Chromodynamics (QCD) dynamics (captured by hadronic form factors). This factorization enables precision tests of the SM and systematic searches for BSM signatures through deviations in experimentally measured Wilson coefficients.

This chapter establishes the complete theoretical framework for $B \rightarrow K^{(*)}\nu\bar{\nu}$ phenomenology. Section 2.1 reviews the SM foundations, emphasizing the electroweak symmetry breaking mechanism (Section 2.1.2) that generates fermion and gauge boson masses, and the origin of FCNCs through the Cabibbo-Kobayashi-Maskawa (CKM) matrix structure (Section 2.1.3). Section 2.2 develops the EFT formalism, with detailed treatment of the WET Lagrangian (Section 2.2.1) and its dimension-six operators relevant for $b \rightarrow s\nu\bar{\nu}$ transitions. The non-perturbative evaluation of hadronic matrix elements through form factor parametrizations is discussed in Section 2.3. Finally, Section 2.4 combines these elements to derive differential decay rates for $B \rightarrow K\nu\bar{\nu}$ (Section 2.4.1) and $B \rightarrow K^*\nu\bar{\nu}$ (Section 2.4.2), demonstrating the complementary Wilson coefficient sensitivities that enable comprehensive BSM searches when both decay modes are analysed simultaneously.

2.1 The Standard Model of particle physics

² The SM of particle physics is a locally gauged quantum field theory that respects Poincaré invariance. It describes three of the four fundamental forces of nature: the electromagnetic, weak, and strong interactions. Gravity is not included in the SM. The basis of this model is its gauge group

$$G_{SM} = \text{SU}(3)_C \times \text{SU}(2)_W \times \text{U}(1)_Y, \quad (2.1)$$

where $\text{SU}(n)$ is the special unitary group

$$\text{SU}(n) := \{U \in \mathbb{C}^n | UU^\dagger = U^\dagger U = \mathbb{1}, \det U = 1\} \quad (2.2)$$

and $U(n)$ is the unitary group

$$U(n) := \{U \in \mathbb{C}^n | UU^\dagger = U^\dagger U = \mathbb{1}\}. \quad (2.3)$$

$\text{SU}(3)_C$ is the gauge group of strong interactions, or QCD. $\text{SU}(2)_W$ is the gauge group of weak isospin, and $\text{U}(1)_Y$ is the gauge group of hypercharge. The number of generators of the underlying Lie algebras defines the number of gauge bosons in the SM. The number of generators of $\text{SU}(n)$ is $n^2 - 1$, resulting in a total of 8 *gluons* from $\text{SU}(3)_C$ and 3 *W* gauge bosons from $\text{SU}(2)_W$. The unitary Lie algebra has n^2 generators. Consequently, one gets 1 gauge boson *B* from $\text{U}(1)_Y$. Gauge bosons carry spin-1.

The matter field content of the SM comprises fermions, which carry spin-1/2. They can be characterized according to their interaction properties with the gauge bosons or, in other words, their transformation properties under the SM gauge group. Under the SM gauge group they transform as shown in Table 2.1. Since right- and left-handed fermions

²This chapter follows References [23–26], unless otherwise specified.

Table 2.1: Transformation properties of **SM** fermions under the gauge symmetry groups $SU(3)_C$, $SU(2)_W$, and $U(1)_Y$. The numbers indicate the representation dimension under each gauge group (singlet, doublet, or triplet), while the hypercharge Y values are given for $U(1)_Y$. Left-handed (L) fermions transform as doublets under $SU(2)_W$, while right-handed (R) fermions are singlets. The index i labels the three generations of fermions [26].

| Fermions | $SU(3)_C$ | $SU(2)_W$ | $U(1)_Y$ |
|------------|-----------|-----------|----------|
| Q_{iL} | 3 | 2 | +1/6 |
| u_{iR} | 3 | 1 | +2/3 |
| d_{iR} | 3 | 1 | -1/3 |
| L_{iL} | 1 | 2 | -1/2 |
| l_{iR} | 1 | 1 | -1 |
| ν_{iR} | 1 | 1 | 0 |

carry different weak isospin ($\pm 1/2$ for doublets and 0 for singlets), the **SM** is a chiral gauge theory [27].

Left-handed quarks are represented by Q_{iL} , right-handed up- and down-type *quarks* by u_{iR} and d_{iR} , left-handed *leptons* by L_{iL} , and right-handed *leptons* by l_{iR} and ν_{iR} . Each of these fields comes in three generations, as indicated by the subscript i ,

$$\begin{aligned}
Q_{iL} &\in [(u, d)_L, (c, s)_L, (t, b)_L] \\
u_{iR} &\in [u_R, c_R, t_R] \\
d_{iR} &\in [d_R, s_R, b_R] \\
L_{iL} &\in [(\nu_e, e)_L, (\nu_\mu, \mu)_L, (\nu_\tau, \tau)_L] \\
l_{iR} &\in [e_R, \mu_R, \tau_R]
\end{aligned} \tag{2.4}$$

The **SM** can be readily augmented by including right-handed neutrinos $\nu_{iR} \in [\nu_{eR}, \nu_{\mu R}, \nu_{\tau R}]$, which have never been observed.

The **SM** has been constructed out of left-handed fermion doublets and right-handed fermion singlets, because W^\pm bosons couple only to left-handed fermions, and because the $SU(2)$ group admits a chargeless 1-dimensional singlet representation and a charged multidimensional representation.

It is an open question why there are exactly three generations of quarks and leptons. However, this number of generations leads to the cancellation of anomalies, i.e. the loss of symmetries of the underlying classical theory after quantization. Anomaly cancellation is a necessity for the consistency of a quantum field theoretical description of nature [27].

2.1.1 The Standard Model Lagrangian

The **SM** Lagrangian contains all renormalizable, gauge invariant operators with the specified matter fields [28]. In the most general form, one can write it as

$$\mathcal{L}_{SM} = \mathcal{L}_{kin} + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{Yukawa}. \quad (2.5)$$

The Lagrangian \mathcal{L}_{kin} includes field dynamics through kinetic terms as well as gauge interactions through covariant derivatives and non-Abelian field strengths. The Lagrangian $\mathcal{L}_{\text{Higgs}}$ contains the Higgs kinematics and potential terms, creating mass terms for the gauge bosons. The Lagrangian \mathcal{L}_{Yukawa} contains all interactions of the Higgs field with fermions, and is hence responsible for generating fermion masses. In this section I will focus on the kinetic Lagrangian \mathcal{L}_{kin} . I will discuss the other two parts in [Section 2.1.2](#).

The **SM** symmetry group G_{SM} can be split into a **QCD** part, $SU(3)_C$ and an electroweak part $SU(2)_W \times U(1)_Y$. Similarly, one can separate the kinetic Lagrangian,

$$\mathcal{L}_{kin} = \mathcal{L}_{QCD} + \mathcal{L}_{EW}, \quad (2.6)$$

into a **QCD** part, \mathcal{L}_{QCD} , and an electroweak part, \mathcal{L}_{EW} .

QCD is described by a non-abelian gauge theory, invariant under $SU(3)$ transformations,

$$q \rightarrow q' = Uq \quad \text{with} \quad U = \exp(i\alpha_a T^a), \quad (2.7)$$

where q is a quark multiplet of $SU(3)$, α_a are real parameters, and T^a are the group generators. For $SU(3)$ there are 8 generators, which are usually chosen as $T^a = \lambda^a/2$, where λ^a are the Gell-Mann matrices (see [Appendix A.1](#)). The generators obey the commutation relation

$$[T^a, T^b] = if^{abc}T^c, \quad (2.8)$$

where f^{abc} are antisymmetric structure constants. Abelian gauge theories, such as Quantum Electrodynamics (**QED**), are described by $U(1)$ transformations, which are characterized by a single generator.

The **QCD** contribution to the kinetic Lagrangian is

$$\mathcal{L}_{QCD} = \bar{q}^i \not{D}_{ij} q^j - \frac{1}{4} G_{\mu\nu}^a G_a^{\mu\nu}. \quad (2.9)$$

Here q^i are the quark fields,

$$q^i = (Q_{1L}^i, Q_{2L}^i, Q_{3L}^i, u_{1R}^i, u_{2R}^i, u_{3R}^i, d_{1R}^i, d_{2R}^i, d_{3R}^i) \quad (2.10)$$

with color index i running from 1 to 3, $\bar{q}^i \equiv q^\dagger \gamma^0$ is the adjoint spinor and γ^μ are the Dirac matrices (see [Appendix A.2](#)). The covariant derivative $\not{D}_{ij} = \gamma_\mu D_{ij}^\mu$ contains the gauge fields A_a^μ and is given by

$$D_{ij}^\mu = \delta_{ij} \partial^\mu - ig_C T_{ij}^a A_a^\mu, \quad (2.11)$$

where g_C is the **QCD** coupling. The gauge field $A^\mu = T^a A_a^\mu$ is a vector field, which transforms under $SU(3)_C$ as

$$A^\mu \rightarrow A'^\mu = U A^\mu U^\dagger - \frac{i}{g_C} (\partial^\mu U) U^\dagger. \quad (2.12)$$

The kinetic terms of the gauge fields enter through the field strength tensor, given by the commutator

$$G_{\mu\nu} = G_{\mu\nu}^a T_a = \frac{i}{g_C} [D_\mu, D_\nu] = \partial_\mu A_\nu - \partial_\nu A_\mu - ig_C [A_\mu, A_\nu]. \quad (2.13)$$

From the last term in the field strength tensor one gets cubic and quartic terms in the gauge fields. This leads to the self-interaction of gluons and is the fundamental reason for confinement.

The electroweak group is the product $SU(2)_W \times U(1)_Y$. Furthermore, the electroweak theory is special because it is a chiral theory. Given the **SM** fermion content in **Equation (2.4)**, one can write down the covariant derivatives for the chiral fields

$$D_\mu \psi_L = (\partial_\mu - igW_\mu - ig'YB_\mu)\psi_L, \quad (2.14)$$

$$D_\mu \psi_R = (\partial_\mu - ig'YB_\mu)\psi_R, \quad (2.15)$$

where the fermion fields are separated into the vectors

$$\psi_L = (Q_{1L}, Q_{2L}, Q_{3L}, L_{1L}, L_{2L}, L_{3L}), \quad (2.16)$$

$$\psi_R = (u_{1R}, u_{2R}, u_{3R}, d_{1R}, d_{2R}, d_{3R}, l_{1R}, l_{2R}, l_{3R}) \quad (2.17)$$

The $W_\mu = \tau_i W_\mu^i$ are the three $SU(2)_W$ gauge fields, $\tau_i = \sigma_i/2$ are the generators, where σ_i are the Pauli matrices (see **Appendix A.1**), and g is a coupling constant. The $U(1)_Y$ gauge field is B_μ , Y is the hypercharge and g' is a coupling constant.

From these one can build the kinetic contribution to the electroweak Lagrangian,

$$\mathcal{L}_{EW} = i\bar{\psi}_L \not{D} \psi_L + i\bar{\psi}_R \not{D} \psi_R - \frac{1}{4} W_{\mu\nu}^i W_i^{\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}. \quad (2.18)$$

The kinetic terms of the gauge fields enter through the field strength tensors

$$W_{\mu\nu} = W_{\mu\nu}^i T_i = \frac{i}{g} [D_\mu, D_\nu] = \partial_\mu W_\nu - \partial_\nu W_\mu - ig[W_\mu, W_\nu], \quad (2.19)$$

$$B_{\mu\nu} = \frac{i}{g'Y} [D_\mu, D_\nu] = \partial_\mu B_\nu - \partial_\nu B_\mu. \quad (2.20)$$

This completes the kinetic Lagrangian of the **SM** gauge group G_{SM} . One can summarize it as

$$\mathcal{L}_{kin} = i\bar{q}^i \not{D}_{ij} q^j + i\bar{\psi}_L \not{D} \psi_L + i\bar{\psi}_R \not{D} \psi_R - \frac{1}{4} G_{\mu\nu}^a G_a^{\mu\nu} - \frac{1}{4} W_{\mu\nu}^i W_i^{\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}. \quad (2.21)$$

To make the theory actually describe the physical world, there is one crucial ingredient missing: masses. So far, neither the gauge bosons nor the fermions are considered to be massive. Massless gauge bosons would imply long-range forces, not observed in nature. The weak interaction has short range, requiring massive gauge bosons. How mass terms are generated while preserving gauge invariance is discussed in the next section.

2.1.2 Spontaneous symmetry breaking

Mass terms such as $m^2 A_\mu A^\mu$ break gauge invariance, since they are not invariant under the gauge transformations such as in Equation (2.12). Furthermore, fermion mass terms also break gauge invariance, since these mix left- and right-handed fermions,

$$m (\bar{\psi}_L \psi_R + \bar{\psi}_R \psi_L) , \quad (2.22)$$

which transform differently under the gauge group.

Introducing gauge invariant terms, which give rise to masses for the gauge bosons and fermions, is the main purpose of the Higgs mechanism. It is built on the idea of spontaneous symmetry breaking [23, 27]. The Higgs mechanism is the simplest way to generate gauge invariant mass terms for the gauge bosons and fermions, by introducing a single complex scalar field Φ .

The newly introduced scalar field Φ is a doublet under the electroweak gauge group $SU(2)_W$, with hypercharge $+1/2$, with four real degrees of freedom (d.o.f.). The Higgs Lagrangian is given by

$$\mathcal{L}_{\text{Higgs}} = (D_\mu \Phi)^\dagger D^\mu \Phi - V(\Phi^\dagger \Phi). \quad (2.23)$$

The covariant derivative D_μ contains the gauge fields,

$$D_\mu = \partial_\mu - igW_\mu - \frac{i}{2}g'B_\mu. \quad (2.24)$$

The beauty of this field lies in its potential, which is given by

$$V(\Phi^\dagger \Phi) = -\mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2, \quad (2.25)$$

which has a non-zero vacuum expectation value (vev) at $\Phi^\dagger \Phi = v^2/2 \equiv \mu^2/(2\lambda)$ for $\mu^2 > 0$ and $\lambda > 0$. One can take the vev to be real and positive, $v > 0$, without loss of generality,

$$\langle \Phi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}. \quad (2.26)$$

In the vicinity of the vev, the theory is not invariant under the gauge transformations of $SU(2)_W \times U(1)_Y$, but only under the electromagnetic $U(1)_{\text{em}}$ transformations. This is colloquially known as breaking of the gauge symmetry from $SU(2)_W \times U(1)_Y$ to $U(1)_{\text{em}}$. The generator of the electromagnetic transformation is the $Q = T_3 + Y$ charge, where T_3 is the third component of the weak isospin and Y is the weak hypercharge. This is known as spontaneous symmetry breaking [23].

In the unitarity gauge, the Higgs field can be written as

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}, \quad (2.27)$$

where $h(x)$ is a real scalar field. The Lagrangian [Equation \(2.23\)](#) then becomes

$$\begin{aligned} \mathcal{L}_{\text{Higgs}} = & \frac{1}{2}(\partial_\mu h)^2 + \frac{(v+h)^2}{8} \left[g^2 (W_\mu^1 W^{1\mu} + W_\mu^2 W^{2\mu}) + (gW_\mu^3 - g'B_\mu)^2 \right] \\ & - \left[\lambda v^2 h^2 + \lambda v h^3 + \frac{\lambda}{4} h^4 \right]. \end{aligned} \quad (2.28)$$

One can identify the mass terms for the gauge bosons and the Higgs field. The Higgs mass is easily read off as $m_h^2 = 2\lambda v^2$. To diagonalize the gauge boson mass terms, one performs a change of basis

$$W_\mu^\pm = \frac{1}{\sqrt{2}} (W_\mu^1 \mp iW_\mu^2), \quad (2.29)$$

$$Z_\mu = \frac{1}{\sqrt{g^2 + g'^2}} (gW_\mu^3 - g'B_\mu), \quad (2.30)$$

$$A_\mu = \frac{1}{\sqrt{g^2 + g'^2}} (g'W_\mu^3 + gB_\mu), \quad (2.31)$$

which are the three massive weak fields W^\pm , Z , and the massless photon field A . Often this is written in terms of the Weinberg angle θ_W , defined as

$$\tan \theta_W = \frac{g'}{g} \quad \text{or} \quad \sin^2 \theta_W = \frac{g'^2}{g^2 + g'^2}. \quad (2.32)$$

The mass terms for the gauge bosons are then

$$\mathcal{L}_{\text{Higgs}} \supset m_W^2 W_\mu^+ W^{-\mu} + \frac{1}{2} m_Z^2 Z_\mu Z^\mu. \quad (2.33)$$

The masses of the W^\pm and Z bosons are

$$m_W = \frac{1}{2} g v, \quad m_Z = \frac{1}{2} \sqrt{g^2 + g'^2} v = \frac{m_W}{\cos \theta_W}. \quad (2.34)$$

The interaction terms between the Higgs field and the gauge bosons are given by

$$\mathcal{L}_{\text{Higgs}} \supset 2 \frac{m_W^2}{v} h W_\mu^+ W^{-\mu} + \frac{m_W^2}{v^2} h^2 W_\mu^+ W^{-\mu} + \frac{m_Z^2}{v} h Z_\mu Z^\mu + \frac{m_Z^2}{2v^2} h^2 Z_\mu Z^\mu. \quad (2.35)$$

This shows that the Higgs field couples to the gauge bosons, which is crucial for the mass generation of the gauge bosons. Furthermore, from [Equation \(2.28\)](#) one can see that the Higgs field has cubic and quartic self-interaction terms. In the mass-diagonal basis, the kinetic Lagrangian \mathcal{L}_{kin} contains interaction terms between the W^\pm , Z and the photon fields of the form

$$W^\pm W^\mp A, \quad W^\pm W^\mp Z, \quad W^\pm W^\mp AA, \quad W^\pm W^\mp ZZ, \quad W^\pm W^\mp AZ. \quad (2.36)$$

The coupling constant e of the photon is related to the weak coupling constant g and the Weinberg angle θ_W by

$$e = g \sin \theta_W = g' \cos \theta_W. \quad (2.37)$$

Neutral-only interactions of the form ZZA , ZZZ , or similar are not present at tree level. These interactions will be discussed in more detail in [Section 2.1.3](#).

Note that the $SU(2)_W \times U(1)_Y$ gauge symmetry is never really broken during spontaneous symmetry breaking. The Higgs mechanism can also be described in a gauge invariant way, without applying the unitary gauge [27]. This can also be seen by looking at the available **d.o.f.**. Before symmetry breaking there is a complex Higgs doublet (4 real **d.o.f.**) plus 4 massless vector fields (8 **d.o.f.**), so 12 **d.o.f.** in total. After symmetry breaking, there is one real Higgs field (1 **d.o.f.**) plus three massive vector fields (9 **d.o.f.**) plus one massless vector field (2 **d.o.f.**), so again 12 **d.o.f.** in total.

The Higgs mechanism is not only responsible for the mass generation of the gauge bosons, but also for the mass generation of the fermions. The fermion mass terms are generated through Yukawa interactions with the Higgs field,

$$\mathcal{L}_{Yukawa} = - \sum_{i,j} \left(y_{ij}^u \bar{Q}_{iL} \tilde{\Phi} u_{jR} + y_{ij}^d \bar{Q}_{iL} \Phi d_{jR} + y_{ij}^l \bar{L}_{iL} \Phi l_{jR} + \text{h.c.} \right), \quad (2.38)$$

where $\tilde{\Phi} = i\sigma_2 \Phi^*$ is the conjugate Higgs doublet and y_{ij}^u , y_{ij}^d , and y_{ij}^l are the Yukawa couplings. I exclude right-handed neutrinos here. The Higgs field couples to the left-handed quark doublets Q_{iL} and lepton doublets L_{iL} , as well as to the right-handed quark singlets u_{iR} , d_{iR} and lepton singlets l_{iR} . After spontaneous symmetry breaking, the Yukawa interactions lead to mass terms for the fermions,

$$\mathcal{L}_{Yukawa} \supset - \sum_{i,j} \left(\frac{v}{\sqrt{2}} \bar{u}_{iL} y_{ij}^u u_{jR} + \frac{v}{\sqrt{2}} \bar{d}_{iL} y_{ij}^d d_{jR} + \frac{v}{\sqrt{2}} \bar{l}_{iL} y_{ij}^l l_{jR} + \text{h.c.} \right). \quad (2.39)$$

The Yukawa couplings are arbitrary complex matrices, which can be diagonalized by a bi-unitary transformation,

$$\begin{aligned} y_{ij}^u &= U_{ik}^u M_{kl}^u (K^u)_{lj}^\dagger, \\ y_{ij}^d &= U_{ik}^d M_{kl}^d (K^d)_{lj}^\dagger, \\ y_{ij}^l &= U_{ik}^l M_{kl}^l (K^l)_{lj}^\dagger, \end{aligned} \quad (2.40)$$

where U^f and K^f are unitary matrices, and M^f are diagonal matrices with the fermion masses on the diagonal, and $f = u, d, l$. The basis of the fermion fields can be freely changed, such that the Yukawa couplings are diagonal,

$$\begin{aligned} u_{iL} &\rightarrow U_{ij}^u u_{jL}, \\ d_{iL} &\rightarrow U_{ij}^d d_{jL}, \\ l_{iL} &\rightarrow U_{ij}^l l_{jL}, \\ u_{iR} &\rightarrow K_{ij}^u u_{jR}, \\ d_{iR} &\rightarrow K_{ij}^d d_{jR}, \\ l_{iR} &\rightarrow K_{ij}^l l_{jR}. \end{aligned} \quad (2.41)$$

This is known as the mass basis. In the mass basis, the Yukawa interactions become

$$\mathcal{L}_{Yukawa} \supset - \sum_i (m_i^u \bar{u}_{iL} u_{jR} + m_i^d \bar{d}_{iL} d_{jR} + m_i^l \bar{l}_{iL} l_{jR} + \text{h.c.}) , \quad (2.42)$$

where the masses of the fermions are proportional to the Yukawa couplings and the Higgs **vev**,

$$m_i^u = \frac{v}{\sqrt{2}} M_{ii}^u, \quad m_i^d = \frac{v}{\sqrt{2}} M_{ii}^d, \quad m_i^l = \frac{v}{\sqrt{2}} M_{ii}^l. \quad (2.43)$$

The interaction terms between the physical gauge bosons and the fermions can be separated into neutral currents involving A_μ and Z_μ and charged currents involving W^\pm . The neutral current interactions are given by

$$\mathcal{L}_{kin,NC} = e A_\mu J_{em}^\mu + \frac{e}{\sin \theta_W} Z_\mu J_Z^\mu. \quad (2.44)$$

The electromagnetic current J_{em}^μ is given by

$$J_{em}^\mu = \sum_f (\bar{f}_L \gamma^\mu Q f_L + \bar{f}_R \gamma^\mu Q f_R), \quad (2.45)$$

where the sum runs over all fermions. The electromagnetic current does not mix left- and right-handed fermions, since it is a vector current. The weak neutral current J_Z^μ is given by

$$J_Z^\mu = \sum_f \frac{1}{\cos \theta_W} (\bar{f}_L \gamma^\mu T_3 f_L - \sin^2 \theta_W J_{em}^\mu). \quad (2.46)$$

The transformations of [Equation \(2.41\)](#) leave the neutral current interactions invariant, making them flavour diagonal.

The charged current does not mix left- and right-handed fermions either, but it only couples to left-handed fermions. The interactions are given by

$$\mathcal{L}_{kin,CC} = \frac{e}{\sqrt{2} \sin \theta_W} [\bar{u}_{iL} \gamma^\mu V^{ij} d_{jL} W_\mu^+ + \bar{\nu}_{iL} \gamma^\mu e_{iL} W_\mu^+ + \text{h.c.}] , \quad (2.47)$$

where $V \equiv U_u^\dagger U_d$ is the **CKM** matrix. The **CKM** matrix V^{ij} is a unitary matrix, which describes the flavour-changing transitions between the different quark flavours. It can be written as

$$V = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}. \quad (2.48)$$

The **CKM** matrix is a complex unitary matrix, which can be parametrized in terms of three mixing angles and one phase. The phase is responsible for Charge-Parity (**CP**) violation in the **SM** [\[23\]](#).

For leptons, there is no equivalent of the **CKM** matrix, if neutrinos are massless. This is because one is free to rotate the neutrinos in the same way as the leptons in [Equation \(2.41\)](#).

If neutrinos are massive, charged leptons and neutrinos need to be diagonalized separately. In this case, one can define a Pontecorvo-Maki-Nakagawa-Sakata (**PMNS**) matrix, which is analogous to the **CKM** matrix for leptons [23]. The charged current interactions are only present for left-handed fermions, since the W^\pm bosons only couple to left-handed fermions. The right-handed fermions do not couple to the W^\pm bosons, but only to the neutral currents.

2.1.3 Flavour changing neutral currents

The **CKM** matrix only appears in the charged current interactions. As a result, it is not possible to have **FCNCs** at tree level in the **SM**. This is a consequence of the gauge structure of the theory: neutral currents (mediated by the Z boson, photon, and gluons) are diagonal in the mass eigenstate basis, while only charged currents (mediated by W^\pm bosons) involve the **CKM** mixing matrix [23, 24].

FCNC transitions can occur at loop level, but the loop contributions themselves are highly suppressed by the Glashow-Iliopoulos-Maiani (**GIM**) mechanism [29]. The **GIM** mechanism describes how loop amplitudes from different intermediate quarks tend to cancel each other, resulting in additional suppression proportional to mass-squared differences between the intermediate quarks [23, 25].

In this work, the focus lies on one **FCNC** transition in particular, the $b \rightarrow s\nu\bar{\nu}$ transition. There are two main contributions to this transition: the box diagram and the penguin diagram, shown in **Figure 2.1**. Both diagrams involve two flavour-changing currents and hence the interaction is additionally suppressed by a factor of $|V_{ts}V_{tb}^*|^2$ or $|V_{cs}V_{cb}^*|^2$. The transition via the u quark is even more strongly suppressed by the smallness of the **CKM** matrix elements $|V_{us}V_{ub}^*|^2$ [10].

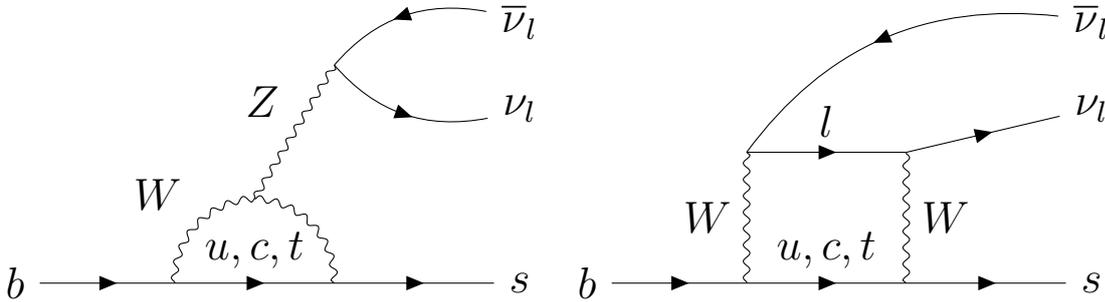


Figure 2.1: Feynman diagrams for the $b \rightarrow s\nu\bar{\nu}$ transition. The left diagram is the penguin diagram, while the right diagram is the box diagram. The W bosons are exchanged in both diagrams, while the Z^0 boson is only exchanged in the penguin diagram. The neutrino ν_l can be any of the three active neutrinos, ν_e , ν_μ , or ν_τ .

2.2 Effective field theories

The central idea of **EFTs** relevant to this work is the separation of scales: high-energy modes are removed from the spectrum, leaving an effective description in terms of light fields and a tower of higher-dimensional local operators suppressed by powers of the heavy scale Λ . This procedure preserves the symmetries of the underlying theory, most notably gauge invariance, and provides a controlled expansion in E/Λ , where E is the typical energy of the process under consideration.

These types of theories are only applicable to energies far below the scale of new physics $E \ll \Lambda$. In the high-energy, ultraviolet (**UV**) regime, the expansion parameter E/Λ becomes large and the theory loses its predictive power.

The construction of an **EFT** follows a systematic procedure [30, 31]. For a separation scale μ , one identifies the relevant light **d.o.f.** that are active (i.e. on-shell) at the energies $E < \mu$. Heavy **d.o.f.**, above the separation scale, are integrated out, generating higher-dimensional operators that encode their effects. The resulting effective Lagrangian takes the form

$$\mathcal{L}_{\text{EFT}} = \mathcal{L}_{d \leq 4} + \sum_{d > 4} \sum_i \mathcal{C}_i^{(d)}(\mu) \mathcal{O}_i^{(d)} + \text{h.c.}, \quad (2.49)$$

where $\mathcal{L}_{d \leq 4}$ contains the renormalizable terms (kinetic terms, gauge interactions, and mass terms for the light **d.o.f.**), $\mathcal{O}_i^{(d)}$ are gauge-invariant operators of mass dimension d , and $\mathcal{C}_i^{(d)}(\mu)$ are the corresponding Wilson coefficients that encode the **HEP** and depend on the separation scale. The expansion is organized by powers of $(E/\Lambda)^d$, making it a controlled approximation when $E \ll \Lambda$.

This Operator Product Expansion (**OPE**) separates the short-distance (perturbative) physics at scales higher than μ (encoded in the Wilson coefficients) from the long-distance (usually non-perturbative) physics lower than μ (encoded in the matrix elements of the operators). The renormalization scale is typically chosen to be of the order of the mass of the decaying hadron, to suppress large logarithms in perturbative calculations. For B decays this is around the b -quark mass $\mu_b \sim 4.2$ GeV.

Physical amplitudes however do not depend on the arbitrary scale μ . The μ dependence of the Wilson coefficients $\mathcal{C}_i^{(d)}(\mu)$, has to be cancelled by the matrix elements of the operators $\langle \mathcal{O}_i \rangle(\mu)$ [30]. Hence, what belongs to the Wilson coefficients and what to the matrix elements is a matter of the concrete choice of μ . This means that a change in μ implies renormalization of both Wilson coefficients and the matrix elements of the operators [30]. Since this renormalization is typically approached as a perturbative expansion in the electromagnetic and strong coupling constants, α and α_s , it typically depends on the renormalization scheme. This scheme dependence is also cancelled between the Wilson coefficients and the matrix elements.

The Wilson coefficients $\mathcal{C}_i^{(d)}(\mu)$ are connected to the full theory through a matching procedure at the scale $\mu \sim \Lambda$, where the **EFT** is required to reproduce the amplitudes of the full theory order by order in perturbation theory. These coefficients encode all effects

of the heavy **d.o.f.** and are process-independent, making the **EFT** framework particularly powerful for phenomenological applications.

A crucial aspect of **EFT** calculations is the treatment of renormalization group evolution [30]. The Wilson coefficients satisfy renormalization group equations that govern their scale dependence:

$$\mu \frac{d}{d\mu} \vec{\mathcal{C}}(\mu) = \gamma^T(\alpha_s(\mu)) \vec{\mathcal{C}}(\mu), \quad (2.50)$$

where $\gamma(\alpha_s)$ is the anomalous dimension matrix and $\vec{\mathcal{C}}(\mu)$ represents the vector of Wilson coefficients. The solution of these equations allows the evolution from the matching scale Λ down to the low-energy scale relevant for the physical process of interest.

2.2.1 Weak Effective Theory

The **WET** is a specific **EFT** that describes physics below the electroweak scale, typically around $\Lambda = M_W \sim 100$ GeV [3–5]. This is a large scale compared to the B -physics scale $\mu_b \sim 4.2$ GeV $\ll \Lambda$, which is the relevant scale for B meson decays. Thus, the weak interaction can be thought of as local at this scale, as propagators of the W and Z bosons can be approximated by a point-like interaction. This factorizes the weak interaction (encoded in the Wilson coefficients) from the electromagnetic and strong interaction (encoded in the hadronic matrix elements of the operators). In this framework, the W , Z , and Higgs bosons, along with the top quark, are integrated out as heavy **d.o.f.**. The remaining light fields are the photon, gluons, and the five light quarks (u, d, s, c, b) and three charged leptons (e, μ, τ) along with their associated neutrinos. If **BSM** physics exists at a scale of Λ or above, these interactions will also be approximately local, and hence encoded in the Wilson coefficients.

The **WET** Lagrangian is organized as an expansion in operators of increasing mass dimension:

$$\mathcal{L}_{\text{WET}} = \mathcal{L}_{\text{QED+QCD}} + \sum_{d>4} \sum_i \mathcal{C}_i^{(d)}(\mu_b) \mathcal{O}_i^{(d)} + \text{h.c.}, \quad (2.51)$$

where $\mathcal{L}_{\text{QED+QCD}}$ contains the renormalizable interactions of **QED** and **QCD** (with five active quark flavours). The dimension-five operators are responsible for neutrino masses via the Weinberg operator. For $b \rightarrow s\nu\bar{\nu}$ transitions, these operators effectively contribute at dimension-six, are generated only at loop level, and are strongly experimentally constrained [32, 33], leading to further suppression. These are therefore typically neglected. Quark flavour-changing interactions first appear at dimension-six, leading to corrections to **SM** processes, and are the primary focus in most phenomenological studies. Since the mass of the initial on-shell B meson limits the maximum momentum transfer in this process, the matrix elements of operators with mass-dimension seven or above are suppressed by at least a factor of $M_B/M_W \simeq 0.06$, which are hence commonly neglected in these types of analyses.

For flavour-changing neutral current processes like $b \rightarrow s\nu\bar{\nu}$, it suffices to discuss the $sb\nu\bar{\nu}$ sector of the **WET** [34]. It is spanned by a subset of local operators of mass-dimension six,

which is closed under the renormalization group to leading order in the Fermi constant G_F [35]. The corresponding Lagrangian density for the $sb\nu\nu$ sector reads [33]³

$$\mathcal{L}_{\text{WET}} \supset -\frac{4G_F}{\sqrt{2}} \frac{\alpha}{2\pi} V_{ts}^* V_{tb} \sum_i C_i(\mu_b) \mathcal{O}_i + \text{h.c.}, \quad (2.52)$$

with α the fine structure constant. Due to the normalization in terms of G_F the newly defined Wilson coefficients

$$C_i(\mu_b) = \frac{\sqrt{2}}{4G_F} \frac{2\pi}{\alpha} \frac{1}{V_{ts}^* V_{tb}} \mathcal{C}_i(\mu_b) \quad (2.53)$$

are dimensionless. They are generally complex-valued and defined in the modified minimal subtraction ($\overline{\text{MS}}$) scheme [33].

Assuming massless neutrinos, the full set of dimension-six operators for the $sb\nu\nu$ sector is given by [33],

$$\begin{aligned} \mathcal{O}_{\text{VL}} &= (\overline{\nu}_L \gamma_\mu \nu_L) (\overline{s}_L \gamma^\mu b_L) \\ \mathcal{O}_{\text{VR}} &= (\overline{\nu}_L \gamma_\mu \nu_L) (\overline{s}_R \gamma^\mu b_R) \\ \mathcal{O}_{\text{SL}} &= (\overline{\nu}_L^c \nu_L) (\overline{s}_R b_L) \\ \mathcal{O}_{\text{SR}} &= (\overline{\nu}_L^c \nu_L) (\overline{s}_L b_R) \\ \mathcal{O}_{\text{TL}} &= (\overline{\nu}_L^c \sigma_{\mu\nu} \nu_L) (\overline{s}_R \sigma^{\mu\nu} b_L), \end{aligned} \quad (2.54)$$

with $\nu_L^c \equiv C \overline{\nu}_L^T$ and where $C = i\gamma^2 \gamma^0$ is the charge conjugation operator and $\sigma^{\mu\nu} \equiv \frac{i}{2} [\gamma^\mu, \gamma^\nu]$. In the above, the subscripts V, S, T represent vector, scalar, and tensor operators, respectively. The subscripts L, R denote the chirality of the quark current. The operators are defined as sums over the neutrino flavours, since this is a property that one currently cannot determine experimentally. If one assumes the existence of only left-handed massless neutrinos, all operators except VL and VR vanish.

In the **SM**, only the vector operator \mathcal{O}_{VL} receives a non-zero Wilson coefficient at the matching scale [36, 37],

$$C_{\text{VL}}^{\text{SM}} = 6.6 \pm 0.1, \quad C_i^{\text{SM}} = 0 \quad \text{for } i \neq \text{VL}. \quad (2.55)$$

This is due to the fact that the charged current interactions in the **SM** only couple to left-handed fermions, and hence only generate left-handed vector operators at low energies. New physics beyond the **SM** can modify these Wilson coefficients, potentially generating non-zero contributions to the scalar and tensor operators or modifying the vector coefficients.

2.3 Hadronic matrix elements

Having established the **WET** framework and the relevant dimension-six operators for $b \rightarrow s\nu\bar{\nu}$ transitions in Section 2.2.1, I now turn to the evaluation of the corresponding

³The superscript (6) on the operators and Wilson coefficients is dropped for brevity.

matrix elements. These matrix elements represent the second crucial component of the factorized amplitude structure discussed in [Section 2.2](#), where the short-distance physics (encoded in the Wilson coefficients) is separated from the long-distance, non-perturbative [QCD](#) dynamics.

These are matrix elements of the effective operators between initial and final quark states. Since quarks are bound inside hadrons, one needs to evaluate the matrix elements between hadronic states, e.g. the initial B meson and final $K^{(*)}$ meson. This gives rise to the so-called hadronic matrix elements, which, due to their hadronic nature, contain the non-perturbative [QCD](#) dynamics.

The evaluation of these matrix elements is inherently non-perturbative since it involves the binding of quarks into hadrons at energy scales where $\alpha_s \sim 1$. This necessitates the use of non-perturbative techniques such as lattice [QCD](#) calculations. The discussion of lattice [QCD](#) calculations is beyond this work, for further details see for example Reference [\[38\]](#).

Since the hadronic matrix elements are Lorentz tensors, they can be decomposed into Lorentz-invariant form factors multiplied by Lorentz structures that depend on the momenta and spin configurations of the initial and final states. The form factor approach provides a systematic way to parametrize these matrix elements in terms of Lorentz-invariant functions that depend only on the momentum transfer q^2 .

For $B \rightarrow K$ transitions, the vector operators in [Equation \(2.54\)](#) \mathcal{O}_{VL} and \mathcal{O}_{VR} give rise to matrix elements of the form [\[33, 39\]](#)

$$\langle K(k) | \bar{s} \gamma^\mu b | B(p) \rangle = f_+(q^2) \left[(p+k)^\mu - \frac{M_B^2 - M_K^2}{q^2} q^\mu \right] + f_0(q^2) \frac{M_B^2 - M_K^2}{q^2} q^\mu, \quad (2.56)$$

where $q = p - k$ is the momentum transfer, $q^2 = (p - k)^2$ is the invariant mass squared of the dineutrino system, and $f_+(q^2)$ and $f_0(q^2)$ are the vector form factors, satisfying the constraint $f_+(0) = f_0(0)$. M_B and M_K are the masses of the B meson and kaon, respectively.

The scalar operators \mathcal{O}_{SL} and \mathcal{O}_{SR} involve matrix elements of scalar currents [\[33, 39\]](#):

$$\langle K(k) | \bar{s} b | B(p) \rangle = f_0(q^2) \frac{M_B^2 - M_K^2}{m_b - m_s}, \quad (2.57)$$

where m_b and m_s are the b and s quark masses, respectively. Interestingly, this matrix element is also proportional to $f_0(q^2)$, which connects the scalar and vector sectors.

The tensor operator \mathcal{O}_{TL} requires the evaluation of [\[33, 39\]](#)

$$\langle K(k) | \bar{s} \sigma^{\mu\nu} b | B(p) \rangle = f_T(q^2) \frac{i}{M_B + M_K} [(p+k)^\mu q^\nu - q^\mu (p+k)^\nu], \quad (2.58)$$

which defines the tensor form factor $f_T(q^2)$.

Pseudoscalar, axial vector, and pseudotensor currents do not contribute to $B \rightarrow K$ transitions, by virtue of the symmetries of [QCD](#). In particular, the invariance of the strong

interaction under C , P , and CP transformations implies that these $B \rightarrow K$ matrix elements vanish.

For $B \rightarrow K^*$ transitions involving a vector meson in the final state, the parametrization is more complex due to the additional polarization **d.o.f.** of the K^* meson [33, 39]. The vector operators give rise to four independent form factors: $V(q^2)$, $A_0(q^2)$, $A_1(q^2)$, $A_2(q^2)$. The scalar operators contribute through the pseudoscalar current, proportional to $A_0(q^2)$, while the scalar current vanishes by Lorentz structure constraints. The tensor operators introduce three additional form factors $T_1(q^2)$, $T_2(q^2)$, and $T_3(q^2)$, with the constraint $T_1(0) = T_2(0)$. It is common to replace the form factors $A_2(q^2)$ and $T_3(q^2)$ with the linear combinations $A_{12}(q^2)$ and $T_{23}(q^2)$, defined in Equations (A.30) and (A.31). The complete parametrization for all operator types is given in Appendix A.3.

The increased complexity of $B \rightarrow K^*$ transitions (seven independent form factors compared to three for $B \rightarrow K$) reflects the richer kinematic structure available when the final state carries spin, providing enhanced sensitivity to new physics through different angular distributions and q^2 dependencies.

The form factors $f_+(q^2)$, $f_0(q^2)$, $f_T(q^2)$ for $B \rightarrow K$ transitions and $V(q^2)$, $A_0(q^2)$, $A_1(q^2)$, $A_2(q^2)$, $T_1(q^2)$, $T_2(q^2)$, $T_3(q^2)$ for $B \rightarrow K^*$ transitions are non-perturbative objects that must be determined using lattice QCD calculations. In the past, lattice QCD calculations were limited to high $q^2 \gtrsim 10 \text{ GeV}^2$ due to computational limitations, such as finite volume effects, but recently for the $B \rightarrow K^{(*)}$ form factors these limitations were overcome and lattice QCD calculations down to $q^2 = 0$ have become available. The results from these two methods can be combined using a suitable series expansion. In this work, form factors are parametrized using the Bourrely-Caprini-Lellouch (BCL) [40] and the Bharucha-Straub-Zwicky (BSZ) [41] parametrizations, which respect analyticity and unitarity constraints. These parametrizations provide a model-independent approach to form factor extrapolation, as they are based on fundamental principles of quantum field theory rather than specific assumptions about the underlying hadronic dynamics.

The series expansion is typically performed in terms of a conformal variable

$$z(q^2, t_0) = \frac{\sqrt{t_+ - q^2} - \sqrt{t_+ - t_0}}{\sqrt{t_+ - q^2} + \sqrt{t_+ - t_0}}, \quad (2.59)$$

where $t_+ = (M_B + M_{K^{(*)}})^2$ and t_0 is a free parameter. This variable z maps the physical q^2 region onto a small interval, improving the convergence of the series expansion.

The BSZ parametrization expresses each form factor as a series expansion in the conformal variable $z(q^2, t_0)$ [41],

$$f(q^2) = \frac{1}{1 - q^2/m_{\text{res}}^2} \sum_{k=0}^K a_k^{(f)} [z(q^2, t_0) - z(0, t_0)]^k, \quad (2.60)$$

where m_{res} is the mass of the lightest resonance with appropriate quantum numbers, using $t_0 \equiv t_+(1 - \sqrt{1 - t_-/t_+})$ with $t_- = (M_B - M_{K^{(*)}})^2$, and the coefficients a_k are fit parameters determined from lattice QCD data. The series is typically truncated at order $K = 2$

Table 2.2: The input parameters used for calculating predictions. The parameters are taken from References [10, 37, 42, 43]. Here α is evaluated at the B -physics scale μ_b . The running of CKM elements is neglected, as their scale dependence is very mild.

| | |
|-----------------------------|----------------------|
| $C_{\text{VL}}^{\text{SM}}$ | 6.6 ± 0.1 |
| $1/\alpha(\mu_b)$ | 133.00 ± 0.15 |
| $ V_{tb}V_{ts}^* $ | 0.041 ± 0.001 |
| τ_{B^0} | 1.519 ± 0.005 ps |
| τ_{B^+} | 1.641 ± 0.004 ps |

or $K = 3$, limiting the number of parameters. The **BSZ** parametrization incorporates known pole structures and ensures proper analytical behaviour, making it particularly suitable for extrapolating lattice **QCD** results from high q^2 to the full kinematic range. Unitarity constraints can be imposed to further reduce the number of free parameters and improve the stability of extrapolations [41].

The **BCL** parametrization follows a similar approach but emphasizes the implementation of unitarity constraints more systematically [40]. In the **BCL** framework, the form factors are expressed as

$$f(q^2) = \frac{1}{1 - q^2/m_{\text{res}}^2} \sum_{k=0}^{K-1} b_k \left[z(q^2, t_0)^k - (-1)^{(k-K)} \frac{k}{K} z(q^2, t_0)^K \right], \quad (2.61)$$

where the coefficients b_k are constrained by unitarity bounds that limit their allowed values. These constraints reduce the parameter space and improve the stability of extrapolations.

2.4 Theory predictions for $B \rightarrow K^{(*)}\nu\bar{\nu}$

The theoretical prediction for $B \rightarrow K^{(*)}\nu\bar{\nu}$ decay rates requires combining the **WET** framework with precise determinations of the hadronic form factors. While one cannot achieve a completely model-independent theoretical description, it is nevertheless possible to capture the effects of a large number of **BSM** theories under mild assumptions. Here, one assumes that potential new **BSM** particles and force carriers have masses at or above the scale of electroweak symmetry breaking, making the **WET** description appropriate. All relevant numerical inputs to predictions in this work are summarized in **Table 2.2**.

2.4.1 Observables for $B \rightarrow K\nu\bar{\nu}$

Since the B meson is a pseudoscalar, the decay is isotropic in the B meson rest frame, and the only kinematically free variable is the squared dineutrino invariant mass q^2 . The kinematic range extends from $q_{\text{min}}^2 = 0$ (assuming massless neutrinos) to $q_{\text{max}}^2 = (M_B - M_K)^2$.

The decay $B \rightarrow K\nu\bar{\nu}$ is governed by the **WET** Lagrangian as described by [Equations \(2.52\)](#) and [\(2.54\)](#). The differential decay rate for $B \rightarrow K\nu\bar{\nu}$ in terms of the **WET** Wilson coefficients is given by [\[33, 34, 44\]](#)

$$\begin{aligned} \frac{d\Gamma}{dq^2} = & 3 \left(\frac{4G_F}{\sqrt{2}} \frac{\alpha}{2\pi} \right)^2 |V_{ts}^* V_{tb}|^2 \frac{\sqrt{\lambda_{BK} q^2}}{(4\pi)^3 M_B^3} \\ & \cdot \left[\frac{\lambda_{BK}}{24q^2} |f_+(q^2)|^2 |C_{VL} + C_{VR}|^2 \right. \\ & + \frac{(M_B^2 - M_K^2)^2}{8(m_b - m_s)^2} |f_0(q^2)|^2 |C_{SL} + C_{SR}|^2 \\ & \left. + \frac{2\lambda_{BK}}{3(M_B + M_K)^2} |f_T(q^2)|^2 |C_{TL}|^2 \right], \end{aligned} \quad (2.62)$$

where M_B and M_K are the masses of the B meson and kaon, respectively, $m_b(\mu_b)$ and $m_s(\mu_b)$ are the $\overline{\text{MS}}$ scheme quark masses evaluated at the B -physics scale μ_b , and $\lambda_{BK} \equiv \lambda(M_B^2, M_K^2, q^2)$ is the Källén function

$$\lambda(a, b, c) = a^2 + b^2 + c^2 - 2ab - 2ac - 2bc. \quad (2.63)$$

As can be inferred from [Equation \(2.62\)](#), the decay is only sensitive to the magnitude of three linear combinations of Wilson coefficients:

$$|C_{VL} + C_{VR}|, \quad |C_{SL} + C_{SR}|, \quad |C_{TL}|. \quad (2.64)$$

This represents a reduction from the 5-dimensional Wilson coefficient space to a 3-dimensional observable space, highlighting the limited sensitivity of current measurements to the full parameter space of the **WET**.

The branching ratio, \mathcal{B} , is connected to the decay rate simply through $\mathcal{B} = \tau\Gamma$, where τ is the decay lifetime [\[23\]](#). To highlight the individual contributions of vector, scalar, and tensor terms from [Equation \(2.62\)](#) to the differential branching ratio, an illustration where individual left-handed Wilson coefficients are set to unity is shown in [Figure 2.2](#). The values for the corresponding 8 hadronic parameters are extracted from a joint theoretical prior Probability Density Function (**PDF**) comprised of the 2021 lattice world average based on results by the Fermilab/MILC and HPQCD collaborations [\[45\]](#), and recent results by the HPQCD collaboration [\[46\]](#) and are reported in [Table 2.3](#).

2.4.2 Observables for $B \rightarrow K^*\nu\bar{\nu}$

The pseudoscalar B meson decays to the vector meson K^* — this decay is isotropic in the B meson rest frame. However, the K^* meson is an unstable resonance that subsequently decays to $K\pi$. The decay $K^* \rightarrow K\pi$ introduces an angular dependence on the angle between the direction of flight of the B and the K in the K^* rest frame. This means that the decay rate in [Equation \(2.65\)](#) exhibits an angular dependence. However, here I

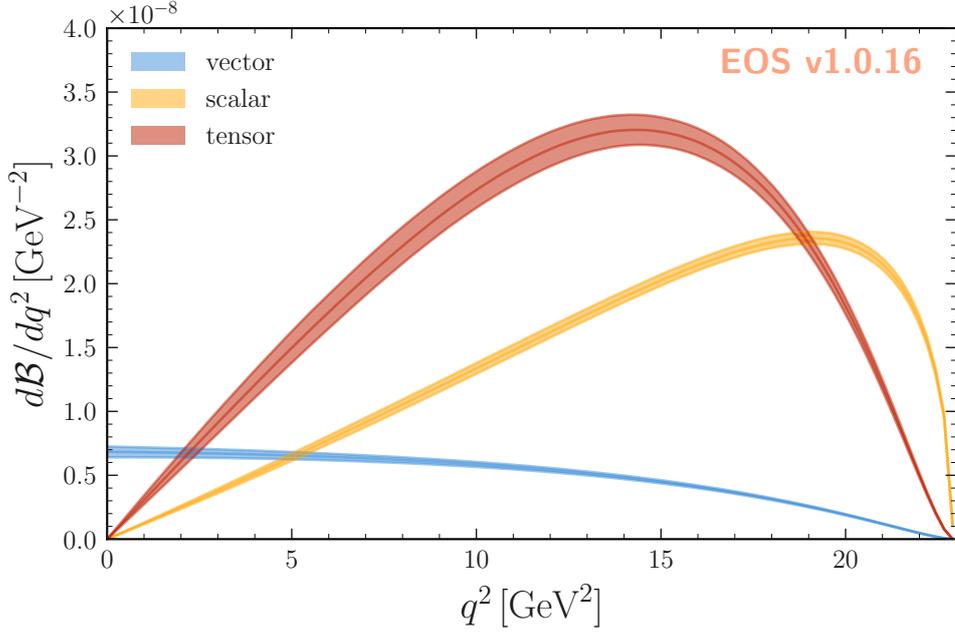


Figure 2.2: Illustration of the variety of shapes of the $B \rightarrow K\nu\bar{\nu}$ branching ratio due to purely vector, scalar, or tensor interactions. Each curve corresponds to setting a single non-zero Wilson Coefficient in Equation (2.62) to unity while keeping all other coefficients at zero. These predictions are obtained using the EOS software [37, 47].

integrate out this angular dependence, leaving q^2 as the only kinematically free variable. The kinematic range extends from $q_{\min}^2 = 0$ (assuming massless neutrinos) to $q_{\max}^2 = (M_B - M_{K^*})^2$.

The decay $B \rightarrow K^*\nu\bar{\nu}$ is governed by the WET Lagrangian as described by Equations (2.52) and (2.54). Its differential decay rate reads [33, 34, 44]

$$\frac{d\Gamma}{dq^2} = 3 \left(\frac{4G_F \alpha}{\sqrt{2} 2\pi} \right)^2 |V_{ts}^* V_{tb}|^2 \frac{\sqrt{\lambda_{BK^*} q^2}}{(4\pi)^3 M_B^3} \cdot [|\mathcal{A}_V|^2 |C_{VL} + C_{VR}|^2 + |\mathcal{A}_A|^2 |C_{VL} - C_{VR}|^2 + |\mathcal{A}_P|^2 |C_{SR} - C_{SL}|^2 + |\mathcal{A}_T|^2 |C_{TL}|^2], \quad (2.65)$$

where the reduced amplitudes multiplying the Wilson coefficients read

$$\begin{aligned} |\mathcal{A}_V|^2 &= \frac{\lambda_{BK^*} |V(q^2)|^2}{12 (M_B + M_{K^*})^2}, \\ |\mathcal{A}_A|^2 &= \frac{8M_B^2 M_{K^*}^2}{3q^2} |A_{12}(q^2)|^2 + \frac{(M_B + M_{K^*})^2 |A_1(q^2)|^2}{12}, \\ |\mathcal{A}_P|^2 &= \frac{\lambda_{BK^*}}{8(m_b + m_s)^2} |A_0(q^2)|^2, \\ |\mathcal{A}_T|^2 &= \frac{32M_B^2 M_{K^*}^2 |T_{23}(q^2)|^2}{3(M_B + M_{K^*})^2} + \frac{4\lambda_{BK^*} |T_1(q^2)|^2}{3q^2} + \frac{4(M_B^2 - M_{K^*}^2)^2 |T_2(q^2)|^2}{3q^2}. \end{aligned} \quad (2.66)$$

Table 2.3: The hadronic parameters used to parametrize the $f_+(q^2)$, $f_0(q^2)$ and $f_T(q^2)$ form factors in the **BSZ** parametrization [48], based on results by the Fermilab/MILC and HPQCD collaborations [45, 46]. The table shows the mean values (first row) and the covariance matrix (upper triangular part).

| $a_0^{0/+}$ | a_1^+ | a_2^+ | a_1^0 | a_2^0 | a_0^T | a_1^T | a_2^T |
|-------------|---------|---------|---------|---------|---------|---------|---------|
| 0.3380 | -0.8845 | -0.1109 | 0.3742 | 0.0787 | 0.3269 | -0.9625 | -0.2780 |
| 0.0001 | 0.0007 | 0.0010 | 0.0008 | 0.0013 | 0.0001 | 0.0006 | 0.0008 |
| | 0.0093 | 0.0216 | 0.0057 | 0.0102 | 0.0004 | 0.0062 | 0.0147 |
| | | 0.0672 | 0.0085 | 0.0171 | 0.0001 | 0.0112 | 0.0425 |
| | | | 0.0062 | 0.0113 | 0.0006 | 0.0047 | 0.0066 |
| | | | | 0.0220 | 0.0010 | 0.0081 | 0.0128 |
| | | | | | 0.0005 | 0.0034 | 0.0058 |
| | | | | | | 0.0319 | 0.0619 |
| | | | | | | | 0.1441 |

One readily finds that the dependence of the observables on the Wilson coefficients is very different in [Equations \(2.62\)](#) and [\(2.65\)](#), respectively. Compared to $B \rightarrow K\nu\bar{\nu}$ decays, the differential $B \rightarrow K^*\nu\bar{\nu}$ decay rate exhibits additional sensitivity to the quantities

$$|C_{\text{VL}} - C_{\text{VR}}|, |C_{\text{SL}} - C_{\text{SR}}|. \quad (2.67)$$

As a consequence, a simultaneous analysis of both decays allows constraining a total of five real-valued (out of ten total real-valued) parameters in the $sb\nu\nu$ sector. Assuming all **WET** Wilson coefficients to be real-valued, this corresponds to constraining the magnitudes of all Wilson coefficients.

To understand the individual contributions to the differential decay rate in [Equation \(2.65\)](#) by vector, scalar, and tensor operators, I provide an illustration of their relative sizes and their shapes in [Figure 2.3](#). This is achieved by setting their respective Wilson coefficients to unity. The values for the corresponding 19 hadronic parameters arise from the Gaussian likelihood provided in Reference [49] and are reported in [Table A.1](#) of [Appendix A.4](#).

2.4.3 The EOS software package

New physics contributions can significantly modify the $B \rightarrow K^{(*)}\nu\bar{\nu}$ prediction by introducing non-zero Wilson coefficients for the scalar and tensor operators or by modifying the vector coefficients. The shape of the q^2 spectrum also provides additional discriminating power between different new physics scenarios, as vector, scalar, and tensor interactions lead to characteristically different kinematic distributions.

The **EOS** software package [37, 47] provides a comprehensive framework for computing theoretical predictions for $B \rightarrow K^{(*)}\nu\bar{\nu}$ and many other processes. It incorporates the **WET** framework, allowing users to specify Wilson coefficients at a chosen scale and compute the corresponding decay rates and distributions.

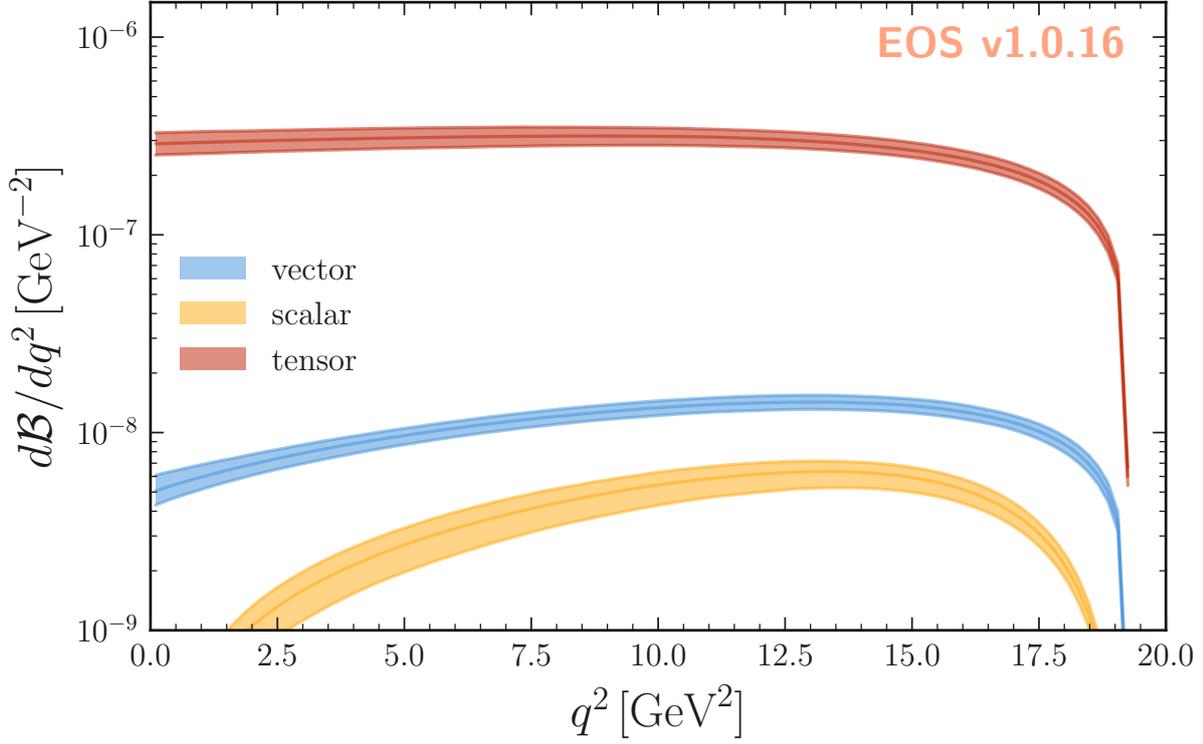


Figure 2.3: Illustration of the variety of shapes of the $B \rightarrow K^* \nu \bar{\nu}$ branching ratio due to purely vector, scalar, or tensor interactions. Each curve corresponds to setting a single (left-handed) non-zero Wilson coefficient in Equation (2.65) to unity while keeping all other coefficients at zero. These predictions are obtained using the EOS software [37, 47]. The values for the corresponding 19 hadronic parameters arise from the Gaussian likelihood provided in Reference [49].

For the $sb\nu\nu$ sector, EOS provides implementations of all five dimension-six operators in Equation (2.54), allowing for arbitrary values of the Wilson coefficients C_{VL} , C_{VR} , C_{SL} , C_{SR} , and C_{TL} . This flexibility makes it particularly suitable for exploring BSM scenarios that deviate from the SM prediction in Equation (2.55).

For $B \rightarrow K^{(*)} \nu \bar{\nu}$ decays EOS can compute a wide range of observables:

Integrated observables:

- Total branching ratios $\mathcal{B}(B \rightarrow K^{(*)} \nu \bar{\nu})$
- Partial branching ratios in specified q^2 bins
- Ratios of branching ratios $\mathcal{R}_{K^*} = \mathcal{B}(B \rightarrow K^* \nu \bar{\nu}) / \mathcal{B}(B \rightarrow K \nu \bar{\nu})$
- Longitudinal polarization fraction F_L for $B \rightarrow K^* \nu \bar{\nu}$

Differential observables:

- Differential branching ratios $d\mathcal{B}/dq^2$
- Normalized q^2 spectra and shape parameters
- Angular distributions for $B \rightarrow K^*\nu\bar{\nu}$

The package incorporates many hadronic form factor results from different lattice QCD collaborations. These form factors are implemented using multiple different parametrizations, with full covariance matrices to properly account for correlations between different form factors.

Furthermore, it is possible to combine different likelihoods, for example from different lattice QCD collaborations, in a statistically consistent manner. This, in turn, leads to more precise determinations of the form factors and hence more accurate predictions for the decay rates.

The package is actively maintained and regularly updated to incorporate new lattice QCD results, improved theoretical calculations, and additional BSM models. All calculations are documented with complete references to the underlying theoretical framework, ensuring reproducibility and traceability of results. The implementation of Equation (2.62) and partially Equation (2.65) was one of my contributions to the EOS package.

This comprehensive implementation makes EOS a valuable tool for precision flavour phenomenology, enabling both theoretical predictions and experimental interpretation within a unified, well-validated framework.

CHAPTER 2. THEORETICAL BACKGROUND

Chapter 3

The Belle II experiment

Experimental validation of theoretical predictions requires sophisticated experimental facilities capable of producing and studying rare processes with high precision. This chapter presents the Belle II experiment and its underlying infrastructure, which provides the experimental data foundation for all the reinterpretation studies discussed in subsequent chapters.

The Belle II experiment, operating at the SuperKEKB electron-positron collider in Japan, is a B factory designed to study flavour physics at unprecedented precision. Building upon the success of its predecessor Belle, Belle II aims to collect 50 times more data, enabling searches for rare processes with branching ratios as small as 10^{-6} or below. This sensitivity is essential for studying flavour-changing neutral current processes like $B \rightarrow K^{(*)}\nu\bar{\nu}$.

The Belle II experiment offers unique capabilities for studying rare B meson decays that are sensitive to **BSM** physics. Operating at the $\Upsilon(4S)$ resonance, which decays dominantly to $B\bar{B}$ pairs, provides a clean experimental environment with well-defined initial conditions and low combinatorial backgrounds compared to hadron colliders. Understanding the experimental setup, data collection methods, and analysis techniques is essential for appreciating both the complexity of extracting rare signals from collision data and the opportunities for model-agnostic reinterpretation that preserve this experimental information.

This chapter establishes the experimental foundation for the measurements and reinterpretations presented in subsequent chapters. **Section 3.1** introduces B factories as precision intensity-frontier experiments. **Section 3.2** describes the SuperKEKB accelerator. **Section 3.3** presents the Belle II detector system, covering the tracking detectors (**Sections 3.3.1 to 3.3.3**), particle identification systems (**Section 3.3.4**), electromagnetic calorimeter (**Section 3.3.5**), K_L^0 and muon detection (**Section 3.3.6**), and the trigger system (**Section 3.3.7**). Finally, **Section 3.4** discusses the computing and software infrastructure, including event reconstruction (**Section 3.4.1**) and Monte Carlo (MC) simulation (**Section 3.4.2**) that enables both data processing and physics analysis.

3.1 B factories

B factories are high-luminosity particle colliders specifically designed to produce and study B mesons, which are composed of a bottom b quark and a lighter u or d quark. In these factories, B mesons are produced by colliding electrons and positrons at a centre-of-mass energy of around 10.58 GeV, to produce $\Upsilon(4S)$ mesons at the highest possible rate. These decay into $B\bar{B}$ meson pairs with a branching ratio of $\mathcal{B} > 0.96$ at 95% confidence level [10].

Whereas experiments at the Large Hadron Collider (LHC) probe the outcomes of proton-proton collisions at the energy frontier — with large centre-of-mass energies of up to 14 TeV — B factories are optimized for the intensity frontier — producing large, clean data sets with a well-defined initial state. This enables precision studies of rare and suppressed flavour processes, offering high sensitivity to indirect effects of BSM physics.

The original motivation for B factories was to test the mechanism of CP violation in the B meson system, as predicted by the CKM matrix. Two B factories, Belle and BaBar, independently observed this prediction for the first time [50, 51].

The successor of Belle, Belle II, now targets an even deeper understanding of flavour physics [52]: *Are there new CP -violating phases in the quark sector?; Does nature have a left-right symmetry, and are there BSM FCNC?; Are there sources of BSM lepton flavour violation?.* The experiment also addresses questions that go beyond the realm of flavour physics, such as: *Is there a dark sector of particle physics at the same mass scale as ordinary matter?.*

A key capability of B factories is their ability to study time-dependent CP violation through the analysis of entangled $B^0\bar{B}^0$ pairs produced in $\Upsilon(4S)$ decays. In the neutral B meson system, the flavour eigenstates and mass eigenstates are not identical, leading to quantum oscillations between B^0 and \bar{B}^0 . This mixing proceeds via flavour-changing neutral current box diagrams (see Section 2.1.3). Because the $B^0\bar{B}^0$ pair is produced in a coherent quantum state, the decay of one meson at a given time projects the other into a definite flavour state. By reconstructing the flavour of both mesons and measuring the time difference between their decays, one can extract time-dependent decay rates. These measurements provide access to both the magnitudes and phases of CKM matrix elements, including the CP -violating phase [24, 52]. Given the short lifetimes of neutral B mesons, such studies are only possible at asymmetric-energy colliders, where the boost of the $\Upsilon(4S)$ along the beam line enables resolution of the two decay vertices.

3.2 Accelerator: SuperKEKB

SuperKEKB is a high-luminosity, asymmetric-energy electron-positron collider located at the High-Energy Accelerator Research Organization (KEK) in Tsukuba, Japan. It operates at a centre-of-mass energy corresponding to the mass of the $\Upsilon(4S)$ resonance. The design luminosity of SuperKEKB is $6 \cdot 10^{35} \text{ cm}^{-2} \text{ s}^{-1}$ — about 30 times higher than that achieved by its predecessor, KEKB [53–55].

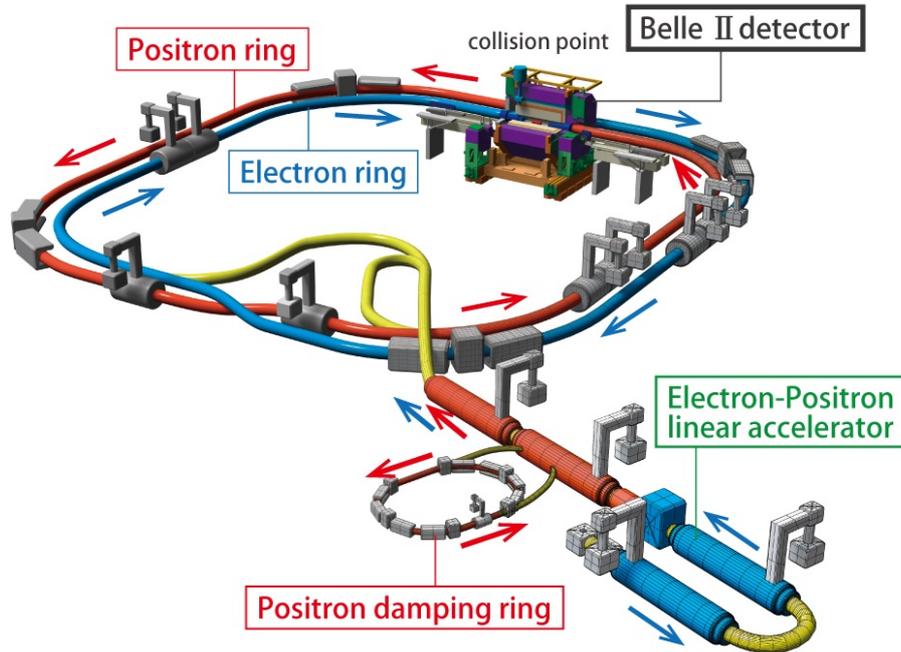


Figure 3.1: A schematic view of the SuperKEKB collider complex. Electron and positron rings intersect at the Belle II detector [53].

At SuperKEKB, electrons and positrons are generated and accelerated through a multi-stage process before colliding at the interaction point within the Belle II detector. Electrons are produced by thermionic emission from an electron gun and are directly injected into the High-Energy Ring (HER) at 7 GeV. Positrons are created when high-energy electrons strike a tungsten target, producing electron-positron pairs. The resulting positron beam has a large emittance — the particles are spread out in position and angle. To reduce this, positrons are passed through a damping ring, where synchrotron radiation reduces their emittance, making the beam narrower and more collimated. The cooled positrons are then injected into the Low Energy Ring (LER) at 4 GeV. Both beams circulate in opposite directions, steered and focused by magnets and accelerated by radiofrequency (RF) cavities, and eventually collide at the Belle II detector’s interaction point. An illustration is shown in Figure 3.1.

The asymmetric beam energies lead to a boost of the centre-of-mass system, which allows for measuring the time-dependent CP symmetry violation by creating a displacement of decay vertices [52].

3.3 The Belle II detector

The Belle II detector is a large-acceptance magnetic spectrometer designed for precision studies of flavour physics in e^+e^- collisions at the SuperKEKB collider in Japan [56, 57]. It is optimized for measurements in the B , D , and τ sectors and searches for BSM physics,

designed to operate at high luminosities and background conditions.

The detector is composed of several subdetectors arranged in a cylindrical geometry around the interaction point, see [Figure 3.2](#). Starting from the innermost layer and moving outward, the tracking system includes a high-resolution Pixel Vertex Detector ([PXD](#)) ([Section 3.3.1](#)) and Silicon Vertex Detector ([SVD](#)) ([Section 3.3.2](#)), which together enable precise vertex reconstruction. These are followed by a large Central Drift Chamber ([CDC](#)) ([Section 3.3.3](#)) that measures charged particle momenta and provides energy loss information for particle identification.

Surrounding the tracking system are specialized detectors for particle identification ([Section 3.3.4](#)): the Time-Of-Propagation counter ([TOP](#)) in the barrel region and the Aerogel Ring-Imaging Cherenkov detector ([ARICH](#)) in the forward endcap. The Electromagnetic Calorimeter ([ECL](#)) ([Section 3.3.5](#)), measures photon and electron energies and contributes to event triggering and particle identification. The entire system is enclosed in a 1.5 T superconducting solenoid magnet, providing the magnetic field necessary for charged particle momentum measurements. The outermost component, the K_L^0 and muon detector ([KLM](#)) ([Section 3.3.6](#)), identifies muons and neutral kaons that interact in iron absorber layers, which simultaneously serve as the return yoke for the superconducting magnet.

3.3.1 Pixel vertex detector

The Belle II [PXD](#) comprises the innermost two layers of the Belle II tracking system and is essential for precisely reconstructing the decay vertices of short-lived particles, especially B mesons [[56](#), [57](#)]. Accurate vertex determination is critical for studies of CP violation and rare decays.

The [PXD](#) uses DEPFET (DEPLETED Field Effect Transistor) pixel sensors, which combine charge collection and signal amplification in each pixel. This design enables a high signal-to-noise ratio and a minimal material thickness, reducing multiple scattering and preserving tracking precision.

Structured as two cylindrical layers located approximately 14 mm and 22 mm from the interaction point, the [PXD](#) is positioned to detect charged particles as early as possible. This proximity maximizes vertex resolution but also exposes the detector to high background rates and radiation levels. These challenges are mitigated through the use of finely segmented pixels, fast readout electronics, and radiation-hardened construction.

As the luminosity of the accelerator increases, the [PXD](#) must handle higher data rates and occupancy. To manage this in the future, a data reduction scheme will use information from the outer tracking detectors to define Regions Of Interest ([ROIs](#)). Only pixel data within these [ROIs](#) are processed, allowing the system to retain essential tracking information while significantly reducing data throughput.

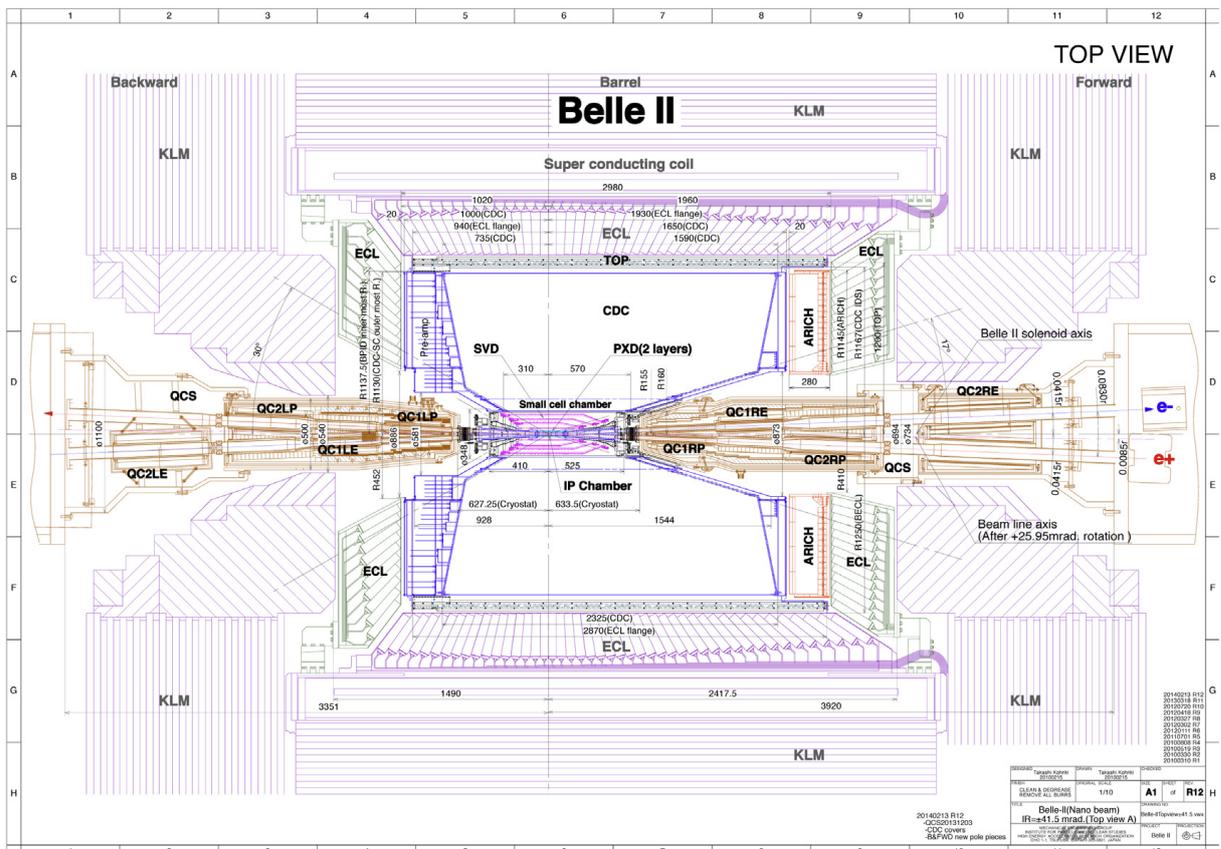


Figure 3.2: A technical design drawing of the Belle II detector [56].

3.3.2 Silicon vertex detector

The **SVD** is the second-innermost component of the Belle II tracking system, surrounding the **PXD** and contributing to precise reconstruction of charged particle trajectories and decay vertices [56, 57]. Together with the **PXD** (Section 3.3.1) and **CDC** (Section 3.3.3), it forms the core of the Belle II tracking system.

The **SVD** consists of four layers of double-sided silicon strip sensors positioned at radii between 38 mm and 135 mm. Each sensor provides 2-dimensional hit information via orthogonal strip readouts, enabling accurate three-dimensional track reconstruction. This complements the high-resolution but localized information from the **PXD**.

With its large acceptance, the **SVD** improves tracking efficiency and vertexing, especially for low-momentum particles and decay chains with displaced vertices. To minimize multiple scattering, the detector is built with a low material budget. Operating in a high-radiation environment, it employs radiation-tolerant electronics. The **SVD** also plays a key role in the **ROI** determination that reduces **PXD** data volume in real time.

3.3.3 Central drift chamber

The **CDC** is the primary tracking detector of Belle II [56, 57]. It reconstructs charged tracks and measures their momenta, provides particle identification via measuring the energy loss, enables identification of low-momentum particles that do not reach outer detectors, and supplies reliable trigger signals for charged particles.

To satisfy these requirements, the **CDC** is designed with large geometrical acceptance, fine spatial resolution, and allows for stable operation under high-luminosity conditions. It consists of a cylindrical volume filled with a low-mass helium-ethane gas mixture to minimize multiple scattering, and incorporates sense wires arranged in alternating axial and stereo layers of wires. Axial wires are aligned parallel to the beam axis and provide measurements in the transverse plane, while stereo wires are tilted at a small angle and supply position information along the beam direction.

As charged particles pass through the **CDC**, they ionize gas molecules, leaving behind a trail of ionization electrons along their paths. These electrons drift toward thin sense wires held at high voltage, under the influence of an electric field. When the electrons reach a sense wire, they create a measurable electrical signal. From the arrival time of these signals and the drift velocity of the electrons in the gas, the distance of closest approach between the particle and the wire can be determined. With many such measurements across multiple layers of wires, the full trajectory of a charged particle can be reconstructed. The particle momentum and charge is inferred from the curvature of the track in a 1.5 T solenoid magnetic field, while the amount of ionization per unit length, or energy loss, provides information on the particle identity.

3.3.4 Particle identification system

The detector employs a sophisticated Particle Identification (**PID**) system designed to distinguish between different charged particle species, such as pions, kaons, and protons, over a wide momentum range. This capability is essential for accurate reconstruction of decay processes and reducing background contamination.

Two complementary Cherenkov detectors comprise the core of the **PID** system: the **TOP** and the **ARICH** detector [56, 57].

3.3.4.1 Barrel: Time-of-propagation counter

The **TOP** detector surrounds the barrel region of the tracking system and exploits the Cherenkov effect, where charged particles travelling faster than the speed of light in a radiator material emit photons at a characteristic angle. Quartz bars serve as radiators and light guides. Cherenkov photons produced by a traversing particle propagate via total internal reflection to photodetectors. By precisely measuring the arrival time and spatial distribution of these photons, the detector reconstructs the Cherenkov angle and thus infers the particle velocity. When combined with momentum measurements from the tracking system, this allows for identification of particle species in the intermediate momentum range.

3.3.4.2 End-cap: Aerogel Ring-Imaging Cherenkov detector

Located in the forward endcap region, the **ARICH** detector complements the barrel-region **TOP** detector by providing **PID** in the forward acceptance. It employs a dual-layer aerogel radiator with varying refractive indices to produce Cherenkov photons, which are detected by arrays of hybrid avalanche photodiodes. The radius of the resulting Cherenkov rings depends on the particle velocity, enabling separation of charged particles. While the **TOP** detector offers excellent **PID** in the barrel region, the **ARICH** detector covers the forward angles. The dual-layer aerogel structure enhances photon yield and angular resolution, improving **PID** efficiency in the forward direction.

3.3.5 Electromagnetic calorimeter

The **ECL** measures the energy and position of electromagnetic showers [52, 56–58]. It is responsible for particle identification, triggering, and the reconstruction of neutral particles, which decay into photons.

The **ECL** consists of scintillating crystals in the barrel and the two endcap sections, covering 90% of the solid angle in the centre-of-mass frame. When a photon or electron enters the calorimeter, it initiates an electromagnetic shower through cascading processes of bremsstrahlung and pair production. High-energy electrons emit photons via bremsstrahlung when deflected by atomic nuclei, while high-energy photons convert to electron-positron pairs. This cascade multiplies until the particles reach low enough energies to be absorbed through ionization and excitation of the crystal lattice. The

scintillation light is produced in proportion to the total deposited energy. The resulting signals provide precise measurements of both the total deposited energy and the shower position [10].

In addition to energy measurement, the **ECL** contributes to particle identification by helping to distinguish between electrons and hadrons based on shower shape and energy-momentum correlation.

3.3.6 K_L^0 and μ detection

The **KLM** forms the outermost component of the detector [56, 57]. Its primary functions are muon identification, detection of K_L^0 mesons through hadronic interactions, and providing trigger and background rejection capabilities.

The **KLM** consists of an alternating sandwich of iron plates and active detector elements, located outside the superconducting solenoid. In the barrel region, the active layers are composed of resistive plate chambers. To withstand higher background rates, the endcaps and innermost barrel layers use plastic scintillator strips with wavelength-shifting fibres, read out by silicon photomultipliers.

Muons are identified based on their ability to penetrate through the detector and absorber material. As charged particles, muons traverse the entire detector, including the **ECL** and solenoid magnet, and continue into the **KLM**. They pass through the detector layers without initiating hadronic or electromagnetic showers. This distinguishes them from hadrons and electrons, which typically stop or shower earlier.

Neutral kaons of the K_L^0 type, on the other hand, do not leave a signal in the inner tracking systems or calorimeter due to their long lifetimes and lack of charge. Instead, they are detected in the **KLM** via hadronic interactions with the absorber material, which produce visible secondary particles in the active detector layers.

3.3.7 Trigger

The trigger system is responsible for selecting potentially interesting events from the large number of particle collisions, reducing the data rate to a manageable level [52, 59]. Its primary goal is to maintain high efficiency for events of physics interest, while rejecting background and uninformative events in real time.

Belle II employs a two-level trigger system: the hardware-based Level-1 (L1) trigger and the software-based High-Level Trigger (HLT). The L1 trigger operates with low latency and uses information from the **ECL**, **CDC**, **KLM** and **TOP** to identify charged tracks, electromagnetic clusters, and muon-like signatures. It reduces the event rate from the initial collision frequency of about 200 MHz to about 30 kHz.

Events accepted by the L1 trigger are processed by the HLT, which performs more refined reconstruction and selection using full detector information, except for **PXD** information. The HLT leverages real-time software algorithms similar to offline reconstruction. The

HLT reduces the data rate further to approximately 10 kHz, making storage and subsequent analysis feasible. Further, it identifies ROIs for the PXD readouts to reduce the data flux.

The selection logic of the trigger system is governed by predefined trigger menus, which define criteria for accepting events based on reconstructed objects such as high-momentum tracks, electromagnetic clusters, and muon candidates. These menus are optimized for various physics goals, including rare decay searches, flavour tagging, and calibration processes.

3.4 Computing and software

The core software framework used at Belle II is `basf2` (Belle Analysis Software Framework 2), a modular C++ framework designed for reconstruction, simulation, and analysis [60]. It supports both real-time data processing and offline analysis.

Once raw data are recorded by the detector and passed through the data acquisition system (DAQ), calibration constants, dependent on accelerator and detector conditions, are determined. Subsequently, the raw data are processed, and the result is stored in smaller mini Data Summary Table (mDST) files, subsets of which serve as inputs to most physics analyses. This reconstruction step is discussed in Section 3.4.1. For individual analyses, or topically related groups of such, data undergo *skimming*, to include more analysis-specific information [57].

Observed collision data must be compared to well-understood theoretical expectations to enable meaningful interpretation. This is mostly done with simulated MC data, produced to replicate both the underlying physics processes and the detector response as accurately as possible (see Section 3.4.2).

3.4.1 Reconstruction

Reconstruction is the process of converting raw detector signals (or their simulated equivalents) into high-level physics objects such as tracks, clusters, and PID likelihoods. It is conventionally divided into tracking and PID, where each step combines data from multiple subdetectors [52, 57].

Tracking The tracking process combines information from the Vertex Detectors (VXD) (PXD and SVD), and CDC to reconstruct charged particle trajectories. This is crucial for determining the momentum and charge of particles, as well as their decay vertices. Track finding collects hits or hit clusters to form track candidates by applying pattern recognition algorithms.

In the CDC, a global algorithm uses all hits simultaneously and applies a Legendre transformation in conformal space to efficiently find circular trajectories from the interaction point. A local algorithm complements this by identifying displaced or short tracks via a

cellular automaton that builds track segments iteratively. The outputs are merged into a final track list [61].

Track finding in the **VXD** uses a cellular automaton to link space points in the detectors into track candidates. Possible combinations are reduced by applying geometric selection criteria, which are defined through a so-called sector map. The sector map is a data-driven lookup table that encodes geometrically allowed hit combinations based on **MC** information [52]. Track candidates from the **VXD** and **CDC** are combined by extrapolating **CDC** tracks to the **VXD** layers and matching them to **VXD** tracks [56].

Track fitting estimates the kinematic Parameters Of Interest (**POIs**) of track candidates, such as momentum and position at the vertex. This is achieved using a Kalman filter-based algorithm, which iteratively combines predictions from a model of the particle's motion with measurements from the detector. The Kalman filter accounts for uncertainties in both the model and the measurements. For the fit, a specific particle hypothesis must be assumed [52].

Particle identification After tracks have been identified, **PID** is performed by combining information from multiple subdetectors. For each reconstructed track, subsystem-specific likelihoods are calculated for various particle hypotheses using observables such as ionization energy loss (dE/dx) in the **CDC**, time-of-propagation in the **TOP** detector, Cherenkov angle measurements in the **ARICH**, electromagnetic shower profiles in the **ECL**, and hit patterns in the **KLM**. These observables are compared to expected responses derived from calibration and simulation. The subsystem likelihoods are then combined into global **PID** likelihoods for six different particle species: π^\pm , K^\pm , p^\pm , d^\pm , e^\pm , μ^\pm . Differences of log likelihoods are used to construct **PID** values, which are used to classify the particle type of each track. Neutral **PID** is also possible using the **ECL** and **KLM** clusters. Thereby, one can identify photons and K_L^0 mesons [52, 56].

All reconstructed data are stored in the **mDST** format [57], which includes a compact yet comprehensive set of reconstructed objects. These serve as the primary input to most physics analyses.

3.4.2 Simulation

Simulated data, also known as **MC** data, is a crucial part of any **HEP** experiment, replicating the full detector response to known physics processes. This ensures accurate comparison between theory and experiment. These simulations provide controlled datasets to interpret experimental data, evaluate detector performance, and design analysis strategies. Further, they allow estimation of signal efficiencies, background modelling, and validation of reconstruction algorithms. Large sets of generic **MC** samples, which simulate all known **SM** processes, are produced to represent the real data as faithfully as possible, whereas specialized, smaller sets of **MC** samples are generated for individual analyses [57].

For Belle II [52], the simulation chain begins with event generation using physics event generators such as **EvtGen** [62] to model decays of B and D mesons into exclusive final

states. `PYTHIA` [63] is used for inclusive decay final states and for the continuum production of light quark pairs.

The generated events are then passed to a detailed `GEANT4`-based detector simulation [64], which models the interaction of particles with the full Belle II detector geometry and materials, including energy deposition, multiple scattering, and detector response.

After the physics interactions are simulated, the digitization step emulates the response of the Belle II electronics. This converts the simulated signals into the same digital format used by the real detector hardware. The digitized output is then processed using the same reconstruction algorithms as real collision data, resulting in reconstructed tracks, clusters, and subsequently particle candidates.

To accurately reflect real experimental conditions, simulated events are overlaid with beam-induced backgrounds and noise sampled from randomly triggered data events. This ensures that the simulation accounts for realistic detector occupancy and noise effects.

Large `MC` samples are produced for signal and background processes under varying beam and detector configurations to provide sufficient statistical power for precision measurements.

The output of the `MC` simulation consists of `mDST` files, which are analysed in the same way as the real data. The `MC` data additionally contains the true event information. A unified reconstruction chain for both real and simulated data guarantee direct comparability and accurate modelling of detector effects.

Chapter 4

Binned statistical models in particle physics

“Why does one need so much collision data? Wouldn’t one new physics event suffice?” — This question is asked quite frequently, and it reveals a common misconception about particle physics searches. My answer usually starts with the probabilistic nature of quantum mechanics, which makes every measurement inherently uncertain. It continues with the challenges posed by theoretical predictions that come with their own uncertainties, and it ends with a discussion about the limitations in modelling our detectors — instruments that, despite their sophistication, cannot perfectly capture every detail of particle interactions. But the core message is always the same: distinguishing a genuine signal from random fluctuations in a sea of background processes requires not just one or two events, but statistically significant evidence accumulated over millions or billions of collisions. This is the foundation of this and the following chapter: the extraction of physics insights from experimental data relies fundamentally and entirely on statistics.

Modern **HEP** analyses employ sophisticated statistical models capable of handling complex, multi-dimensional data while rigorously accounting for various sources of uncertainty — both statistical fluctuations inherent to the measurement process and systematic effects arising from theoretical and experimental limitations. This chapter introduces the binned likelihood models that form the mathematical backbone of contemporary **HEP** statistical analyses. These models provide the language in which experimental results are expressed, constraints are communicated, and limits are set. Understanding this framework is essential for the reinterpretation methods presented in later chapters, as reinterpretation fundamentally involves translating experimental constraints from one theoretical context into another using the same statistical machinery.

Section 4.1 reviews probability theory, including frequentist and Bayesian interpretations, and introduces likelihood-based inference. **Section 4.3** derives the Poisson distribution and binned likelihood functions that arise naturally in counting experiments. **Section 4.4** presents the **HistFactory** specification — a standardized framework for binned likelihood models in particle physics — detailing its modifier types for parametrizing systematic un-

certainties and its Bayesian extension. [Section 4.4.3](#) introduces `pyhf`, a modern Python implementation of the `HistFactory` formalism, providing portable JSON-based models and efficient inference through automatic differentiation. Finally, [Section 4.4.4](#) presents extensions developed for this work: methods for arbitrary parameter correlations ([Section 4.4.4.1](#)) and custom modifiers ([Section 4.4.4.2](#)) that enable the flexible reinterpretation approaches in later chapters.

4.1 Probability and random variables

The fundamental concept of uncertainty in measurements can have many origins. For example, it could derive from a lack of knowledge of a system, e.g. the exact details of a particle detector. On the other hand, it could also stem from a more fundamental unpredictability such as the nondeterministic property of quantum mechanics. A measurement which depends on such undetermined factors is said to be *random*.

To which degree this randomness manifests itself can be classified with the concept of *probability*. Probability can be defined with three simple axioms. Consider a sample space S , where to each subset $A \subset S$ one assigns a real number $p(A)$ called the probability, satisfying the axioms [65]

- $p(A) \geq 0$ for every subset $A \subset S$
- For two disjoint subsets A, B ($A \cap B = \emptyset$), $p(A \cup B) = p(A) + p(B)$
- The probability assigned to the sample space is unity $p(S) = 1$

A variable that takes a specific value for every element of the sample space S is called a random variable.

The *conditional probability* is defined as

$$p(A|B) = \frac{p(A \cap B)}{p(B)}, \quad (4.1)$$

provided $p(B) \neq 0$.

Two subsets are said to be *independent* if their joint probability factorizes:

$$p(A \cap B) = p(A)p(B). \quad (4.2)$$

This independence condition immediately implies that $p(A|B) = p(A)$ and $p(B|A) = p(B)$, meaning knowledge of one event provides no information about the other.

Using the symmetry property $A \cap B = B \cap A$, the joint probability can be expressed in two equivalent ways:

$$p(A \cap B) = p(A|B)p(B) = p(B \cap A) = p(B|A)p(A). \quad (4.3)$$

Rearranging this fundamental relationship yields *Bayes' theorem*:

$$p(A|B) = \frac{p(B|A)p(A)}{p(B)}. \quad (4.4)$$

In this formulation, $p(A)$ represents the *prior probability*, $p(B|A)$ the *likelihood*, $p(B)$ the *marginal likelihood*, and $p(A|B)$ the *posterior probability* [65].

4.1.1 Interpretations of probability

Given a function satisfying the above axioms, it is still necessary to specify how to interpret the elements of the sample space and how to assign and interpret the probability values. There are two main interpretations: the *frequentist* and *subjective* interpretations [65].

One important point to keep in mind is that physics does not care about the interpretation of probability — conclusions about the physical world from measurements should not depend on it.

4.1.1.1 Frequentist probability

The *frequentist* interpretation assigns probability to the relative number of occurrences of some event A , in the limit of infinite trials [65–67],

$$p(A) = \lim_{N \rightarrow \infty} \frac{N_A}{N} \quad (4.5)$$

where A occurs N_A times in N trials. This is a natural interpretation for particle physics, where processes are counted over large numbers of independent collisions. However, it is not applicable when repeated measurements are infeasible, and prior information can only be incorporated through the data itself, typically via constraints or auxiliary measurements.

4.1.1.2 Subjective probability

In the *subjective*, or *Bayesian*, interpretation, probability is associated with the degree of belief in some hypothesis A [65, 67, 68],

$$p(A) = \text{degree of belief that hypothesis } A \text{ is true.} \quad (4.6)$$

This interpretation can handle situations where the frequentist approach fails, such as assigning probabilities to theories defined by unknown constants.

The subjective interpretation is closely related to Bayes' theorem, formulated as

$$p(\text{theory}|\text{data}) = \frac{p(\text{data}|\text{theory})p(\text{theory})}{p(\text{data})}. \quad (4.7)$$

A key advantage is the method of prior updating: posteriors from one analysis can serve as priors for subsequent data, iteratively refining one's degree of belief.

The main criticism is that results depend on the prior, whose assignment is not always clear. Further, computing $p(\text{data})$ is often intractable, so one typically samples from the posterior: $\text{theory} \sim p(\text{theory}|\text{data})$.

4.1.1.3 The common ground

Given either interpretation, the goal of any statistical analysis is to make inferences about the underlying probability distribution of the data. The mathematics behind the two interpretations is the same, and Bayes' theorem is valid under both.

The physics lies in the model of the observed data $p(\text{data}|\text{theory})$, which is common ground for both interpretations. Frequentist inference is based on the likelihood alone, while Bayesian inference additionally incorporates prior beliefs $p(\text{theory})$. In both cases, obtaining the best possible model of the data is paramount.

4.2 The nature of statistical inference

Consider the measurement of a random variable x , distributed according to some PDF (see Appendix B.1) $p(x|\boldsymbol{\theta})$, depending on a set of unknown parameters $\boldsymbol{\theta} = (\theta_1, \dots, \theta_m)$.

For a statistical analysis, most effort should go into the modelling of the assumed data-generating process

$$x \sim p(x|\boldsymbol{\theta}). \quad (4.8)$$

The model should be chosen such that it is as close as possible to the true data-generating process.

When treating $p(x|\boldsymbol{\theta})$ as a function of $\boldsymbol{\theta}$ for fixed observed data x , it is called the *likelihood* function. If the hypothesized PDF and parameter values are correct, one expects a large value for the likelihood. Conversely, parameter values far away from the true values should yield a low probability for the measurements obtained. It is important to note that the likelihood is not a probability distribution of $\boldsymbol{\theta}$, as it is generally not normalized [65],

$$\int dx p(x|\boldsymbol{\theta}) = 1 \quad \forall \boldsymbol{\theta}, \quad \text{but} \quad \int d\boldsymbol{\theta} p(x|\boldsymbol{\theta}) \neq 1. \quad (4.9)$$

In particle physics the power lies in repeated experiments. Consider a *sample* of N independent measurements of x , $\mathbf{x} = (x_1, \dots, x_N)$. The likelihood function is then given by

$$p(\mathbf{x}|\boldsymbol{\theta}) = \prod_{i=1}^N p(x_i|\boldsymbol{\theta}). \quad (4.10)$$

This equation highlights an important property of likelihood methods: *the power of combination*. It demonstrates the possibility to combine independent measurements simply by multiplying their likelihoods. This implies that likelihoods of completely different samples \mathbf{x} and \mathbf{y} , depending on parameters $\boldsymbol{\theta}_x$ and $\boldsymbol{\theta}_y$, respectively, can be combined,

$$p(\mathbf{x}, \mathbf{y}|\boldsymbol{\theta}_x \cup \boldsymbol{\theta}_y) = p(\mathbf{x}|\boldsymbol{\theta}_x)p(\mathbf{y}|\boldsymbol{\theta}_y), \quad (4.11)$$

to improve constraints on parameters, which both measurements are sensitive to, $\rho = \theta_x \cap \theta_y$. This has very powerful consequences for parameter estimates and their uncertainties.

Since the likelihood function is commonly a product of small numbers, its actual value may become very small and hence hard to deal with computationally. Due to this reason, it is useful to consider twice the negative log-likelihood, $-2 \log p(\mathbf{x}|\boldsymbol{\theta})$, to avoid dealing with very small numbers. The factor of 2 is a convenience, explicable when considering twice the negative log-likelihood for normally distributed data points, dropping all constant terms,

$$-2 \log p(\mathbf{x}|\boldsymbol{\theta}) = \sum_{i=1}^N \frac{(x_i - \mu(\boldsymbol{\theta}))^2}{\sigma_i^2}, \quad (4.12)$$

which is the well known $\chi^2(\boldsymbol{\theta})$ variable, which is distributed according to a χ^2 -distribution. Here σ_i corresponds to the uncertainty of x_i .

4.3 Counting experiments

In particle physics one performs counting experiments. For example, given that one particle X decays to particle Y with a probability P , the probability to observe n $X \rightarrow Y$ events out of N decays of X can be calculated.

If all the decays were *distinguishable* (the order of events matters), the probability is [65],

$$p(n|N, P) = P^n (1 - P)^{N-n}, \quad (4.13)$$

which simply is the product of any event with probability P happening n times out of N distinguishable trials.

For *indistinguishable* events (only the overall number of a particular decay matters), there is an additional factor accounting for the combinatorics. Adding this factor results in the *binomial distribution* [65],

$$p(n|N, P) = \frac{N!}{n!(N-n)!} P^n (1 - P)^{N-n}. \quad (4.14)$$

What if there are not only two possible outcomes? One might want to consider X decaying into a full alphabet of particles, $X \rightarrow A$, or $X \rightarrow B$, \dots , $X \rightarrow Z$ each happening with a probability P_A, P_B, \dots, P_Z . Then the probability to get n_A $X \rightarrow A$ events, n_B $X \rightarrow B$ events, \dots , is given by the *multinomial distribution*,

$$p(n_A, n_B, \dots, n_Z | N, P_A, P_B, \dots, P_Z) = \frac{N!}{n_A! n_B! \dots n_Z!} P_A^{n_A} P_B^{n_B} \dots P_Z^{n_Z}, \quad (4.15)$$

where $N = \sum_{i=A}^Z n_i$ and $\sum_{i=A}^Z P_i = 1$.

In principle, one could perform analyses in particle physics with the binomial distribution of Equation (4.14) or its extension, the multinomial distribution of Equation (4.15). In

general, one wants to compare the *expected number* of $X \rightarrow Y$ events to the actually observed ones. The expected number of events is simply $\nu = NP$, where commonly N is very large (experiments take a lot of data) and P is very small (actually observing $X \rightarrow Y$ is very rare).

In the limit of large N and small P , with some finite $\nu = NP$, the binomial distribution in Equation (4.14) becomes the *Poisson distribution* [65],

$$p(n|\nu) = \frac{\nu^n}{n!} e^{-\nu}. \quad (4.16)$$

4.3.1 Binned likelihoods

Given a total number of N observations of a random variable x distributed according to a PDF $p(x|\boldsymbol{\theta})$, in the case of N becoming very large, it becomes difficult to evaluate the negative log-likelihood for all points in the sample \mathbf{x} . One often simplifies by binning the data in a histogram of B bins. This results in a set of observed data yields $\mathbf{n} = (n_1, \dots, n_B)$ and expected yields $\boldsymbol{\nu} = (\nu_1, \dots, \nu_B)$ in each bin. The expected yields are given by

$$\nu_b(\boldsymbol{\theta}) = N \int_{\text{bin } b} dx p(x|\boldsymbol{\theta}), \quad (4.17)$$

where the integral boundaries correspond to the bin boundaries of the given bin.

Since the individual observations, and hence also the bins are independent, and given that each event must fall into any of the bins, the full PDF is given by a multinomial distribution (Equation (4.15)) with $P_b = \nu_b/N$ [65],

$$p(\mathbf{n}|\boldsymbol{\nu}) = \frac{N!}{n_1! \dots n_B!} \left(\frac{\nu_1}{N}\right)^{n_1} \dots \left(\frac{\nu_B}{N}\right)^{n_B}. \quad (4.18)$$

Empty bins, i.e. $n_i = 0$ for some i , are of no concern, but expectation values $\nu_i = 0$ would result in $p(\mathbf{n}|\boldsymbol{\nu}) = 0$.

So far, the discussion has considered the situation where the total number of events was fixed to N . But the total number of events might be a random variable itself, distributed according to a Poisson distribution,

$$p(N|\nu_{tot}) = \frac{\nu_{tot}^N}{N!} e^{-\nu_{tot}}. \quad (4.19)$$

Here $N = \sum_{b=1}^B n_b$ and $\nu_{tot} = \sum_{b=1}^B \nu_b$, where the expected yields are now a function of the expected total yields,

$$\nu_b(\nu_{tot}, \boldsymbol{\theta}) = \nu_{tot} \int_{\text{bin } b} dx p(x|\boldsymbol{\theta}). \quad (4.20)$$

In this case the full PDF is given by

$$p(\mathbf{n}|\boldsymbol{\nu}) = \frac{\nu_{tot}^N}{N!} e^{-\nu_{tot}} \frac{N!}{n_1! \dots n_B!} \left(\frac{\nu_1}{\nu_{tot}}\right)^{n_1} \dots \left(\frac{\nu_B}{\nu_{tot}}\right)^{n_B}, \quad (4.21)$$

where the probability of an event landing in bin i is now determined through the total expected yield, $P_b = \nu_b/\nu_{tot}$.

Interestingly, the above equation can be simplified to

$$p(\mathbf{n}|\boldsymbol{\nu}(\boldsymbol{\theta})) = \prod_{b=1}^B \text{Pois}(n_b|\nu_b) = \prod_{b=1}^B \frac{\nu_b(\boldsymbol{\theta})^{n_b}}{n_b!} e^{-\nu_b(\boldsymbol{\theta})}. \quad (4.22)$$

Hence, it is found that a multinomial PDF with an overall Poisson constraint on the total number of observations is equivalent to the product of Poisson PDFs for individual histogram bins.

4.4 HistFactory

Data analyses in particle physics have many common features: they are based on counting experiments which compare collider data with simulated MC data, are subject to statistical and systematic fluctuations (as any analysis of data), and many are histogram-based due to the overall large data yields. Simulated data are easily classifiable into categories, such as *signal* (events that one is interested in) and *background* (events that one is less interested in) in simple cases.

To unify data analyses in particle physics to a certain degree, a mathematical specification of *closed-world*¹ statistical models was developed: **HistFactory** [70]. It has become a gold standard in the particle physics community, across experiments, since it is flexible enough to fit the purpose of many analyses, while providing a limited set of building blocks, which maintains simplicity.

The **HistFactory** model applies in cases where,

- one conducts an analysis of binned data,
- one can separate the events into categories of physics processes (*samples*²),
- uncertainties can be modelled as either shape or normalization variations to the histogram.

The hierarchy of a **HistFactory** model consists of C *channels* — each a distinct binned distribution with B_c bins. Each channel consists of S_c *samples* — distinct physics processes. Furthermore, variations due to statistical or systematic effects are specified as parametrized histogram variations, named *modifiers*.

The basis of the **HistFactory** model is the Poisson PDF derived in Equation (4.22) [70, 71],

$$p_{\text{data}}(\mathbf{n}|\boldsymbol{\theta}) = \prod_{c \in C} \prod_{b \in B_c} \text{Pois}(n_{cb}|\nu_{cb}(\boldsymbol{\theta})). \quad (4.23)$$

¹Closed-world statistical models contain a set of given building blocks for model building, whereas open-world models allow for custom components to be added [69].

²A **HistFactory** specific wording, not to be confused with a statistical sample.

The expected yields are given by nominal expected yields for each channel c , sample s and bin b , ν_{csb}^{nom} , subject to parametrized variations, due to given uncertainties in the expectation. Nominal yields are commonly obtained from **MC** simulation and subsequent reconstruction and event selection. These parametrized variations can either be multiplicative, $\kappa(\boldsymbol{\theta})$, or additive, $\Delta(\boldsymbol{\theta})$. The resulting expected yields are given by

$$\nu_{cb}(\boldsymbol{\theta}) = \sum_{s \in S_c} \nu_{csb} = \sum_{s \in S_c} \left(\prod_{\kappa \in \boldsymbol{\kappa}} \kappa_{csb}(\boldsymbol{\theta}) \right) \left(\nu_{csb}^{\text{nom}} + \sum_{\Delta \in \boldsymbol{\Delta}} \Delta_{csb}(\boldsymbol{\theta}) \right). \quad (4.24)$$

The model parameters, controlling the variations, fall into two categories: *unconstrained* and *constrained* parameters,

$$\boldsymbol{\theta} = (\boldsymbol{\eta}, \boldsymbol{\chi}). \quad (4.25)$$

Unconstrained parameters, $\boldsymbol{\eta}$, are free — their variation only changes the likelihood value through the variation of the expected yields. Constrained parameters, $\boldsymbol{\chi}$, are subject to given constraint **PDFs**, given *auxiliary data* \mathbf{a} ,

$$p_{\text{constr.}}(\mathbf{a}|\boldsymbol{\chi}) = \prod_{\chi \in \boldsymbol{\chi}} p(a_\chi|\chi). \quad (4.26)$$

Note that this is a product of individual constraint terms for each parameter $\chi \in \boldsymbol{\chi}$. Hence, each parameter is considered independent by assumption. How one can nevertheless deal with correlated parameters is discussed in **Section 4.4.4**.

One can also separate the model parameters into two alternative categories: *POIs* and *nuisance parameters*,

$$\boldsymbol{\theta} = (\boldsymbol{\mu}, \boldsymbol{\psi}). \quad (4.27)$$

Parameters of interest, $\boldsymbol{\mu}$, as the name suggests, are parameters from which valuable knowledge is extracted. Nuisance parameters, $\boldsymbol{\psi}$, are commonly not of interest, but are needed to parametrize the model — these are usually used to describe sources of uncertainties. Although not strictly true for any model, in the context of this work, **POIs** are always unconstrained and nuisance parameters are always constrained parameters.

Within the frequentist interpretation of statistics, constraint terms can only be introduced through auxiliary measurements. This can be for example a control region measurement, constraining the normalization of one or multiple samples in the model.

The full **HistFactory** statistical model is given as the product of **Equations (4.23)** and **(4.26)**,

$$p(\mathbf{n}, \mathbf{a}|\boldsymbol{\eta}, \boldsymbol{\chi}) = p_{\text{data}}(\mathbf{n}|\boldsymbol{\eta}, \boldsymbol{\chi}) p_{\text{constr.}}(\mathbf{a}|\boldsymbol{\chi}) = \prod_{c \in C} \prod_{b \in B_c} \text{Pois}(n_{cb}|\nu_{cb}(\boldsymbol{\eta}, \boldsymbol{\chi})) \prod_{\chi \in \boldsymbol{\chi}} p(a_\chi|\chi). \quad (4.28)$$

4.4.1 Modifier types

In the `HistFactory` statistical model, parametrized variations to binned distributions are controlled through *modifiers*. Since the `HistFactory` model is closed-world, only a fixed set of modifications are available within the framework. They can be classified into multiplicative, $\kappa(\theta)$, or additive, $\Delta(\theta)$ modifiers, as in Equation (4.24). Each modifier is parametrized by one or a set of parameters, with a corresponding constraint term. Furthermore, each modifier requires certain input specifications, which are used to compute the auxiliary data, depending on the type. Modifier input specifications commonly include variations of the histogram. Modifier parameters then interpolate between the nominal histogram, the variations, and beyond.

The available modifier types in the `HistFactory` specification and their respective constraint terms are listed here [70, 72].

Uncorrelated shape — `shapesys` This is a bin-wise, multiplicative modifier for a fixed channel \bar{c} and sample \bar{s} , $\kappa_{\bar{c}\bar{s}b}$, introducing one parameter per bin. The required input is the absolute uncertainty on the yield of each bin $\sigma_{\bar{c}\bar{s}b}$, for the given channel and sample. The amount of auxiliary data $a_{\bar{c}\bar{s}b}$ per bin is determined by asking “which effective number of entries would give that relative uncertainty”. That is given by

$$\frac{\sqrt{a_{\bar{c}\bar{s}b}}}{a_{\bar{c}\bar{s}b}} = \frac{\sigma_{\bar{c}\bar{s}b}}{\nu_{\bar{c}\bar{s}b}^{\text{nom}}} \rightarrow a_{\bar{c}\bar{s}b} = \left(\frac{\nu_{\bar{c}\bar{s}b}^{\text{nom}}}{\sigma_{\bar{c}\bar{s}b}} \right)^2 \quad (4.29)$$

The corresponding constraint term is

$$p(\mathbf{a}|\kappa_{\bar{c}\bar{s}}) = \prod_{b \in B_{\bar{c}}} \text{Pois}(a_{\bar{c}\bar{s}b} | a_{\bar{c}\bar{s}b} \kappa_{\bar{c}\bar{s}b}). \quad (4.30)$$

Statistical uncertainty — `staterror` Expected bin yields are often derived from `MC` datasets. Consequently, they necessarily carry an uncertainty due to the finite sample size of the datasets. In principle, one could add an *uncorrelated shape* modifier for each sample, to account for this uncertainty, but this would yield an excessively large number of parameters with limited utility. To circumvent this, one can quadratically combine the overall statistical uncertainty per bin, over the collection of all samples [73]. This is exactly what the `staterror` modifier does, adding a bin-wise, multiplicative modifier, for a given channel \bar{c} , $\kappa_{\bar{c}b} = \sum_{s \in S_c} \kappa_{\bar{c}\bar{s}b}$. The required input is the absolute uncertainty on the yield of each bin $\sigma_{\bar{c}sb}$, for each of the samples in a given channel to be included.

The constraint term is a product of normal `PDF`s over bins,

$$p(\mathbf{a}|\kappa_{\bar{c}}) = \prod_{b \in B_{\bar{c}}} \mathcal{N}(a_{\bar{c}b} = 1 | \kappa_{\bar{c}b}, \sigma_{\bar{c}b}^{\text{rel}}), \quad (4.31)$$

where the auxiliary data are $a_{\bar{c}b} = 1$, and

$$\sigma_{\bar{c}b}^{\text{rel}} = \sqrt{\sum_{s \in S_c} \sigma_{\bar{c}sb}^2} / \sum_{s \in S_c} \nu_{\bar{c}sb}^{\text{nom}}. \quad (4.32)$$

Correlated shape uncertainty — `histosys` Systematic effects mostly affect the whole histogram in a correlated manner, affecting both shape and normalization of the histogram. By varying certain parameters of underlying assumptions on either the theoretical or experimental model by one standard deviation (assuming Gaussianity), one can compute upward, Δ^{up} , and downward, Δ^{dn} , variations of the nominal bin yields. These variations serve as the input data for the `histosys` modifier. This is an additive modifier, where a single interpolation parameter α controls the modification to the expected bin yields,

$$\Delta(\alpha) = \begin{cases} \alpha|\Delta^{\text{up}} - \nu^{\text{nom}}| & \alpha \geq 1 \\ \sum_{n=0}^6 k_n \alpha^n & |\alpha| < 1 \\ \alpha|\Delta^{\text{dn}} - \nu^{\text{nom}}| & \alpha \leq -1 \end{cases}, \quad (4.33)$$

where the parameters k_n are fixed by the boundary conditions at $\alpha = 0, \pm 1$, as well as the continuity conditions (matching of first and second derivative) at $\alpha = \pm 1$. The corresponding constraint term for α is a normal distribution with unit width and auxiliary data $a = 0$,

$$p(a|\alpha) = \mathcal{N}(a = 0|\alpha, \sigma = 1). \quad (4.34)$$

Normalization uncertainty — `normsys` As evident from the name, with this modifier one can introduce an uncertainty on the overall normalization of a particular sample in a given channel \bar{c} . This is a multiplicative modifier $\kappa_{\bar{c}s}(\alpha)$, parametrized by a single parameter α , which interpolates between the upward and downward multiplicative factor at one standard deviation, $\kappa_{\bar{c}s}(+1) = \kappa^{\text{up}}$ and $\kappa_{\bar{c}s}(-1) = \kappa^{\text{dn}}$, given as input data. The nominal value is $\kappa(0) = 1$. The corresponding constraint term for α is a normal distribution with unit width and auxiliary data $a = 0$,

$$p(a|\alpha) = \mathcal{N}(a = 0|\alpha, \sigma = 1). \quad (4.35)$$

Luminosity uncertainty — `lumi` Histograms created from simulated `MC` samples are commonly scaled to the overall luminosity of the data itself, for a truthful comparison to the overall theoretically predicted rate. This is a global multiplicative modifier, with a luminosity scale κ_L and uncertainty σ_L as inputs. Global here means that it affects all samples to which the modifier is added in the same way. The constraint term is a normal distribution,

$$p(\kappa_L|\kappa) = \mathcal{N}(\kappa_L|\kappa, \sigma_L). \quad (4.36)$$

Unconstrained normalization — `normfactor` This modifier scales the histogram by a free parameter μ . This is commonly used to measure the signal strength of a given process, as the `POI`. There is no input data needed. Furthermore, there is no constraint term, such that a deviation from the nominal value $\mu = 1$ only affects the data likelihood.

Data-driven shape — `shapefactor` This is a free, bin-wise multiplicative modifier for a fixed channel \bar{c} and sample \bar{s} , $\kappa_{\bar{c}\bar{s}b}$. No additional input data are needed, and no constraint term is added to the likelihood. This is for example used to model the shape

of a background process which is not well known, but can be constrained by data in a control region. The parameters $\kappa_{\bar{c}sb}$ are then fitted simultaneously in the signal and control region.

4.4.2 Bayesian HistFactory

A well-defined statistical model is the basis for any inference problem. In [Section 4.1.1.2](#) it was discussed that Bayesian inference is based on the posterior. Starting from HistFactory likelihood $p(\mathbf{n}, \mathbf{a}|\boldsymbol{\eta}, \boldsymbol{\chi})$ in [Equation \(4.28\)](#), one can formulate a posterior model $p(\boldsymbol{\eta}, \boldsymbol{\chi}|\mathbf{n}, \mathbf{a})$, given that corresponding priors are defined for all parameters $\boldsymbol{\eta}$ and $\boldsymbol{\chi}$. Given that parameters in the HistFactory model are independent, it is possible to separate the prior distributions for constrained and unconstrained parameters, $p(\boldsymbol{\eta})$ and $p(\boldsymbol{\chi})$, respectively. The posterior model is then given by

$$p(\boldsymbol{\eta}, \boldsymbol{\chi}|\mathbf{n}, \mathbf{a}) \propto p(\mathbf{n}, \mathbf{a}|\boldsymbol{\eta}, \boldsymbol{\chi}) p(\boldsymbol{\eta}) p(\boldsymbol{\chi}), \quad (4.37)$$

where the normalization is just the total integral over the right-hand side. Separating the HistFactory likelihood into its *data* and *constraint* parts as in [Equation \(4.28\)](#), one obtains

$$p(\boldsymbol{\eta}, \boldsymbol{\chi}|\mathbf{n}, \mathbf{a}) \propto p_{\text{data}}(\mathbf{n}|\boldsymbol{\eta}, \boldsymbol{\chi}) p(\boldsymbol{\chi}|\mathbf{a}) p(\boldsymbol{\eta}). \quad (4.38)$$

where $p(\boldsymbol{\chi}|\mathbf{a}) \propto p(\mathbf{a}|\boldsymbol{\chi})p(\boldsymbol{\chi})$ has been defined. This last step is a classic example of *Bayesian updating* [68], where the posterior of a previous measurement is used as a prior for a subsequent one. In this case, the constraint likelihood of [Equation \(4.28\)](#), is used to update the PDF of the constraint parameters given in $p(\boldsymbol{\chi})$ ³.

Since the constraint terms are a set of default PDFs, as discussed in [Section 4.4.1](#), the priors $p(\boldsymbol{\chi})$ can also be set to default PDFs [74]. It is convenient to choose *conjugate priors*, which are priors that yield a posterior of the same type as the prior itself. This is especially useful for the HistFactory model, since it allows for an algebraic expression of the posterior, which can be computed efficiently. The constraint likelihoods are either normal or Poisson PDFs. Conjugate priors are discussed in [Appendix B.2](#).

4.4.3 HistFactory in Python: pyhf

The classic implementation of the HistFactory model is within the ROOT framework [75]. Implemented through RooFit and RooStats, the ROOT-based HistFactory provides a flexible configuration via XML files and extensive C++/Python interoperability. However, this approach often requires a full ROOT environment, entails non-trivial software dependencies, and is less convenient to modern analysis workflows.

In contrast to the original ROOT-based implementation, pyhf (pure Python HistFactory) is a modern, ROOT-independent reimplement of the HistFactory statistical model [71, 72]. Written entirely in Python, pyhf is designed to support binned likelihood fits in a lightweight and modular environment. It enables analyses to be developed,

³This is sometimes referred to as an *ur-prior*.

shared, and executed without the need for the ROOT framework, facilitating integration into contemporary scientific software ecosystems.

`pyhf` offers two minimizers to choose from — `scipy`-based minimizers [76] for fast and robust minimization, without uncertainty quantification, and `minuit` [77] for a more thorough analysis and detailed results. A key feature of `pyhf` is its support for gradient-based optimisation and automatic differentiation through backends such as `JAX` [78]. These capabilities allow efficient computation of likelihoods and parameter gradients, which is particularly advantageous for large-scale parameter inference (Section 5.1) and enables compatibility with modern machine learning tools and hardware acceleration.

Model definitions in `pyhf` are specified using a structured JSON schema, making them portable, human-readable, and version-controllable. This declarative approach promotes reproducibility and transparency, and aligns with Findable, Accessible, Interoperable, Reusable (FAIR) principles in data preservation and open science. JSON-based models can be easily shared and reused across platforms.

`pyhf` provides a streamlined and efficient implementation of frequentist hypothesis testing using the profile likelihood ratio as the core test statistic (Section 5.2.1). `pyhf` leverages asymptotic formulae derived in Reference [79] to approximate the sampling distributions of various test statistics. This allows for fast calculation of P -values and discovery significances without resorting to computationally expensive pseudo-experiments (Section 5.2.1). This approach provides scalable inference, particularly useful for setting confidence intervals and exclusion limits (Section 5.4).

As an example, a simple `pyhf` model with one bin is shown below, with a single channel `singlechannel`, containing two samples: `signal` and `background`.

```
{
  "channels": [
    {
      "name": "singlechannel",
      "samples": [
        {
          "name": "signal",
          "data": [10],
          "modifiers": [
            { "name": "mu", "type": "normfactor", "data": null }
          ]
        },
        {
          "name": "background",
          "data": [74],
          "modifiers": [
            {
              "name": "bkg_shape_unc",
              "type": "histosys",
              "data": {
```

```

        "lo_data": [70.0],
        "hi_data": [80.0]
    }
}
]
}
]
}
]
}

```

Nominal bin values are provided for both samples. The `signal` includes a normalization factor modifier `mu` representing the **POI**. The `background` includes a shape systematic uncertainty modelled with a `histosys` modifier, defined by low and high variation templates (Section 4.4.1). More details on the syntax of adding modifiers, can be found in the [pyhf documentation](#) [72].

As `pyhf` is the implementation of the `HistFactory` model, there is `bayesian pyhf`, which is the implementation of the Bayesian `HistFactory` model [74]. It takes the `pyhf` likelihood and translates it into a joint posterior distribution as described in Section 4.4.2. The posterior model is implemented in the PyMC [80] framework, which implements a suite of Markov Chain Monte Carlo (**MCMC**) sampling algorithms⁴ to finally obtain posterior samples.

4.4.4 Extended features for `HistFactory` and `pyhf`

As with any model or software, `HistFactory` and `pyhf` also come with limitations and drawbacks. In this section I want to discuss two of these limitations and provide solutions to them.

4.4.4.1 Arbitrarily correlated uncertainties

In `pyhf`, model parameters — particularly nuisance parameters associated with systematic uncertainties — are, by default, treated as statistically uncorrelated across different measurements or channels. This assumption simplifies the construction of the likelihood function by modelling each nuisance parameter with its own independent constraint term. However, in realistic analyses, some systematic effects are shared across multiple samples or channels and must be correlated to avoid under- or overestimating the total uncertainty.

To introduce *complete correlations* in `pyhf`, the same modifier name is used across all affected samples and channels within the JSON-based model specification. When `pyhf` parses the model, it identifies modifiers with matching names and treats them as a single global nuisance parameter, applying their effect consistently wherever referenced. A prerequisite for this is, that the parameters share the same constraint term. *Complete*

⁴A good overview of sampling algorithms can be found in Reference [68].

anti-correlation can be achieved by swapping the up- and down-variations in the modifier definition of one of the modifiers and then correlating them by the same naming scheme.

While `pyhf` supports full correlations between nuisance parameters through shared modifier names, it currently lacks support for defining arbitrary correlations between parameters, such as specifying non-diagonal correlation matrices. This limitation means users cannot directly encode partial correlations or more complex dependence structures among parameters within the model. As a result, all parameter correlations must in principle be either full (via shared modifiers) or absent, which may restrict modelling flexibility in analyses requiring nuanced treatment of correlated uncertainties.

To solve this issue, I have developed `pyhfcorr` [81] — a tool to arbitrarily correlate parameters in a `pyhf` JSON format. It is based on eigendecomposing the covariance matrix of the parameters, resulting in new independent parameters as superpositions of the original ones [34].

Eigendecomposition is a useful method for decorrelating a set of parameters by a unitary transformation. One starts with a covariance matrix C , which is symmetric. Hence, one can always decompose it as

$$C = USU^T \quad (4.39)$$

where $UU^T = \mathbf{1}$. The columns of the transformation matrix U are the eigenvectors of C . The eigenvalues, $s_i = S_{ii}$, are the variances in the rotated space. The standard deviations are $\sigma_i = \sqrt{s_i}$.

If one wants to incorporate the variances, one can define a new transformation matrix $Z = U\sqrt{S}$. Z is column-wise composed of the eigenvectors of C , each of which is scaled by the corresponding standard deviation.

The `pyhf` modifier parameters, \mathbf{p} , describe the contribution of each of these scaled eigenvectors. Hence, to rotate from these parameters to the standard deviation vector for the correlated parameters, $\boldsymbol{\alpha}$, one can apply

$$\boldsymbol{\sigma}_\alpha = Z\mathbf{p}. \quad (4.40)$$

In that way, the modifier parameters are all interpretable in the same way as the usual `pyhf` modifier parameters, i.e. $p_i = \pm 1$ corresponds to a shift of $\pm\sigma_i$ along the i^{th} eigenvector direction.

This approach allows the specification of a correlation matrix for a set of modifier parameters in the `pyhf` model, which are rotated into a new basis of uncorrelated parameters. Based on this, a new `pyhf` model in terms of the rotated parameters can be defined. The number of parameters is always conserved with this method. Furthermore, the output of the `pyhfcorr` tool is a valid `pyhf` model, such that the full inference suite of `pyhf` can still be applied without drawbacks.

With `pyhfcorr` one can extend an existing `pyhf` JSON specification by a `correlations` entry. There one can provide the parameter names and correlation matrix of the to-be

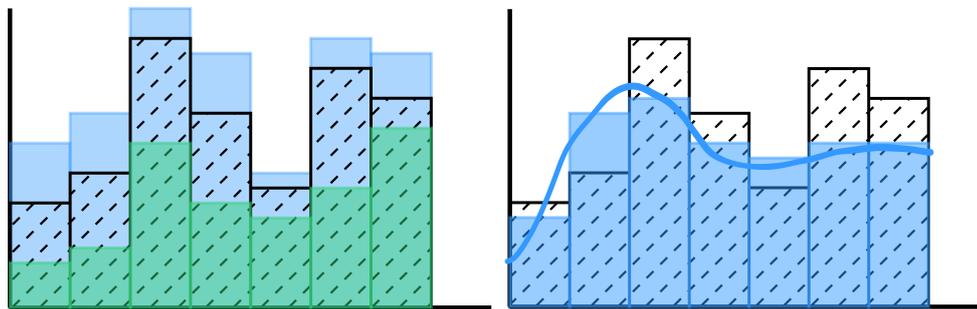


Figure 4.1: This figure illustrates two types of modifiers in `pyhf`. In both panels, the nominal expected yields are represented by the black dashed histogram. *Left*: An additive modifier that alters the expected yields by interpolating between the nominal yields and their variations, as is typical for the default `pyhf` modifiers. These variations are depicted by the blue and green histograms. *Right*: A custom modifier that scales the expected yields in a specific channel using a parametrized function. The function is shown as blue line. The resulting variations, shown by the blue histogram, follow the prescribed custom function.

correlated parameters. As an example, to correlate two parameters `a` and `b`, present in the `pyhf` model, with a correlation coefficient of 0.5, one simply adds:

```
spec = {
    "channels" : ...,
    "correlations": [
        {
            "name": "decorrelate_a_b",
            "vars": ["a", "b"],
            "corr": [[1.0, 0.5], [0.5, 1.]],
        }
    ]
}

new_spec = pyhfcorr.decorrelate.decorrelate(spec)
new_model = pyhf.Model(new_spec)
```

This results in a new `pyhf` model with parameters `a` and `b` replaced by parameters `decorrelate_a_b[0]` and `decorrelate_a_b[1]`, corresponding to the parameters in the decorrelated space.

4.4.4.2 Custom modifiers

`pyhf` is a closed-world statistical model, with a fixed set of modifiers. In some cases, however, it may be desirable to introduce *custom modifiers* — that is, parametrized functions not included in the default set, which directly control the expected yields in a given channel. This is illustrated in [Figure 4.1](#).

There are multiple approaches to map the output of a function onto variations within a histogram. Here, I implement one variant of a custom modifier in `pyhf`, where a parametrized function returns bin-wise weights for a given sample. This function can be defined either directly in Python or as a string expression using the `numexpr` package [82]. The custom modifier operates multiplicatively, ensuring compatibility with other modifiers acting on the same sample. Formally, the custom function can be expressed as

$$\kappa_{\bar{c}s b}(\boldsymbol{\theta}) = \nu_{\bar{c}s b}(\boldsymbol{\theta}) / \nu_{\bar{c}s b}^{\text{nom}}, \quad (4.41)$$

where $\nu_{\bar{c}s b}(\boldsymbol{\theta})$ denotes the new bin yields computed as a function of the parameters $\boldsymbol{\theta}$. The strength of this approach lies in the fact that $\nu_{\bar{c}s b}(\boldsymbol{\theta})$ is an *arbitrary* function, enabling users to construct an “open world” `HistFactory` model [69].⁵ This flexibility is particularly valuable, as it extends beyond the standard `HistFactory` statistical model implemented in `pyhf`, allowing users to incorporate any necessary custom modifications. A limitation of this approach is that, to distribute the `pyhf` model, both the JSON specification and the implementation of the custom function are required — unless the function is defined using `numexpr`.

An example of how the custom modifier can be added to a `pyhf` model is shown below. The modifier is implemented as part of the `redist` package [83], which I will discuss further in [Chapter 7](#).

```
import pyhf
import numpy as np
from scipy import stats
from redist import custom_modifier

def custom_func(pars):
    x = np.linspace(0, 5, 50)
    return lambda: stats.norm.pdf(x, loc=pars['mu'], scale=pars['sig'])

spec = {
    "channels": [{
        "name": "singlechannel",
        "samples": [{
            "name": "signal",
            "data": [1.]*50,
            "modifiers": [{
                "name": "custom_modifier",
                "type": "custom",
                "data": {"expr": 'gauss_func'}
            }]
        }]
    }]
}
```

⁵All models remain binned.

```

new_params = {
    'mu': {'inits': (1.,), 'bounds': ((-5, 5),), 'paramset_type':
    ↪ 'unconstrained'},
    'sig': {'inits': (1.,), 'bounds': ((-5, 5),), 'paramset_type':
    ↪ 'unconstrained'}
}

modifier_set = custom_modifier.add(
    'custom', ['mu', 'sig'], new_params, namespace={'gauss_func':
    ↪ custom_func}
)

model = pyhf.Model(spec, validate=False, modifier_set=modifier_set)

```

The custom function `custom_func` returns a Gaussian PDF, which is then used to scale the expected yields in the channel `singlechannel`. The custom modifier is added to the `pyhf` JSON specification as a modifier of type “custom”, an arbitrary identifier chosen by the user. This identifier is also provided to `custom_modifier.add`, which registers the custom modifier within the `pyhf` model. The custom function accepts a dictionary of parameters, defined in the `new_params` dictionary and passed to `custom_modifier.add`. This dictionary allows specification of initial values, parameter bounds, constraint types, and can even include correlated parameters. Correlations among parameters are internally handled via decorrelation techniques, as discussed in [Section 4.4.4.1](#). The custom function is referenced within the namespace as `gauss_func`, which is also supplied to `custom_modifier.add`.

The `pyhf` model can then be constructed with an extended modifier set (including the custom modifier) and the JSON specification. The parameters `mu` and `sig` are now model parameters. The custom function is evaluated during model fitting, where the parameters `mu` and `sig` are optimized to fit the data. The expected yields of the above example are shown in [Figure 4.2](#). The expected yields are computed for three different parameter values of `mu` and `sig`, demonstrating the flexibility of the custom modifier.

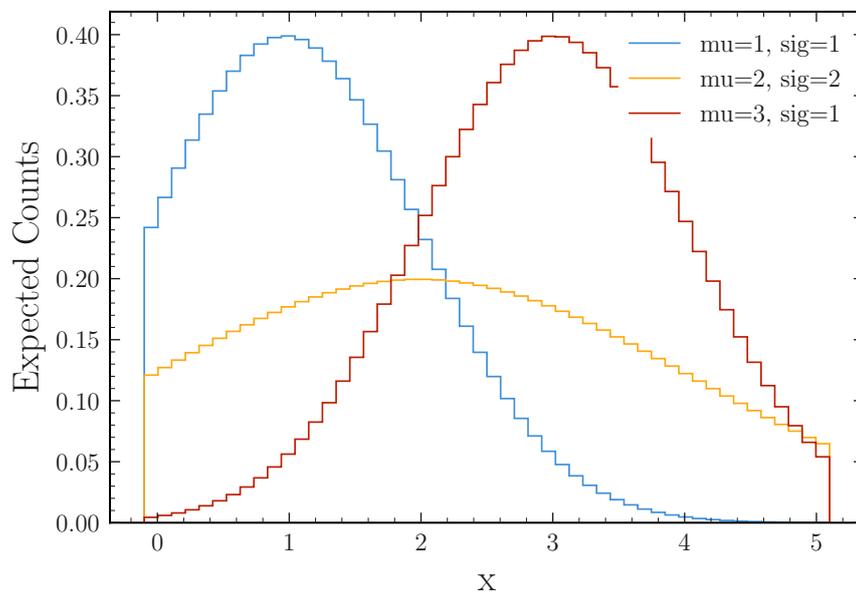


Figure 4.2: Expected yields of the custom modifier example for three different parameter values of μ and σ .

Chapter 5

Statistical inference in particle physics

Building upon the foundational concepts of probability and binned likelihood models established in the previous chapter, this chapter addresses their practical application to extracting physics knowledge from experimental data. Statistical inference provides the mathematical framework for drawing scientifically rigorous conclusions about theoretical parameters from observations while properly accounting for both statistical and systematic uncertainties.

Modern particle physics experiments present interesting challenges: high-dimensional parameter spaces, intricate correlations between observables, and the demanding task of distinguishing genuine signals from background processes. This chapter develops a comprehensive statistical inference framework tailored to address these complexities, with particular emphasis on techniques essential for **BSM** searches and precision measurements.

Section 5.1 covers parameter estimation from both frequentist (maximum likelihood) and Bayesian (maximum a posteriori) perspectives. I discuss fundamental properties of estimators including bias, consistency, and the Cramér-Rao bound, demonstrating their equivalence under uniform priors. **Section 5.2** addresses point-wise comparison through hypothesis testing and global model comparison using Bayes factors, including the Neyman-Pearson lemma, nested hypotheses, and weighted P -values. **Section 5.3** formulates the goodness-of-fit testing framework. Finally, **Section 5.4** covers interval estimation via Bayesian credible intervals and frequentist confidence intervals using Neyman construction, with emphasis on exclusion limits central to **BSM** searches.

5.1 Parameter estimation

Any empirical scientific endeavour involves extracting information from data. One common objective is to infer the numerical values of parameters believed to govern the data-generating process — such as the mass of a particle, specific coupling constants, or branch-

ing ratios. This section discusses some of the basic ideas that flow into the procedure of parameter estimation [65].

The goal is to learn about the underlying PDF from a sample $\mathbf{x} = (x_1, \dots, x_N)$ by constructing a likelihood model $p(\mathbf{x}|\boldsymbol{\theta})$ depending on parameters $\boldsymbol{\theta} = (\theta_1, \dots, \theta_M)$. The procedure of determining these parameters is known as *parameter estimation*, yielding *estimators* $\hat{\boldsymbol{\theta}}(\mathbf{x})$. Since estimators are functions of random data, they are themselves random variables with a sampling distribution $p(\hat{\boldsymbol{\theta}}|\boldsymbol{\theta}_0)$ depending on the true parameter values $\boldsymbol{\theta}_0$.

Key properties of estimators include [65]:

- **Bias:** $\mathbf{b} = \mathbf{E}[\hat{\boldsymbol{\theta}}] - \boldsymbol{\theta}_0$. An *unbiased* estimator has $\mathbf{b} = 0$.
- **Consistency:** An estimator is *consistent* if $\lim_{N \rightarrow \infty} \mathbf{E}[\hat{\boldsymbol{\theta}}] = \boldsymbol{\theta}_0$.
- **Variance:** The estimator variance $\mathbf{V}[\hat{\boldsymbol{\theta}}]$ quantifies the spread of estimates across repeated experiments.

The *Cramér-Rao bound* [66, 67] establishes a fundamental limit: the variance of any unbiased estimator satisfies $\mathbf{V}[\hat{\boldsymbol{\theta}}] \geq \mathbf{I}(\boldsymbol{\theta})^{-1}$, where $\mathbf{I}(\boldsymbol{\theta})$ is the Fisher information [67]. An estimator achieving this bound is optimal. A proof of the Cramér-Rao bound can be found in [Appendix B.3](#). Fisher information is described in more detail in [Appendix B.4](#).

5.1.1 Method of maximum likelihood

An intuitive way to find the best estimator is to choose the one that maximizes the probability of observing the obtained data

$$\hat{\boldsymbol{\theta}}(\mathbf{x}) = \operatorname{argmax}_{\boldsymbol{\theta}} p(\mathbf{x}|\boldsymbol{\theta}). \quad (5.1)$$

$\hat{\boldsymbol{\theta}}(\mathbf{x})$ is the *Maximum Likelihood Estimator (MLE)* of $\boldsymbol{\theta}$.

The MLE corresponds to the point in parameter space where the likelihood function is maximized. This can be found by solving the condition

$$\nabla p(\mathbf{x}|\boldsymbol{\theta})|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} = 0. \quad (5.2)$$

The roots of these equations are the parameter values that give the maximal probability for the observed data.

5.1.1.1 Asymptotic properties of maximum likelihood estimators

The properties of any estimator can be broadly classified into two categories: *asymptotic* properties, which are valid in the limit of large sample sizes ($N \rightarrow \infty$), and *finite-sample* properties, which describe behaviour for a fixed and typically small number of observations [67]. This section focuses on the asymptotic properties of the MLE, which possesses several attractive features.

In the asymptotic limit, the **MLE** has the following properties [65, 66]: it is *consistent* and *asymptotically normal*, which means that in the large sample limit $\sqrt{N}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)$ converges to a normal distribution with minimum variance,

$$\sqrt{N}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \rightarrow^d \mathcal{N}(\mathbf{0}, \mathbf{I}(\boldsymbol{\theta}_0)^{-1}) . \quad (5.3)$$

These properties are proven in [Appendix B.5](#) and [Appendix B.6](#).

5.1.2 Maximum a posteriori estimate

In the Bayesian interpretation of probability, parameters are treated as fundamentally random variables, rather than fixed but unknown quantities (frequentist interpretation). Accordingly, there is no notion of a single *true parameter* value. All available knowledge about the parameters, after observing the data, is encoded in the *posterior distribution* — a multidimensional **PDF** over the parameter space that may also capture correlations among parameters.

Nonetheless, often a particular point in a parameter space is desired. The most common choice in this case is the *posterior mode* or *Maximum A Posteriori estimator (MAP)*,

$$\hat{\boldsymbol{\theta}}_{\text{MAP}}(\mathbf{x}) = \operatorname{argmax}_{\boldsymbol{\theta}} p(\boldsymbol{\theta}|\mathbf{x}) . \quad (5.4)$$

This estimate corresponds to the point in parameter space where the posterior distribution is maximized. Hence, the **MAP** estimator satisfies

$$\nabla_{\boldsymbol{\theta}} p(\boldsymbol{\theta}|\mathbf{x})|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_{\text{MAP}}} = 0 . \quad (5.5)$$

5.1.3 Equivalence of maximum likelihood and maximum a posteriori estimators

The choice between frequentist and Bayesian interpretations of probability should not change the underlying physics of the data-generating process. It is therefore of interest to identify the conditions under which the **MLE** and the **MAP** yield identical results.

Starting from the definition of the **MAP** estimator in [Equation \(5.4\)](#), and substituting the expression for the posterior via Bayes' theorem, the result is

$$\hat{\boldsymbol{\theta}}_{\text{MAP}}(\mathbf{x}) = \operatorname{argmax}_{\boldsymbol{\theta}} p(\boldsymbol{\theta}|\mathbf{x}) = \operatorname{argmax}_{\boldsymbol{\theta}} \left(\frac{p(\mathbf{x}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\mathbf{x})} \right) . \quad (5.6)$$

Given that the argmax function is invariant under scaling by a constant factor, the marginal likelihood $p(\mathbf{x})$ can be dropped,

$$\hat{\boldsymbol{\theta}}_{\text{MAP}}(\mathbf{x}) = \operatorname{argmax}_{\boldsymbol{\theta}} (p(\mathbf{x}|\boldsymbol{\theta})p(\boldsymbol{\theta})) . \quad (5.7)$$

If the prior $p(\boldsymbol{\theta})$ is constant (i.e. uniform over the parameter space), it does not influence the location of the maximum, and the **MAP** estimator reduces to the **MLE**

$$\hat{\boldsymbol{\theta}}_{\text{MAP}}(\mathbf{x}) = \operatorname{argmax}_{\boldsymbol{\theta}} p(\boldsymbol{\theta}|\mathbf{x}) = \operatorname{argmax}_{\boldsymbol{\theta}} p(\mathbf{x}|\boldsymbol{\theta}) = \hat{\boldsymbol{\theta}}(\mathbf{x}) . \quad (5.8)$$

Therefore, in the special case of a uniform prior, the **MAP** and **MLE** estimators are equivalent.

5.2 Model comparison

Scientific progress is based on iteratively improving theoretical models, given observed data. But given a set of different theories that claim to describe the data-generating process, which one should be chosen, and how? How can the performance of one model over another be quantified?

Answering these questions will be the goal of this section. A distinction is made between two complementary approaches to model comparison: the point-wise and the global comparison. The point-wise compares models at their respective best-fit point (Section 5.2.1), whereas the global comparison evaluates the average performance over the full parameter space (Section 5.2.2).

5.2.1 Point-wise model comparison: hypothesis testing

The goal of most analyses is to test how well a predicted model or *hypothesis* describes the data. A hypothesis H would, for example, be the PDF of a random variable x , $p(x|H)$.

Commonly, a benchmark model — the null hypothesis, H_0 — is chosen, which predicts a PDF $p(x|H_0)$. One then wants to compare the performance to alternative hypotheses H_1, H_2, \dots which predict different PDFs $p(x|H_1), p(x|H_2), \dots$

For example, in particle physics, the null hypothesis could be the SM expectation of some process, excluding any contribution from a given new physics process. The alternative hypotheses would include the contribution of the given new physics process that one is targeting to observe.

To measure the agreement between a given prediction and the observed data, it would be helpful to have a function of the data that reflects this. This function is called the *test statistic* $t(\mathbf{x})$. Since the test statistic is a function of a random variable, it itself is a random variable. Hence, each hypothesis H_i predicts a PDF for the test statistic, $p(t|H_i)$. Commonly, smaller values of the test statistic correspond to better agreement between the prediction and the data.

The test statistic can be used to define a *cut* or *decision boundary* t_{cut} , which separates the *rejection region*, $t > t_{cut}$, from the *acceptance region*, $t < t_{cut}$. If the observed value of the test statistic falls into the rejection region, one rejects the null hypothesis. If, on the other hand, the measured test statistic falls into the acceptance region, one rejects the alternative hypothesis. For an illustration see Figure 5.1. The decision boundary is usually chosen such that the probability for t to be observed there, under the assumption of the null hypothesis, is some chosen value α ,

$$\alpha = \int_{t_{cut}}^{\infty} dt p(t|H_0) = 1 - F(t_{cut}|H_0), \quad (5.9)$$

which is also known as the *size* of the test. Here, $F(t_{cut}|H_0)$ is the Cumulative Distribution Function (CDF) (see Appendix B.1.1) up to the cut. Typical values for α are $\alpha = 0.05$ or $\alpha = 0.1$.

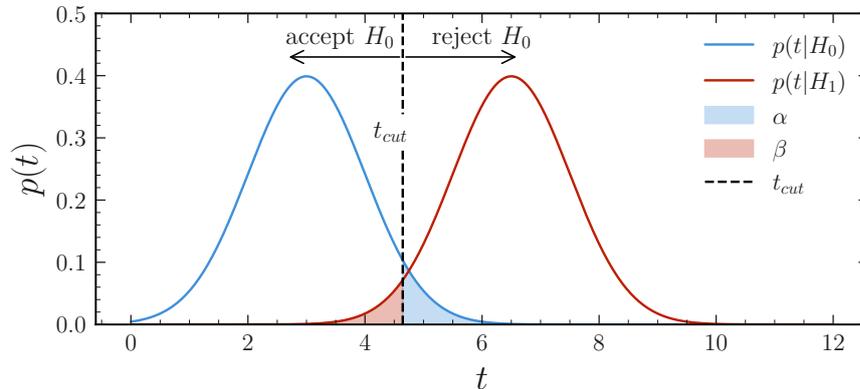


Figure 5.1: The PDFs of the test statistic, given two different hypotheses. The chosen value of $\alpha = 0.05$ defines t_{cut} , which separates the rejection region (right) from the acceptance region (left). The integrals of the shaded regions correspond to α and β , respectively.

Any hypothesis test involves two types of errors: rejecting H_0 when it is true (*Type I error*, probability α), and failing to reject H_0 when H_1 is true (*Type II error*, probability $\beta = \int_{-\infty}^{t_{cut}} dt p(t|H_1)$). The quantity $1 - \beta$ is the *power* of the test. An ideal test statistic maximizes power by clearly separating the PDFs of H_0 and H_1 .

Given a set of observations, an observed value for the test statistic, t_{obs} , can be calculated. Subsequently, the probability of obtaining a result as compatible or less with H_0 than the one actually *observed* can be determined,

$$P(t_{obs}|H_0) = \int_{t_{obs}}^{\infty} dt p(t|H_0) = 1 - F(t_{obs}|H_0), \quad (5.10)$$

which is known as the *P-value*.¹ This is illustrated in [Figure 5.2](#). If $P(t_{obs}|H_0) < \alpha$, one rejects the null hypothesis. If $P(t_{obs}|H_0) > \alpha$, one does not reject the null hypothesis.

5.2.1.1 Weighted *P*-values

If there is low statistical power, but one observes $P(t_{obs}|H_0) < \alpha$, should the null hypothesis still be rejected? The power of the test is a measure of how well the test can discriminate between the null and the alternative hypothesis. If the power is low, the test is not sensitive enough to prioritize one hypothesis over the other. In this case, the null hypothesis should not be rejected. To protect against this case, weighted *P*-values can be considered.

Consider the case, where the hypothesis to be excluded is the null hypothesis H_0 , which describes the case where signal plus background is observed ($H_0 = s + b$), according to

¹Not to be confused with the symbol for probability used in the previous sections

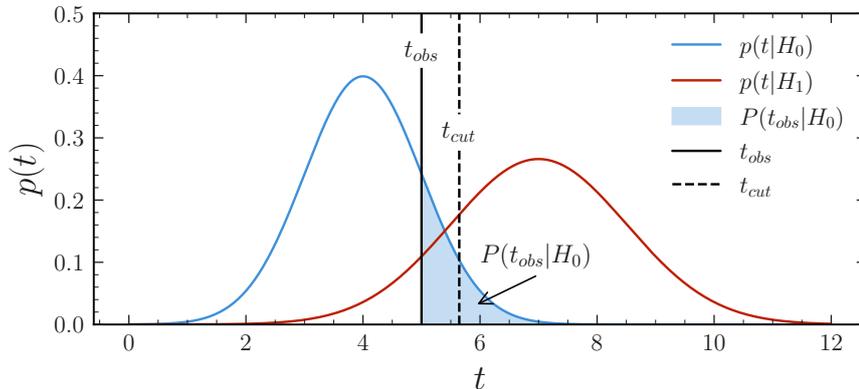


Figure 5.2: The **PDFs** of the test statistic, given two different hypotheses. The chosen value of $\alpha = 0.05$ defines t_{cut} , which separates the rejection region from the acceptance region. The observed value of the test statistic t_{obs} defines the lower bound of the integration for the P -value.

some model. The alternative hypothesis H_1 describes the case where background only is observed ($H_1 = b$).

A P -value can be defined

$$P(t_{obs}|s+b) = \int_{t_{obs}}^{\infty} dt p(t|s+b), \quad (5.11)$$

which is the probability of obtaining a measurement result as compatible or less with the $s+b$ hypothesis.

Given the choice of a rejection region, or equivalently a value of α , the $s+b$ hypothesis is excluded if

$$P(t_{obs}|s+b) < \alpha. \quad (5.12)$$

The problem with this approach is that hypotheses to which there is little or no sensitivity might be excluded. By sensitivity, it is meant that the power of the test is low, and the **PDFs** of the test statistic are almost indistinguishable, or in other words largely overlapping. This could be the case if, for example, there are very few signal events compared to the number of background events. In this case, the **PDFs** will be very similar.

To protect against this problem, $P(t_{obs}|s+b)$ can be compared to

$$P(t_{obs}|b) = \int_{t_{obs}}^{\infty} dt p(t|b), \quad (5.13)$$

which is the probability of obtaining a measurement result as compatible or less with the b hypothesis. One can define the *weighted* P -value

$$CL_s \equiv \frac{P(t_{obs}|s+b)}{P(t_{obs}|b)} < \alpha \quad (5.14)$$

as exclusion condition [84]. This condition takes the b hypothesis into account, which is equivalent to saying “the $s + b$ hypothesis is excluded only if the result is sufficiently more compatible with the b hypothesis”. If there is similar compatibility of the measurement with both of the hypotheses, i.e. a lack of sensitivity, $s + b$ should not be excluded.

If the two PDFs $p(t|s + b)$ and $p(t|b)$ are widely separated, then $P(t_{obs}|b)$ will be close to unity and $CL_s \approx P(t_{obs}|s + b)$. Contrarily, if the two PDFs are largely overlapping, then $P(t_{obs}|b) \approx P(t_{obs}|s + b)$ and $CL_s \approx 1$.

Equation (5.14) shows that $CL_s \geq P(t_{obs}|s + b)$ in all possible cases. The hypotheses excluded by $CL_s < \alpha$ are a subset of the ones excluded by $P(t_{obs}|s + b) < \alpha$. Hence, the criterion $CL_s < \alpha$ limit is *weaker* than the one defined by $P(t_{obs}|s + b) < \alpha$. This is the reason why the CL_s criterion is classified as *conservative*.

5.2.1.2 Functional form of the test statistic

In general, one wants to find a test statistic with optimal power for a given size α . The *Neyman-Pearson lemma* states that the likelihood ratio,

$$\lambda = \frac{p(\mathbf{x}|H_0)}{p(\mathbf{x}|H_1)}, \quad (5.15)$$

leads to the most powerful rejection region. A proof can be found in [Appendix B.7](#). Based on this a test statistic can be defined as

$$t = -2 \log \lambda. \quad (5.16)$$

Most commonly used test statistics are based on the likelihood ratio [79].

The null hypothesis is rejected if the data are too compatible with the alternative hypothesis,

$$t = -2 \log \lambda > t_{cut}, \quad (5.17)$$

A nice property of the likelihood ratio test is that it is invariant under reparametrization of the data,

$$\frac{p(y | H_0)}{p(y | H_1)} = \frac{|J_f|^{-1} p(x = f^{-1}(y) | H_0)}{|J_f|^{-1} p(x = f^{-1}(y) | H_1)} = \frac{p(x | H_0)}{p(x | H_1)}. \quad (5.18)$$

The Jacobian factors $|J_f|^{-1}$ cancel. This is a very useful property, as it allows the choice of the most convenient parameterization for the problem, without any loss of power.

5.2.1.3 Distribution of the test statistic

One technical question remains — what is the PDF of the test statistic under a hypothesis H , parametrized by θ , $p(t|\theta)$?

There is no unique answer to this question. It depends on the particular analysis case under consideration. A general, but also computing intensive method of finding $p(t|\theta)$ for a general functional form of t is through MC methods [67]. For a given parameter point θ , the procedure goes as follows:

1. Sample a large set of data points from the likelihood

$$\mathbf{x} \sim p(\mathbf{x}|\theta). \quad (5.19)$$

2. For each sampled dataset \mathbf{x} , compute the test statistic $t(\mathbf{x})$.
3. Construct the empirical distribution of the test statistic $p(t|\theta)$ from the computed values, through kernel density estimation or histograms.

This procedure can be computationally intensive, especially for high-dimensional parameter spaces and functional forms of the test statistic, which involve optimisation steps in the evaluation (e.g. likelihood ratios).

In [Section 5.1.1.1](#) the asymptotic behaviour of maximum likelihood estimators was discussed. Similarly, one can utilize asymptotic approximations to derive a functional form for the **PDF** of a test statistic. References [\[79, 85, 86\]](#) provide further discussions and collection of asymptotic distributions for a number of important test statistics.

5.2.1.4 Nested hypotheses

In the case where the null hypothesis is a special case of the alternative hypothesis $H_0 \subset H_1$, H_0 is said to be *nested* within H_1 .

For example the **POI** could be a signal strength² μ and the nuisance parameters could be the background normalization ψ . Assume the null hypothesis to be the background only hypothesis, $H_0 : \mu_0 = 0$, and the alternative hypothesis to be the signal plus background hypothesis, $H_1 : \mu > 0$. The value of ψ is not of interest.

To compare the hypotheses, it is necessary to compare how well the data are described by the **POI** of the null hypothesis compared to any other **POI** value μ_0 , irrespective of the nuisance parameters. In this case the likelihood ratio test statistic is [\[79\]](#)

$$t(\mu_0) = -2 \log \lambda(\mu_0) = -2 \log \frac{p(\mathbf{x}|\mu_0, \hat{\psi})}{p(\mathbf{x}|\hat{\mu}, \hat{\psi})}. \quad (5.20)$$

The $\hat{\mu}$ and $\hat{\psi}$ are the *unconditional* ML estimator, which maximize the likelihood function. The $\hat{\psi}$ is the *conditional* ML estimators, which maximize the likelihood for the specified μ_0 . The procedure of finding the conditional ML estimators is known as *profiling* [\[79\]](#).

From the definition of $\lambda(\mu)$, it follows that $0 < \lambda(\mu) < 1$, with $\lambda(\mu) \rightarrow 1$ implying good agreement between data and hypothesized μ . Hence, increasing compatibility between data and μ occurs as $t(\mu) \rightarrow 0$.

This test statistic does not distinguish between $\mu_0 > \hat{\mu}$ or $\mu_0 < \hat{\mu}$. Hence, one can achieve a small *P*-value through either too large or too small values of μ_0 . This results in a two-sided confidence interval ([Section 5.4.2](#)).

²A multiplicative factor of the overall expected signal yields.

Wald [85] demonstrated that for a single **POI**, the test statistic can be approximated as

$$t(\mu) = \frac{(\hat{\mu} - \mu)^2}{\sigma^2} + \mathcal{O}(1/\sqrt{N}) , \quad (5.21)$$

where σ^2 is the variance of the estimator $\hat{\mu}$. In the large-sample limit, this variance is given by the inverse of the Fisher information, $\sigma^2 = \mathbf{I}(\mu)^{-1}$ (see [Section 5.1.1.1](#)).

If the estimator $\hat{\mu}$ is normally distributed and the $\mathcal{O}(1/\sqrt{N})$ correction term is neglected, the test statistic $t(\mu)$ asymptotically follows a χ^2 -distribution with one **d.o.f.** This result is known as *Wilks' theorem* [86]. A proof is provided in [Appendix B.8](#).

5.2.1.5 Significance

Often one converts P -values into *significance levels* Z , which describe the number of standard deviations from a Gaussian distributed variable [79],

$$P(t_{obs}|H_0) = 1 - \Phi(Z) , \quad (5.22)$$

where Φ is the **CDF** of a standard normal distribution, $\mathcal{N}(0,1)$. Inverting the above equation, the significance level is given by

$$Z = \Phi^{-1}(1 - P(t_{obs}|H_0)) , \quad (5.23)$$

where Φ^{-1} is the inverse of Φ . For a normally distributed test statistic,

$$t_{obs} - E[t|H_0] = Z\sigma , \quad (5.24)$$

inverting the definition of Φ .

The significance level hence corresponds to the number of standard deviations from the mean of the **PDF** predicted by H_0 . Commonly used benchmarks for the significance level are $Z = 3$ ($P(Z = 3) = 1.35 \cdot 10^{-3}$) for *evidence* and $Z = 5$ ($P(Z = 5) = 2.87 \cdot 10^{-7}$) for an *observation*. In the latter case, this translates to: “for repeated experiments one expects one out of $1/2.87 \cdot 10^{-7} = 3.48 \cdot 10^6$ measurement outcomes to show a deviation as large as the one observed”.

5.2.2 Global model comparison: Bayes factors

A global model comparison focuses on the model, independently of a given parameter point. Here, the objective is to compare probabilities of a theoretical model, or hypothesis³ H , given the observed data \mathbf{x} , $p(H|\mathbf{x})$. Using Bayes theorem, one can write

$$p(H|\mathbf{x}) = \frac{p(\mathbf{x}|H)p(H)}{p(\mathbf{x})} . \quad (5.25)$$

³In this context, the term *model* is more commonly used. Here, models are referred to as hypotheses, to draw the parallels with the pointwise comparison in [Section 5.2.1](#).

If one wants to compare two hypotheses H_0 and H_1 , one common option is to compare the ratio

$$\frac{p(H_1|\mathbf{x})}{p(H_0|\mathbf{x})} = \frac{p(\mathbf{x}|H_1)p(H_1)}{p(\mathbf{x}|H_0)p(H_0)}. \quad (5.26)$$

Commonly, one assigns the same probability to either hypothesis, such that $p(H_1) = p(H_0)$. The *marginal likelihood* or *evidence* $p(\mathbf{x}|H)$ is obtained by marginalizing the likelihood over the full, or a subset of the parameter space, weighted by the prior,

$$p(\mathbf{x}|H) = \int d^M\theta p(\mathbf{x}|\theta, H)p(\theta|H). \quad (5.27)$$

This is the probability of the observed data assuming H is the true underlying hypothesis. The Bayes factor is defined as the ratio of marginal likelihoods [68],

$$B = \frac{p(\mathbf{x}|H_1)}{p(\mathbf{x}|H_0)}. \quad (5.28)$$

A Bayes factor of $B > 1$ favours hypothesis H_1 , whereas $B < 1$ favours H_0 . More specifically, Jeffreys [87] has introduced a scale for interpreting the strength of evidence provided by the Bayes factor, shown in Table 5.1.

Table 5.1: Jeffreys’s scale for interpreting Bayes factors. The ranges of $\log_{10} B$ indicate the strength of evidence in favour of hypothesis H_1 over H_0 [87].

| $\log_{10} B$ | Evidence for H_1 over H_0 |
|---------------|-------------------------------|
| 0.0 – 0.5 | Barely worth mentioning |
| 0.5 – 1.0 | Substantial |
| 1.0 – 1.5 | Strong |
| 1.5 – 2.0 | Very strong |
| > 2.0 | Decisive |

When interpreting the Bayes factor, there is one important aspect to consider: the prior sensitivity. Whereas the support of priors for unconstrained parameters is less critical for the posterior result, it matters largely for the Bayes factor [88]. This is because the *average* likelihood in the *full* parameter space is being computed. Hence, if the parameter space is increased to regions where the likelihood is very small, a smaller marginal likelihood will be obtained. This is in general an advantage, as it penalizes overly complex models. On the other hand, it ties the prior choice even tighter with the model.

To illustrate this, consider the case of a model hypothesis H_1 , parametrized by one parameter θ , and another unparametrized model hypothesis H_0 [89]. In addition, assume that the model hypothesis H_0 is nested within H_1 , in other words there exists a parameter

point θ_0 where $p(\mathbf{x}|\theta_0, H_1) = p(\mathbf{x}|H_0)$. Take the prior of θ to be uniform with a width of $\Delta\theta$ and the likelihood $p(\mathbf{x}|\theta, H_1)$ to be sharply peaked around the value $\bar{\theta}$, with a width of $\delta\theta$, which lies in the region of support of the prior. Under these assumptions, the marginal likelihood can then be approximated as

$$p(\mathbf{x}|H_1) = \int d\theta p(\mathbf{x}|\theta, H_1)p(\theta|H_1) \approx p(\mathbf{x}|\bar{\theta}, H_1) \frac{\delta\theta}{\Delta\theta}. \quad (5.29)$$

The Bayes factor is

$$B = \frac{p(\mathbf{x}|H_1)}{p(\mathbf{x}|H_0)} \approx \frac{p(\mathbf{x}|\bar{\theta}, H_1)}{p(\mathbf{x}|\theta_0, H_1)} \frac{\delta\theta}{\Delta\theta}. \quad (5.30)$$

The ratio

$$\frac{p(\mathbf{x}|\bar{\theta}, H_1)}{p(\mathbf{x}|\theta_0, H_1)} \geq 1, \quad (5.31)$$

since the data can always be fit equally well or slightly better with an additional parameter.

The second term in [Equation \(5.30\)](#), $\delta\theta/\Delta\theta$, sometimes referred to as the *Occam's razor* term, penalizes the more complex model H_1 for including parameter space which is unfavoured by the data. Hence, the Bayes factor only favours more complex models, when the likelihood ratio compensates for the penalty of the Occam's razor term.

5.3 Goodness-of-fit

Goodness-of-fit testing can be interpreted as a hypothesis test against the optimal model — namely, the model that reproduces the data exactly. A well motivated measure for the goodness-of-fit is the P -value, calculated from the [PDF](#) of the test statistic $p(t_{\text{gof}})$, where

$$P_{\text{gof}} = \int_{t_{\text{obs}}}^{\infty} dt_{\text{gof}} p(t_{\text{gof}}) = 1 - \text{CDF}_p(t_{\text{obs}}), \quad t_{\text{gof}} = -2 \ln \frac{p(\mathbf{x}|\hat{\boldsymbol{\theta}})}{p_{\text{sat}}}. \quad (5.32)$$

Here $p(\mathbf{x}|\hat{\boldsymbol{\theta}})$ is the best-fit likelihood. The saturated likelihood, p_{sat} , is the best possibly achievable likelihood [\[90\]](#).

The value of t_{gof} indicates how close one comes to the saturated (“perfect”) value of the likelihood (smaller is better, $t_{\text{gof}} = 0$ would correspond to a “perfect” fit). The saturated likelihood has zero free parameters.

For a Gaussian likelihood with N independent data points with means μ_i and variances σ_i^2 , the saturated model sets $x_i = \mu_i$, and the goodness-of-fit statistic reduces to the familiar χ^2 form $t_{\text{gof}} = \sum_i (x_i - \mu_i)^2 / \sigma_i^2$. Each standardized residual contributes a χ^2 with 1 [d.o.f.](#). By independence $t_{\text{gof}} \sim \chi_N^2$, and fitting m parameters reduces the [d.o.f.](#) to $N - m$. A derivation is given in [Appendix B.8](#).

For the HistFactory likelihood $p(\mathbf{n}, \mathbf{a}|\boldsymbol{\eta}, \boldsymbol{\chi})$ ([Section 4.4](#)) the saturated likelihood is

$$p_{\text{sat}}(\tilde{\boldsymbol{\chi}}) = \text{Pois}(\mathbf{n} | \mathbf{n}) C(\mathbf{a} | \tilde{\boldsymbol{\chi}}), \quad (5.33)$$

which is the product of the Poisson likelihood with expectation values equal to the observed data counts and the constraint likelihood with parameters $\tilde{\chi}$ maximizing the constraint term. In the asymptotic limit, $p(t_{\text{gof}})$ will equal a χ^2 -distribution with $N - m$ **d.o.f.**, where in this case N is the number of bins in the model and m the number of free parameters η .

5.4 Limit setting

In interval estimation, the goal is to find a parameter range

$$\theta_a \leq \theta \leq \theta_b , \quad (5.34)$$

which contains the true value of the parameter θ_0 with a certain probability $1 - \alpha$,

$$P(\theta_a \leq \theta_0 \leq \theta_b) = 1 - \alpha . \quad (5.35)$$

In an idealized setting, one assumes access to the sampling distribution of the estimator θ , given the true parameter value θ_0 , denoted by $p(\theta|\theta_0)$. Under this assumption, the interval can be determined by solving

$$P(\theta_a \leq \theta_0 \leq \theta_b) = \int_{\theta_a}^{\theta_b} d\theta p(\theta|\theta_0) = 1 - \alpha . \quad (5.36)$$

Hence, one could directly determine the interval boundaries θ_a and θ_b from the parameter **PDF**. Note that such an interval is not necessarily unique. However, since the true distribution $p(\theta|\theta_0)$ is typically unknown in practice, this represents an idealized scenario.

The Bayesian approach to interval estimation comes closest to this scenario. Here one does not have access to the **PDF** $p(\theta|\theta_0)$, but to the parameter **PDF** $p(\theta|\mathbf{x})$. In this case the obtained interval is known as a *credible interval*. This is discussed in [Section 5.4.1](#).

In the frequentist case, one constructs intervals for a large range of parameter values, utilizing the **PDF** of a test statistic $t(\mathbf{x})$. This is known as *Neyman construction* and results in *confidence intervals* of the parameter. This is discussed in [Section 5.4.2](#).

5.4.1 Credible intervals

Bayesian credible intervals are defined as the range of parameter values that contain the true parameter value with a certain probability $1 - \alpha$, given the observed data. Given the posterior distribution $p(\theta|\mathbf{x})$, the credible interval can be found by determining the interval boundaries θ_a and θ_b such that

$$\int_{\theta_a}^{\theta_b} d\theta p(\theta|\mathbf{x}) = 1 - \alpha . \quad (5.37)$$

The credible interval is then given by $[\theta_a, \theta_b]$. Hence, Bayesian intervals provide a direct probabilistic interpretation, namely, the probability that the true parameter value lies within the interval is equal to $1 - \alpha$.

Generally, this interval is not unique. There are multiple options for defining the boundaries, leading to different credible intervals:

- **Highest Density Interval (HDI)**: This is the smallest possible interval, containing $1 - \alpha$ probability. It has the property that the probability density within the interval is never lower than that outside the interval [68]. This includes all parameter values $\theta : p(\theta|\mathbf{x}) > W$, for some W , such that [91]

$$\int_{\theta:p(\theta|\mathbf{x})>W} d\theta p(\theta|\mathbf{x}) = 1 - \alpha . \quad (5.38)$$

- **Central interval**: This is an interval $[\theta_a, \theta_b]$, centred around a specific value, usually the posterior mode or mean.
- **Equal-tailed interval**: This interval ensures the same probability mass in the upper and lower tails of the posterior distribution. One calculates the boundaries such that

$$P(\theta < \theta_a|\mathbf{x}) = P(\theta > \theta_b|\mathbf{x}) = \frac{\alpha}{2} . \quad (5.39)$$

- **Upper/lower interval**: These one-sided intervals specify either an upper or lower bound on the parameter, containing a given posterior probability mass. For an upper interval,

$$P(\theta < \theta_b|\mathbf{x}) = 1 - \alpha, \quad (5.40)$$

and for a lower interval,

$$P(\theta > \theta_a|\mathbf{x}) = 1 - \alpha. \quad (5.41)$$

These are particularly useful in setting exclusion bounds.

In the case of a multidimensional parameter space $\boldsymbol{\theta} \in \mathcal{R}^M$, the above results are easily generalizable. In such cases the credible interval spans over a region $R \subset \mathcal{R}^M$ in parameter space, such that

$$\int_R d^M\theta p(\boldsymbol{\theta}|\mathbf{x}) = 1 - \alpha . \quad (5.42)$$

Further, one can obtain credible intervals for a subset of parameters $\boldsymbol{\theta}' \subseteq \boldsymbol{\theta}$ by marginalizing the posterior over all but the **POIs**.

5.4.2 Confidence intervals

Frequentist confidence intervals are constructed by inverting the probability statement in [Equation \(5.35\)](#) — that is, instead of asking what is the probability that a fixed interval contains the true parameter, one constructs an interval that, in repeated experiments, would contain the true value with the specified coverage probability. This method is known as *Neyman construction* [67].

This inversion can be achieved by summarizing the data with a test statistic $t(\mathbf{x})$, with a PDF $p(t|\theta)$, and then solving for the interval boundaries $t_a(\theta)$ and $t_b(\theta)$ for a given point θ ,

$$P(t_a < t < t_b|\theta) = \int_{t_a}^{t_b} dt p(t|\theta) = 1 - \alpha. \quad (5.43)$$

Repeating this for a range of θ values, one constructs the confidence intervals in t -space $[t_a(\theta), t_b(\theta)]$.

The uniqueness of these intervals can be given by further specifying interval properties such as:

- **Central interval:** Here the interval is constructed to have equal probability in the upper and lower tails,

$$\int_{-\infty}^{t_a} dt p(t|\theta) = \int_{t_b}^{\infty} dt p(t|\theta) = \frac{\alpha}{2}. \quad (5.44)$$

- **Upper/lower interval:** These intervals specify either an upper or lower bound on the parameter. For an upper bound,

$$\int_{-\infty}^{t_b} dt p(t|\theta) = 1 - \alpha, \quad (5.45)$$

and for a lower interval,

$$\int_{t_a}^{\infty} dt p(t|\theta) = 1 - \alpha. \quad (5.46)$$

These are particularly useful in setting exclusion bounds.

Suppose the confidence intervals $[t_a(\theta), t_b(\theta)]$ have been constructed for a range of θ values. These can be inverted to obtain the confidence intervals in θ -space, $[\theta_a(t), \theta_b(t)]$. The required confidence interval for θ is then given by the intersection of an observed value t_{obs} with $t_a(\theta)$ and $t_b(\theta)$, or equivalently by $[\theta_a(t_{\text{obs}}), \theta_b(t_{\text{obs}})]$. This is illustrated in [Figure 5.3](#).

To verify that the interval $[\theta_a, \theta_b]$ is truly the desired confidence interval, suppose a large number of experiments are performed. The proportion of values $t(\mathbf{x})$ falling into a range $[t_a(\theta_0), t_b(\theta_0)]$, for the true value θ_0 , is $1 - \alpha$ by definition. Therefore, the proportion of intervals $[\theta_a, \theta_b]$, corresponding to the t values in $[t_a(\theta_0), t_b(\theta_0)]$, covering the true value θ_0 is $1 - \alpha$, that is

$$P(\theta_a < \theta < \theta_b) = 1 - \alpha. \quad (5.47)$$

A method which gives an interval $[\theta_a, \theta_b]$, satisfying [Equation \(5.36\)](#), possesses the property of *coverage*. As long as [Equation \(5.43\)](#) is satisfied, the Neyman construction method has exact coverage [\[67\]](#).

The extension to a multidimensional parameter space is more complicated for the frequentist case. This is mainly due to the fact that one must find the distribution of the test statistic in multiple dimensions. In a limited number of cases, this can be approximated asymptotically [\[67\]](#). If asymptotics are not applicable, the **MC** approach is only feasible for a small number of parameters.

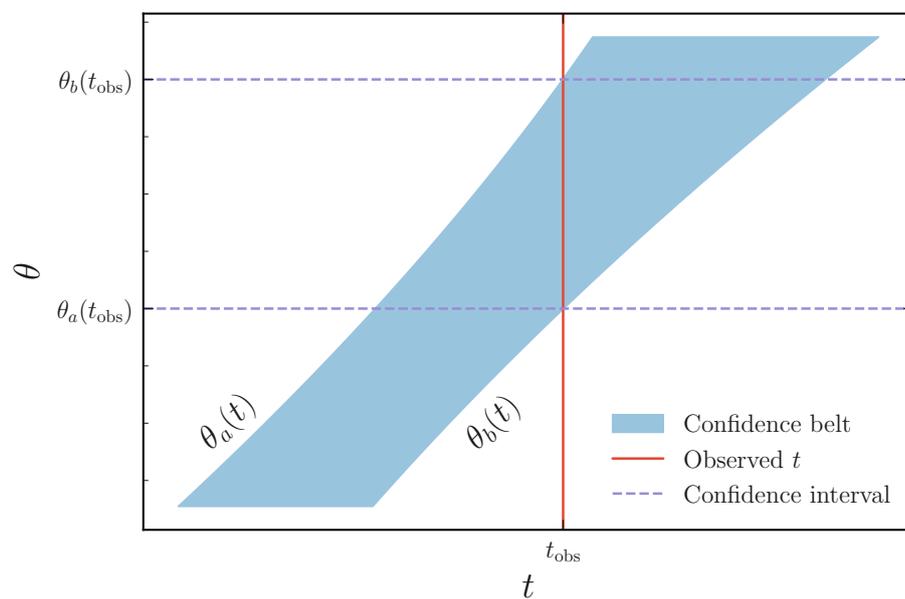


Figure 5.3: Neyman construction of confidence intervals. The confidence interval is constructed horizontally by scanning through values of θ , determining $t_a(\theta)$ and $t_b(\theta)$ such that $P(t_a(\theta) < t < t_b(\theta)) = 1 - \alpha$.

Chapter 6

Ongoing reinterpretation efforts in particle physics

The reinterpretation of published analysis results has emerged as a transformative methodology in **HEP**, allowing theorists and experimentalists alike to extract maximal value from existing datasets. By reconstructing and modifying statistical models, this practice enables comprehensive exploration of **SM** and **BSM** scenarios without necessitating new data collection or analysis designs. Reinterpretation also facilitates precision measurements, model comparisons, and global fits [69].

Collider experiments are unique, resource-intensive endeavours. Once an analysis is published, the original authors often move on, making further direct access to their insight and methods difficult. Thus, a structured strategy for the preservation and reuse of analysis data products and methods is essential to long-term scientific productivity [92–94].

This chapter reviews the essential components of reinterpretation infrastructure: data preservation (Section 6.1), analysis preservation (Section 6.2), statistical model serialization (Section 6.3), classification of reinterpretation approaches (Section 6.4), and available public tools (Section 6.5). Finally, Section 6.6 discusses remaining challenges and future goals for the community.

6.1 Data preservation

Data are the currency of the modern **HEP** community. Experimental collaborations highly value their collected collision data and rarely release it to the public. This is understandable, given the amount of work and dedication needed to run colliders, detectors, and finally collect the data.

Experimental collaborations, such as ATLAS and CMS, have begun releasing data through the CERN Open Data portal [95], often adhering to the **FAIR** data principles. Due to the complexity of **HEP** analyses, however, **FAIR** compliance alone does not ensure reproducibility of measurement results [93].

Collaborations not only benefit from large computing resources, allowing them to process the data in the first place, but also from internal expertise, which allows for correct analyses and interpretations. These computation- as well as knowledge-bottlenecks limit the usability of large open datasets for the wider **HEP** community.

Smaller, analysis-specific data products are more restricted in their versatility, but benefit greatly from their usability and accessibility. In this case, it is recommended to preserve [92]:

- **Event-level data**, including both reconstructed data and **MC** simulations.
- **Generator-level data**, necessary for fast reinterpretation pipelines and Rivet-based workflows [96].
- **Analysis code and logic**, which includes the scripts and frameworks used to perform the selection and statistical interpretation.
- **Detector performance and response**: Either through unfolding or via response matrices.
- **Full statistical models**, which are the full probability models $p(\mathbf{x}|\boldsymbol{\theta})$, with transparent treatment of nuisance parameters and their correlations.
- **Analysis data products**, which encode derived data, such as selection efficiencies and limits.

Having these data products in addition to a full analysis preservation, discussed in **Section 6.2**, ensures the long-term reproducibility of experimental results and increases their credibility.

6.2 Analysis preservation

Analysis preservation allows for the processing of the preserved data products (**Section 6.1**). This enables, on the one hand, the reproduction of the analysis data products or results, and on the other hand, modifications to the analysis and/or input data, enabling reinterpretation.

Either the full analysis workflow — from the initial collider data to the final result — or a lightweight approximation of the analysis can be preserved, each coming with its own set of advantages and disadvantages.

Full analysis preservation This involves storing the complete analysis software stack, including detector simulation, reconstruction, and physics analysis code, ideally within containerized or virtualized environments. While this is the most accurate form of preservation, it is also the most computationally expensive and hardest to maintain due to the diversity of constantly evolving software ecosystems (even within a single experiment).

Lightweight preservation More suitable for rapid reinterpretation over large model spaces, lightweight formats omit full reconstruction details in favour of simplified implementations. These are typically based on fast simulations, cut-and-count approximations, or generator-level analysis, and are easier to share publicly.

Both methods serve complementary roles. Lightweight approaches enable rapid exploration of high-dimensional parameter spaces and identification of regions compatible with data, with combinations of multiple analyses further boosting sensitivity at minimal computational cost. Full analysis preservation becomes valuable once interesting regions are identified and precision is prioritized.

This work presents a reinterpretation method that bridges the gap between generator-level and reconstruction-level observables, combining the computational efficiency of lightweight approaches with accurate results. Reinterpretation is particularly valuable for exploring large theory parameter spaces rather than for precision measurements at individual points. Since experimental analyses are optimized for specific scenarios (typically the **SM** or particular **BSM** models), reinterpretation studies sacrifice some sensitivity compared to dedicated analyses. However, combining multiple lightweight analyses can achieve competitive or even superior sensitivity across broader parameter regions — a capability unattainable with resource-intensive full-stack approaches. The identification of interesting parameter space regions through reinterpretation should motivate dedicated experimental analyses, thereby strengthening the feedback loop between theoretical and experimental communities in the search for new physics.

6.3 Statistical models and their preservation

Reinterpretation efforts are mostly hindered by the lack of access to full statistical models. This lack of availability can be attributed to the technical challenges that come with the serialization of statistical models. These challenges lie mainly in the complexity of the models and the need for high customization to the underlying analysis. The earliest reusable models were often implemented in `RooFit` [75] and stored in `RoosWorkspaces`, which offered a structured way to encode complex likelihoods within the `ROOT` framework. However, these formats were tightly coupled to C++/`ROOT` environments, limiting portability and long-term preservation. The introduction of the `pyhf` library [71, 72], a pure Python implementation of the `HistFactory` framework [70], has enabled the preservation and reuse of full statistical models in a platform-independent, JSON-based format.

To support long-term preservation and interoperability of statistical models, the HEP Statistics Serialization Standard (**HS3**) [97] has been proposed. **HS3** defines a standard JSON-based schema for encoding statistical models, ensuring compatibility with existing tools such as `pyhf` [71, 72], `RooFit` [75], and `zfit` [98]. It also supports the inclusion of metadata and documentation that describes the analysis context. The specification is maintained in a public repository by the community [97], promoting transparency and ease of access.

6.4 Types of reinterpretation

One can classify reinterpretation efforts as follows [34, 93]¹:

- *Kinematic reinterpretation* or *recasting*, which includes testing an alternative physics process with different kinematic distributions. This typically requires propagating the new signal hypothesis through the detector simulation or applying response matrices to account for changes in efficiencies and acceptance regions. It is the most demanding form of reinterpretation and often relies on external tools.
- *Model updating*, which refines either theoretical predictions or experimental inputs. The main goal is an overall reduction of uncertainties. Technically, this can be viewed as a subclass of kinematic reinterpretation where the analysis logic remains unchanged.
- *Combinations* of datasets and measurements across experiments or channels. This is crucial for reducing uncertainties on shared parameters or for deriving global parameter constraints, where different channels may have different sensitivities to some parameters. For such combinations, it is necessary that the underlying model assumptions are mutually consistent.

6.5 Public tools for reinterpretation

Several public frameworks have been developed to facilitate the reinterpretation of HEP data. A representative overview is provided in Reference [94].

RECAST (full²) Allows theorists to submit signal models, which are recast by experimental workflows within experimental collaborations, using full simulation and statistical frameworks. Provides high accuracy at the cost of internal computational resources [99].

CheckMATE (full) Performs full MC simulations including detector effects, and applies implemented experimental analyses with their original cuts. It yields ratios of predicted to excluded cross-sections [100, 101].

MadAnalysis 5 (full) Allows both detector-level recasting and generator-level analyses with user-defined cuts or integrated public analyses. Input includes MC events and analysis code. Computational cost varies from moderate to high depending on simulation depth. Inference outputs include CL_S exclusion limits [102, 103].

GAMBIT / ColliderBit (full/lightweight) Embedded within the global-fitting framework GAMBIT, ColliderBit handles event generation and fast detector simulation, integrating the results into global fits. Provides simplified likelihoods for calculating exclusion limits [104].

¹This section is based on the review article [93], reformulated in Reference [34].

²See Section 6.2.

- Rivet** (lightweight) Designed for validating **MC** generators. Applies truth-level particle-level analyses to generator-level events, relying on published measurements or unfolded distributions. It provides truth and detector level distributions [96].
- Contur** (lightweight) Builds on Rivet and utilizes unfolded (particle-level) differential cross-sections to derive exclusion contours in **BSM** model space, using existing measurements [105].
- pyhf** (full) Performs reinterpretation via JSON-serialized full statistical models (**HistFactory**), allowing complete likelihood recalculations without **MC** simulation. Requires a **pyhf** JSON model, signal templates, and background data. Computational cost is low to moderate, depending on the complexity of likelihood evaluations. Requires full simulation and analysis for additional signal models [70, 71].
- mapyde** (full/lightweight) Strings together event generation, fast detector simulation, analysis selection, and likelihood interpretation in an automated pipeline. Input includes model parameters, detector acceptances and selection efficiencies, and **pyhf** models. It requires high computational cost but offers end-to-end reproducibility [106, 107].
- SModels** (lightweight) Decomposes **BSM** models into simplified model spectra [108] and confronts them with a database of experimental limits from **LHC** searches, using efficiency maps and published cross-section constraints. Provides a simplified likelihood for inference. Operates with low computational cost as no **MC** generation is involved, and is ideal for rapid screening of large model spaces [109, 110].

Each tool occupies a distinct niche in the reinterpretation landscape. Lightweight methods excel at speed and broad survey capability, while full-stack or hybrid tools offer fidelity and statistical rigour. Tools such as **pyhf** bridge between these paradigms by facilitating direct likelihood reinterpretation without simulating events.

The reinterpretation framework presented in this work occupies a unique position in this landscape, combining the computational efficiency of lightweight approaches with the statistical rigour of full likelihood-based methods. Unlike purely generator-level tools such as **Rivet** and **Contur**, this method incorporates detector-level effects through parametrized response functions, bridging the gap between truth-level and reconstruction-level observables without requiring full detector simulation. This approach maintains compatibility with **pyhf** statistical models while enabling rapid exploration of high-dimensional parameter spaces — a capability that distinguishes it from resource-intensive full-stack tools like **CheckMATE** or **RECAST**.

6.6 Challenges and future goals

Despite major progress in reinterpretation infrastructure, significant challenges remain. A key difficulty is the diversity of analysis implementations, statistical frameworks, and data formats. Even within a single experiment, variations in software environments and

detector simulation chains make consistent preservation and reuse difficult [93].

Another challenge is the resource-intensive nature of full reinterpretations. Detector-level reinterpretation requires access to full simulation and reconstruction chains, which is often impractical due to computing demands. This motivates the development of lightweight reinterpretation approaches, such as fast simulation or parameterized efficiencies. However, such methods may sacrifice precision or require validation against results based on fully-preserved analyses.

Machine-learning-based analyses introduce new complexities for reinterpretation. Sharing a trained model requires not only the architecture and weights, but also a detailed specification of input features, preprocessing, and software dependencies [93]. Standardization and tooling for machine learning preservation are still in early stages.

Future goals include [93]:

- Establishing community-wide standards for analysis preservation and statistical model sharing.
- Facilitating reinterpretation-aware design of new analyses, including public documentation of selection logic and systematics.
- Creating a centralized, searchable portal of recastable analyses with accompanying statistical models and metadata.
- Enabling reinterpretation-driven model testing in global fits to identify gaps in experimental coverage.

Chapter 7

Model-agnostic likelihoods for the reinterpretation of particle physics results

¹ Building upon the review of existing reinterpretation approaches, this chapter presents a novel method that addresses many of the limitations identified in current reinterpretation strategies. The approach bridges the gap between computationally expensive full reanalysis and potentially biased simplified methods by introducing a technique based on so-called *model-agnostic likelihoods*.

This reinterpretation method enables direct inference on theoretical parameters while correctly accounting for changes in the kinematic distributions. The method offers several key advantages: inference is typically very fast; it requires only minimal additional information beyond the likelihood itself; it is general in terms of applicability across different models and analyses; and the resulting likelihoods are easily distributable, making them a powerful tool for enhancing the long-term impact and reusability of **HEP** measurements.

The development of this method represents a significant step forward in making experimental results more accessible to the broader theoretical physics community, thereby increasing the scientific return on the substantial investments in experimental particle physics.

This chapter is structured as follows: I first introduce the method in **Section 7.1**, followed by a discussion of its implementation in **Section 7.2**. In **Section 7.3**, I present applications of the method to two toy examples from flavour physics, the $B \rightarrow K\nu\bar{\nu}$ and $B \rightarrow K^*\nu\bar{\nu}$ decays. Finally, I discuss the significance and implications of the method in **Section 7.4**.

¹The material presented in this chapter has been published in the European Physical Journal C [34]. While the content largely follows that of the referenced paper, editorial modifications have been made, including restructuring of the presentation, rephrasing of portions of the text, and expansion of selected explanations. As this project is the result of collaborative work, the specific contributions of each co-author are documented in **Appendix C.1**.

7.1 Reinterpretation through reweighting

The reinterpretation method described here is based on updating the distributions of the observable variables, given changes in the underlying kinematic distribution.

The **PDF** of reconstructed events $p(x)$ results from folding the **PDF** of a theoretical kinematic prediction $p(z)$ with the conditional distribution $p(x|z)$ and the indicator function $\mathbb{1}_\varepsilon(x)$,

$$p(x) = \frac{1}{\varepsilon} \int dz \mathbb{1}_\varepsilon(x) p(x|z) p(z). \quad (7.1)$$

Here, the reconstruction variable x represents one or multiple observable variables, and the kinematic variable z represents one or multiple kinematic **d.o.f.**. The function $\mathbb{1}_\varepsilon(x)$ models the selection criteria for a reconstructed event, and $p(x|z)$ is the conditional probability of measuring a reconstructed configuration x given an underlying particle configuration z . The overall reconstruction efficiency ε acts as a normalization factor for the **PDF** $p(x)$. The **PDF** $p(z)$ corresponds to the normalized kinematic distribution of a theoretical prediction $\sigma(z)$,

$$p(z) = \frac{\sigma(z)}{\sigma}. \quad (7.2)$$

The number density of expected events, given a total integrated luminosity L , is $\nu(x) = L \sigma \varepsilon p(x)$ and reads

$$\nu(x) = L \int dz \varepsilon(x|z) \sigma(z) = \int dz \nu(x, z), \quad (7.3)$$

where both reconstruction and selection are combined into $\varepsilon(x|z) = \mathbb{1}_\varepsilon(x) p(x|z)$, and $\nu(x, z) = L \varepsilon(x|z) \sigma(z)$ can be thought of as a *joint number density*, analogous to the joint **PDF** $p(x, z) = p(x|z) p(z)$. This represents a *folding procedure* in which the theoretical prediction is convolved with $\varepsilon(x|z)$ (and scaled by L) to obtain the expected number of events [65]. In this way, the kinematic distribution can be mapped to the reconstructed space, taking both reconstruction and selection efficiencies into account.

The reinterpretation task involves determining the number density $\nu_1(x)$ of an *alternative* theoretical prediction $\sigma_1(z)$. This can be obtained by reweighting the joint number density $\nu_0(x, z)$ according to the kinematic *null* distribution $\sigma_0(z)$ via

$$\begin{aligned} \nu_1(x) &= L \int dz \varepsilon(x|z) \sigma_1(z) \\ &= L \int dz \varepsilon(x|z) \sigma_0(z) \frac{\sigma_1(z)}{\sigma_0(z)} \\ &= \int dz \nu_0(x, z) w(z). \end{aligned} \quad (7.4)$$

The weight factor $w(z)$ is simply the ratio of the theoretically predicted alternative kinematic distribution to the null distribution.

This reweighting process solely requires knowledge of the joint null number density $\nu_0(x, z)$. Together with the weight factor, this is sufficient to predict the number density according to an alternative theory.

7.1.1 Discrete reweighting

In practical applications, the continuous joint number density is typically not analytically obtainable and requires estimation through **MC** simulations. To address this, one can discretize the reweighting method by representing the joint number density as a multidimensional matrix ν_{xz} in bins of $x \times z$. This is done alongside the binning of the theoretically predicted distribution σ_z and weight factor w_z in the kinematic **d.o.f.** z . Consequently, the discrete joint number density has dimension $\dim(x) \cdot \dim(z)$.

The discrete joint number density can be obtained by integrating the continuous joint number density over each $x \times z$ bin,

$$\nu_{xz} = \int_{\text{bin } x} \int_{\text{bin } z} dx' dz' \nu(x', z'), \quad (7.5)$$

where the integral boundaries are the bin boundaries of each x and z bin, respectively. The binned weights are given by

$$w_z = \frac{\sigma_{1,z}}{\sigma_{0,z}} = \frac{\int_{\text{bin } z} dz' \sigma_1(z')}{\int_{\text{bin } z} dz' \sigma_0(z')}. \quad (7.6)$$

The reweighting step of **Equation (7.4)** becomes

$$\nu_{1,x} = \sum_z \nu_{0,xz} w_z. \quad (7.7)$$

This discrete approach offers an advantage due to its lower computational cost compared to event-by-event reweighting. The cost of this simplification is a loss of accuracy due to the binning in the kinematic **d.o.f.** z . However, this loss is controllable by increasing the number of bins. The binning should be chosen such that the joint number density and weight function factorize approximately in each bin. In principle, an arbitrarily fine binning can be chosen such that the uncertainty due to this loss is negligible compared to other sources of uncertainty (provided enough **MC** samples are available; see **Appendix C.2**).

Crucially, only a fixed set of samples from the joint **PDF** $p_0(x, z)$ based on the null prediction is required, and no new samples need to be produced for the reinterpretation. Therefore, publishing the (binned) joint null number density and knowledge of the underlying kinematic null distribution is sufficient to perform the reinterpretation of a given measurement.

7.1.2 Limitation

The proposed reweighting approach is a lightweight and accurate way of obtaining new signal templates given a joint number density and the kinematic null distribution. Still, it does have one main limitation: if the phase space contains regions with $w(z) \gg 1$, the effective MC sample size decreases. Put differently, large weights are assigned to regions that are sparsely populated with MC samples obtained from the null distribution. Furthermore, if the null distribution lacks support in a region of phase space, $\sigma_0(z) \rightarrow 0$, it can happen that $w(z) \rightarrow \infty$ when reweighting to an alternative distribution. In this case, no MC samples exist in the given region. The only solution is to re-analyse new samples produced according to the alternative distribution.

7.2 Implementation of the reweighting method

Using the reweighting method, one can construct likelihood functions for particle physics analyses directly parametrized in terms of theory parameters. Although the reweighting method is independent of any likelihood formalism, it is showcased here in terms of the HistFactory formalism [70] as a baseline statistical model. For a review of the HistFactory formalism, I refer to Section 4.4. In terms of implementation, I use the pyhf [71, 72] package, the Python implementation of the HistFactory model, introduced in Section 4.4.3.

To obtain the data likelihood p_{data} (see Equation (4.23)) for any theoretical model, one needs to calculate the event rates of the corresponding signal template. This requires an implementation of Equation (7.7), and hence an extension of the pyhf codebase. Since the method prescribes changes to event rates, these can be applied via a multiplicative modifier. Therefore, pyhf is extended with a *custom modifier* (see Section 4.4.4.2) that computes the modifications to the nominal event rates according to the procedure in Equation (7.7). This modifier is expressed as a function of the underlying theory parameters of the alternative kinematic distribution.

The constraint likelihood $p_{\text{constr.}}$ (see Equation (4.26)) contains all constraint terms of these underlying theory parameters, which can be correlated in general. By definition, modifier parameters in pyhf are treated as uncorrelated. To correctly include correlated parameter constraints in the statistical model, the theory parameters are decorrelated using matrix eigendecomposition (discussed in Section 4.4.4.1). One normally-constrained pyhf modifier parameter is then assigned to each of these decorrelated parameters.

7.3 Application of the reweighting method to toy examples

A central aim of this work is to motivate the experimental HEP community to make use of the proposed method. This will, in turn, enable subsequent model-agnostic phenomenological analyses of experimental results. Here, I detail the full reinterpretation of two toy

examples from low-energy flavour physics.

In these toy examples, no experimental data are used, but rather two datasets of simulated samples, where one dataset acts as *real* data. The *MC data* is produced according to the **SM** prediction. The *real data* is produced by assuming that **BSM** physics affects the example process. To simulate detector and other analysis selection effects, observables in both datasets are smeared according to estimated detector resolutions, and event yields are scaled by an efficiency map.

By comparing the produced datasets using either Bayesian or frequentist methods, the aim is to recover the chosen **BSM** parameters starting from the **SM**. This is possible only because the likelihood is made a function of the theory parameters, allowing shape changes in the kinematic distribution due to **BSM** physics to be directly taken into account.

As a general result, the goal is to compute the posterior distribution in the space of theory parameters, given the two simulated datasets and prior parameter constraints. To obtain a posterior distribution from a `HistFactory` likelihood, I use the Bayesian `HistFactory` formalism, which I discuss in [Section 4.4.2](#), and is implemented in `bayesian pyhf` [74].

7.3.1 $B \rightarrow K\nu\bar{\nu}$

The recent measurements of the total rate of $B \rightarrow K\nu\bar{\nu}$ decays by the Belle II collaboration [2, 111] hint at an excess of signal events compared to the **SM** expectation. This has triggered substantial interest in the **HEP** phenomenology community to interpret this excess as a sign of **BSM** physics and to extract the corresponding model parameters [112–114]. In this subsection, I study the performance of the proposed approach using simulated $B \rightarrow K\nu\bar{\nu}$ data.

7.3.1.1 Weak Effective Theory parametrization

To capture the effects of **BSM** physics in $B \rightarrow K\nu\bar{\nu}$ decays, I work within the **WET**, which is a low-energy effective field theory that describes both **SM** and potential **BSM** effects using a common set of parameters. Details about the **WET** are given in [Section 2.2.1](#). Here, it is assumed that potential new **BSM** particles and force carriers have masses at or above the scale of electroweak symmetry breaking.

The $sb\nu\nu$ **WET** Lagrangian is constructed from the five mass-dimension 6 operators given in [Equations \(2.52\) and \(2.54\)](#). Presently, the only measured observable is the differential decay rate for $B \rightarrow K\nu\bar{\nu}$, discussed in [Section 2.4.1](#) and simulated in this example. The decay rate in terms of the Wilson coefficients of the operators is given in [Equation \(2.62\)](#). From this prediction, it is clear that this decay is sensitive only to linear combinations of vector, scalar, and tensor Wilson coefficients: $|C_{VL} + C_{VR}|$, $|C_{SL} + C_{SR}|$, $|C_{TL}|$ (see [Equation \(2.64\)](#)). The kinematic shape of the differential branching ratio, due to the vector, scalar and tensor interactions, is illustrated in [Figure 2.2](#).

The hadronic matrix elements of the contributing operators are described by the three independent hadronic form factors $f_+(q^2)$, $f_0(q^2)$ and $f_T(q^2)$, as discussed in [Section 2.3](#).

In this work, the form factors are parametrized following the BSZ parametrization [41] (see Section 2.3), which is truncated at order $K = 2$. The values and uncertainties for the corresponding 8 hadronic parameters are extracted from a joint theoretical prior PDF comprised of the 2021 lattice world average based on results by the Fermilab/MILC and HPQCD collaborations [45] and recent results by the HPQCD collaboration [46]; see Table 2.3. Correlations between the hadronic parameters are taken into account through their respective covariance matrices and implemented as discussed in Section 7.2.

For the SM prediction, only the vector left-handed Wilson coefficient C_{VL} is non-zero (see Equation (2.55)), leading to a predicted branching ratio of [37, 47]

$$\mathcal{B}(B^+ \rightarrow K^+ \nu \bar{\nu})_{\text{SM}} = (4.8 \pm 0.2) \cdot 10^{-6}. \quad (7.8)$$

The theoretical uncertainty is dominated by the form factor uncertainties from lattice QCD, with smaller contributions from CKM matrix elements and higher-order QCD corrections.

The Belle II experiment, which found the first evidence for this decay, observes more events than expected in the SM. The ratio of observed to expected events is 4.6 ± 1.3 [2] and the measured branching ratio is

$$\mathcal{B}(B^+ \rightarrow K^+ \nu \bar{\nu})_{\text{exp}} = (2.3 \pm 0.7) \cdot 10^{-5}. \quad (7.9)$$

For later use, I define a benchmark point in the space of Wilson coefficients that roughly reproduces the measured branching ratio after correcting for the efficiency. Assuming all Wilson coefficients to be real, it reads

$$\begin{aligned} |C_{VL}| &= 10, & |C_{VR}| &= 4, \\ |C_{SL}| &= 3, & |C_{SR}| &= 1, \\ |C_{TL}| &= 1. \end{aligned} \quad (7.10)$$

As shown in Equation (2.62), the decay rate of $B \rightarrow K \nu \bar{\nu}$ is only sensitive to three linear combinations of Wilson coefficients shown in Equation (2.64). The projection of this benchmark point onto this subspace reads

$$\begin{aligned} |C_{VL} + C_{VR}| &= 14, \\ |C_{SL} + C_{SR}| &= 4, \\ |C_{TL}| &= 1. \end{aligned} \quad (7.11)$$

7.3.1.2 Datasets

To make this example as realistic as possible, the setup is designed similarly to what has been done in the Belle II analysis.

The MC data are produced according to the SM prediction (*null* hypothesis) of the differential branching ratio $d\mathcal{B}(B \rightarrow K \nu \bar{\nu})/dq^2$, where the Wilson coefficients are set to the values in Equation (2.55). The number of samples produced is equivalent to the

Table 7.1: The number of $B \rightarrow K\nu\bar{\nu}$ samples produced for this study, corresponding to an equivalent of 362 fb^{-1} and 50 ab^{-1} integrated luminosity at the SuperKEKB collider. *Generated* and *reconstructed* samples correspond to the numbers prior and post efficiency correction.

| Luminosity | $B\bar{B}$ events | MC generated \ \ reconstructed | Data generated \ \ reconstructed |
|-----------------------|---------------------------|--------------------------------|----------------------------------|
| 362 fb^{-1} | $\sim 3.87 \cdot 10^8$ | $1.86 \cdot 10^3$ | $1.05 \cdot 10^4$ |
| | | $2.41 \cdot 10^2$ | $1.14 \cdot 10^3$ |
| 50 ab^{-1} | $\sim 5.35 \cdot 10^{10}$ | $2.57 \cdot 10^5$ | $1.45 \cdot 10^6$ |
| | | $3.21 \cdot 10^4$ | $1.58 \cdot 10^5$ |

number of signal events seen or expected for a given integrated luminosity. Samples are produced for 362 fb^{-1} integrated luminosity, which corresponds to the total collider data used in Reference [2], and for 50 ab^{-1} integrated luminosity corresponding to the total target luminosity of the Belle II experiment [52]. The estimated number of $B\bar{B}$ events is multiplied by the SM branching ratio of Equation (7.8) to obtain a rough estimate for the number of MC events to produce, prior to the efficiency correction (see below).

The real data are produced according to a BSM prediction (*alternative* hypothesis). Following the observations of more events than predicted in the SM [2], the previously defined benchmark point in Equation (7.10) is used. The estimated number of $B\bar{B}$ events is multiplied by the BSM branching ratio, $\mathcal{B}(B \rightarrow K\nu\bar{\nu}) \approx 2.71 \cdot 10^{-5}$ [47, 115], to obtain a rough estimate for the number of data events to produce, prior to the efficiency correction (see below).

The number of produced events is listed in Table 7.1. Unless stated otherwise, all numerical values, figures, and tables provided in the following are obtained from studies that assume the 50 ab^{-1} sample size. Samples of the decay's probability distribution for both the null and the alternative hypothesis are produced using the EOS software in version 1.0.11 [115] (see Section 2.4.3).

To simulate the detector resolution, the q^2 value of each sample is shifted by a value drawn from a normal distribution of width 1 GeV^2 . This roughly corresponds to the Belle II detector resolution.

Furthermore, an efficiency map is applied to the samples according to the function

$$\varepsilon(q^2) = 0.4 \exp(-5 q^2/M_B^2), \quad (7.12)$$

which mimics the efficiency obtained in Reference [2].

The reconstruction variable is chosen to be the reconstructed momentum transfer q_{rec}^2 , obtained from the kinematic variable q^2 by applying detector and efficiency corrections. The binnings of the kinematic and reconstruction variables differ:

- For the reconstruction variable, a balance needs to be struck between the number of events in each bin (to ensure sufficient statistical power) and the number of bins (to ensure sensitivity to differences in the shape of the q^2 distribution). Eight equally-spaced bins are chosen for the reconstruction variable.
- For the kinematic variable, the number of bins is determined as follows. The convergence of the expected yields from the reweighted model is studied as the number of kinematic bins increases. For this study, the detector resolution smearing is removed. This is done for 100 randomly chosen theoretical models (see [Appendix C.2](#) for further details). These models correspond to normally-distributed variations $\sim \mathcal{N}(0, 10)$ of the **WET** parameters with respect to the **SM** parameter point. A width of $\sigma = 10$ is chosen to ensure that the models are sufficiently different from the **SM** point while still being within a reasonable range of the **WET** parameters. The aim is to ensure convergence at the level of 1% accuracy. It is found that using 24 equally-spaced bins for the kinematic variable ensures this aim. [Figure 7.1](#) illustrates this procedure for the benchmark point in [Equation \(7.10\)](#), where the changes for 5, 15 and 25 bins in the kinematic variable are shown.

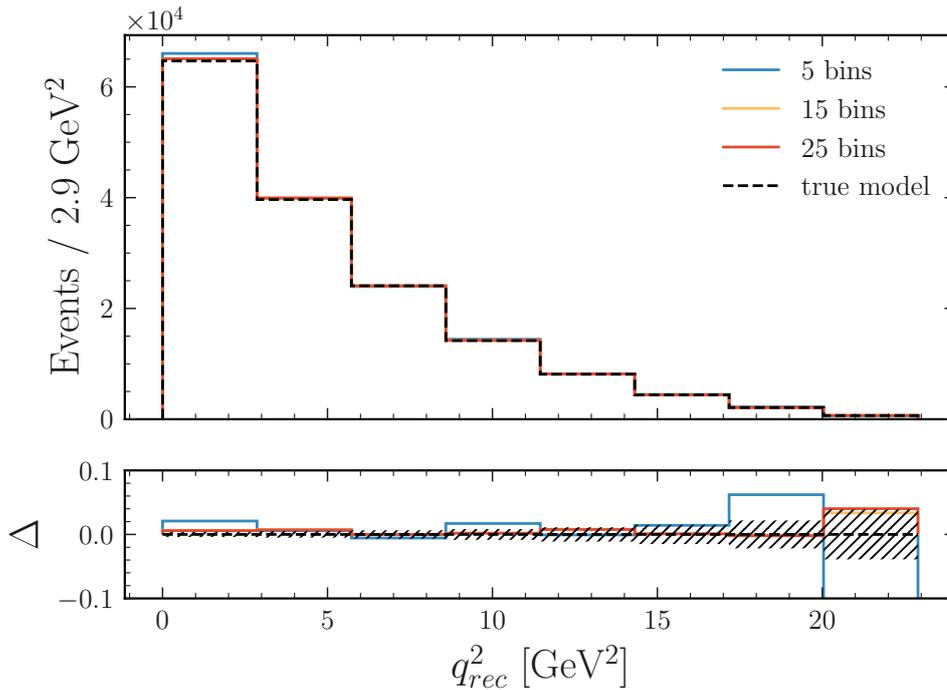


Figure 7.1: The null histogram yields, reweighted to the benchmark point in [Equation \(7.10\)](#), as a function of kinematic bins (red). The histogram yields of the true model, sampled from the probability density function at the benchmark point (black). The bottom plot shows the relative difference Δ of the reweighted models to the true model. The statistical uncertainty of the true model is shown as the hatched region.

Both datasets, according to the null (**SM**) and alternative (**BSM**) hypotheses, and their corresponding changes after detector resolution smearing and efficiency correction, are

shown in Figure 7.2.

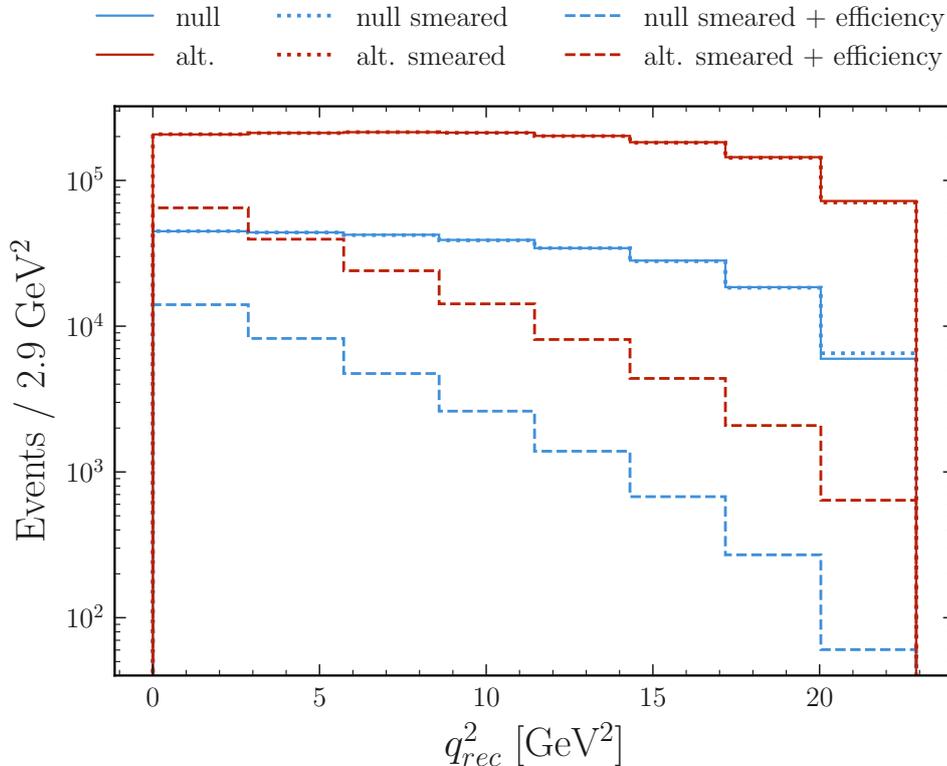


Figure 7.2: Both the null/**SM** (blue) and alternative/**BSM** (red) $B \rightarrow K\nu\bar{\nu}$ datasets, according to the pure theoretical prediction, after detector resolution smearing and efficiency correction.

The null (**SM**) and alternative (**BSM**) predictions have also been calculated using the EOS software [47, 115] and are shown in Figure 7.3. The null joint number density is obtained by binning the **MC** data in a 2-dimensional histogram of the reconstruction variable q_{rec}^2 against the kinematic variable q^2 , as shown in Figure 7.4.

7.3.1.3 Full statistical model

To build the posterior according to the statistical model described in Section 7.2, all parameters of the likelihood and their corresponding priors are collected.

The theoretical parameters include the **WET** parameters and the hadronic parameters. The **WET** parameters correspond to the three independent linear combinations of Wilson coefficients that enter the theoretical description of $B \rightarrow K\nu\bar{\nu}$ decays; see Equation (2.62). These are $|C_{VL} + C_{VR}|$, $|C_{SL} + C_{SR}|$, and $|C_{TL}|$. While the Wilson coefficients are — in general — complex-valued parameters, the overall phase of the **WET** Lagrangian Equation (2.52) is not observable. Moreover, an inspection of the differential decay rate in Equation (2.62) shows only sensitivity to the absolute values of the three linear combina-

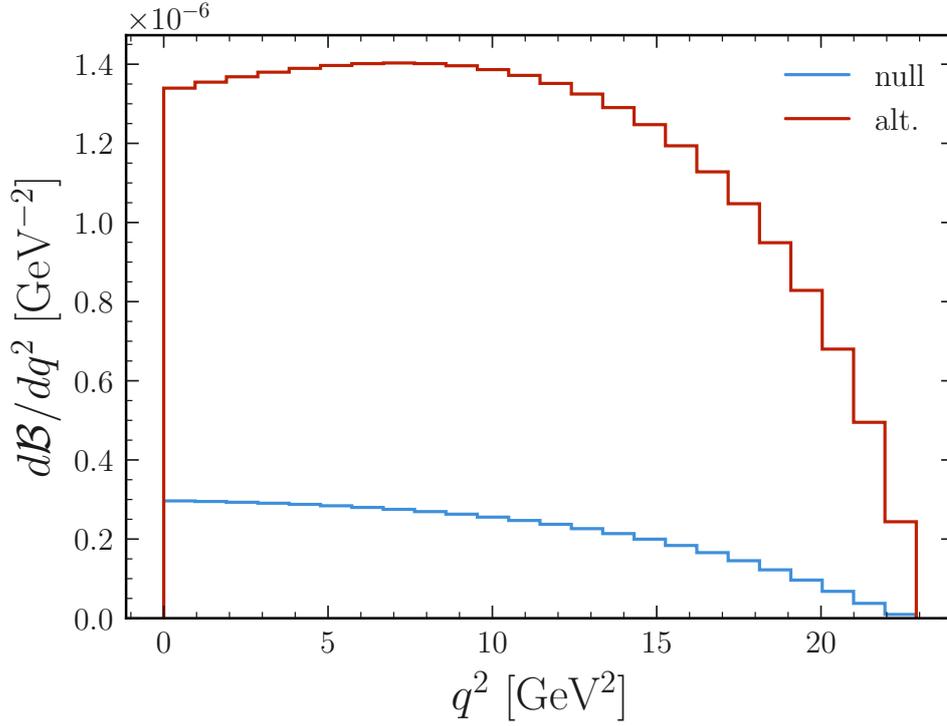


Figure 7.3: The bin-integrated null/SM (blue) and alternative/BSM (red) predictions for the differential branching ratio $d\mathcal{B}(B \rightarrow K\nu\bar{\nu})/dq^2$.

tions of Wilson coefficients. Hence, each linear combination is represented as a positive real-valued number. Their prior is chosen as the uncorrelated product of uniform distributions with support

$$\begin{aligned}
 5 &< |C_{VL} + C_{VR}| < 20, \\
 0 &< |C_{SL} + C_{SR}| < 15, \\
 0 &< |C_{TL}| < 15.
 \end{aligned} \tag{7.13}$$

These ranges are motivated with the benchmark point, Equation (7.11), in mind and wide enough to ensure that the posterior can explore the parameter space sufficiently. The prior of the hadronic parameters, discussed in Section 7.3.1.1, is a multivariate normal distribution, which is implemented as a sequence of independent univariate normal distributions, as discussed at the end of Section 7.2.

The experimental constraint includes one parameter per bin of the reconstruction variable, representing the statistical uncertainty of the MC yields. The priors for these parameters are normal distributions $\mathcal{N}(1, 1/\sqrt{N_b})$, where N_b is the total yield in reconstruction bin b . For the purpose of this proof-of-concept study, further (systematic) sources of uncertainty are not accounted for.

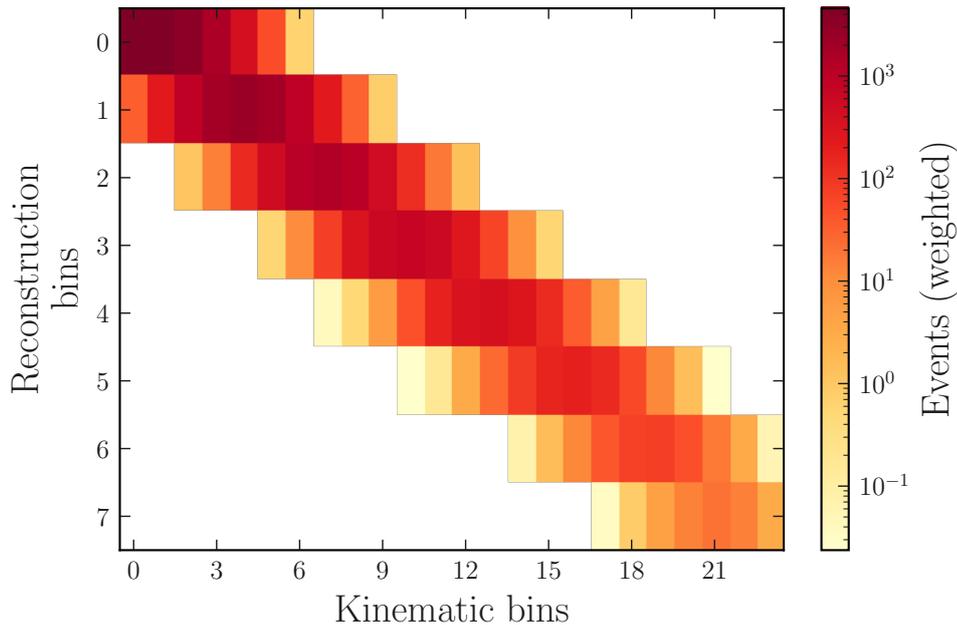


Figure 7.4: The null joint number density, showing the 8 bins of the reconstruction variable on the vertical axis and the 24 bins of the kinematic variable on the horizontal axis.

7.3.1.4 Reinterpretation results

Having built a model-agnostic likelihood function from the toy data, I investigate the potential of the approach to constrain the Wilson coefficients. Using **MCMC** sampling, the 3-dimensional marginal posterior distribution of the Wilson coefficients is obtained. The posterior is sampled using 8 independent chains², each with 100 000 samples. The values at the mode of the full posterior agree with those of the benchmark point outlined in **Equation (7.10)**. The full set of 2-dimensional marginalizations of this posterior and the resulting intervals at 68% and 95% probability are shown in **Figure 7.5**. The posterior mode and the 95% **HDI** (see **Section 5.4.1**) for the magnitudes of the Wilson coefficients are shown in **Table 7.2**.

The marginal posterior is found to peak at the benchmark point given in **Equation (7.11)**. The posterior mode agrees well with the benchmark point, especially for the 50 ab^{-1} sample, where the statistical precision is higher. The 95% **HDI** contains the benchmark point for all three parameters in both datasets. The deviations from the benchmark point observed in the 362 fb^{-1} sample are due to statistical fluctuations, arising from a combination of variance and potential bias. These deviations are substantially reduced in the 50 ab^{-1} sample, where the likelihood provides much stronger constraints on the parameters. An anti-correlation between the scalar and tensorial Wilson coefficients is visible in their marginalized 2-dimensional distribution. This behaviour is expected, as both the tensor and scalar terms in **Equation (2.62)** contribute most significantly at larger

²Chains are initialized randomly in the parameter space and run independently. This allows for better exploration of the parameter space and provides a way to assess convergence of the sampling [68].

Table 7.2: The mode of the posterior, and HDI at 95% for the (sums of the) WET Wilson coefficients in Equation (2.62), derived from the posterior in Figure 7.5.

| Parameters | 362 fb ⁻¹ | | 50 ab ⁻¹ | |
|-----------------------------------|----------------------|--------------|---------------------|--------------|
| | Mode | 95% HDI | Mode | 95% HDI |
| C _{VL} + C _{VR} | 13.9 | [12.0, 15.1] | 14.0 | [13.3, 14.7] |
| C _{SL} + C _{SR} | 4.2 | [0.0, 5.5] | 4.0 | [3.3, 4.7] |
| C _{TL} | 0.5 | [0.0, 4.6] | 1.0 | [0.0, 2.2] |

values of q^2 , where the efficiency (Equation (7.12)) is low. Moreover, this anti-correlation weakens as the statistical power of the data increases, demonstrating that it is driven primarily by limited statistics rather than fundamental parameter degeneracies.

Overall, the observed agreement with the benchmark Wilson coefficients provides a successful closure test for the reinterpretation method.

7.3.2 Combination of $B \rightarrow K\nu\bar{\nu}$ and $B \rightarrow K^*\nu\bar{\nu}$

A limitation of studying solely the $B \rightarrow K\nu\bar{\nu}$ process is that its sensitivity to the WET Wilson coefficients is limited to the three linear combinations shown in Equation (2.64). In this example, I showcase the power of combining data on $B \rightarrow K\nu\bar{\nu}$ and $B \rightarrow K^*\nu\bar{\nu}$ decays. These decays exhibit *complementary sensitivity* to the Wilson coefficients due to their different hadronic spin and orbital angular momentum configurations. This example is inspired by the ongoing Belle II analysis of the $B \rightarrow K^*\nu\bar{\nu}$ decay discussed in Chapter 11. For the sake of simplicity in this example, effects of additional kinematic variables in the decay chain $B \rightarrow K^*(\rightarrow K\pi)\nu\bar{\nu}$, such as the helicity angle θ_K of the kaon and the $K\pi$ invariant mass, are neglected. For the application of the proposed method to real-world examples, all kinematic variables should be included in the joint number density for full sensitivity to variations in these variables.

Moreover, this example demonstrates from a technical perspective that the method and its implementation also work for combined pyhf models, providing full access to the complementarity in sensitivity.

7.3.2.1 $B \rightarrow K^*\nu\bar{\nu}$ Weak Effective Theory parametrization

The decay $B \rightarrow K^*\nu\bar{\nu}$ is governed by the same WET Lagrangian as described by Equations (2.52) and (2.54). Its differential decay rate in terms of the Wilson coefficients of the operators is given in Equation (2.65), discussed in Section 2.4.2. Compared to $B \rightarrow K\nu\bar{\nu}$ decays, the differential $B \rightarrow K^*\nu\bar{\nu}$ decay rate exhibits additional sensitivity to the quantities $|C_{VL} - C_{VR}|$, $|C_{SL} - C_{SR}|$ (see Equation (2.67)). As a consequence, a simultaneous analysis of both decays allows constraining the magnitudes of all Wilson

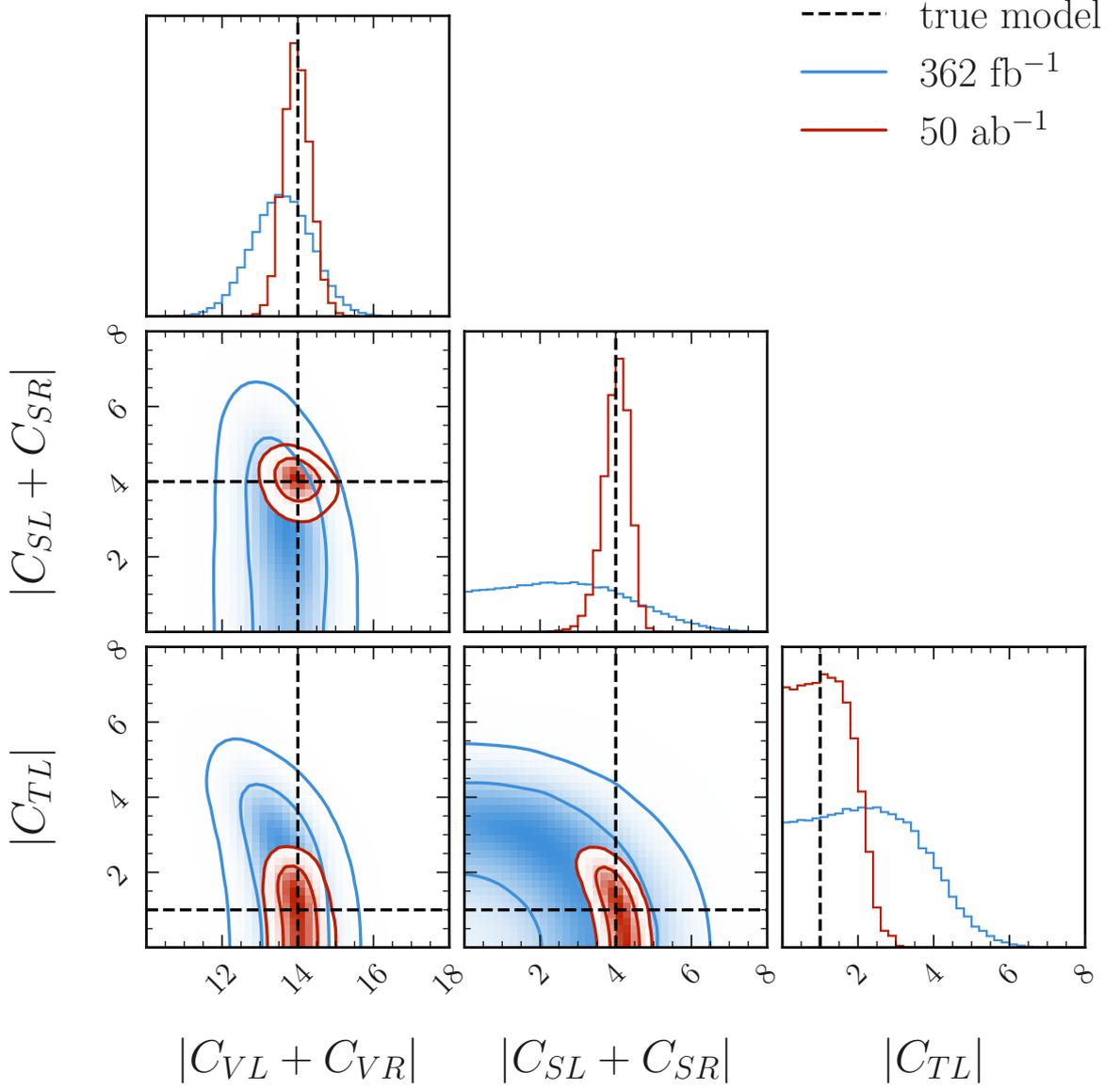


Figure 7.5: The marginalized posterior distributions, obtained by **MCMC** sampling from the $B \rightarrow K\nu\bar{\nu}$ likelihood. On the diagonal, the 1-dimensional marginal distributions of the Wilson coefficients in [Equation \(2.64\)](#) are shown. The contours on the 2-dimensional plots correspond to 68% (inner) and 95% (outer) probability. The dashed lines indicate the true underlying model ([Equation \(7.11\)](#)).

Table 7.3: The number of $B \rightarrow K^* \nu \bar{\nu}$ samples produced for this study, corresponding to an equivalent of 50 ab^{-1} integrated luminosity at the SuperKEKB collider. *Generated* and *reconstructed* samples correspond to the numbers prior and post efficiency correction.

| Luminosity | $B\bar{B}$ events | MC generated \ \ reconstructed | Data generated \ \ reconstructed |
|-----------------------|---------------------------|--|--|
| 362 fb^{-1} | $\sim 3.87 \cdot 10^8$ | $3.61 \cdot 10^3$ $8.46 \cdot 10^2$ | $1.05 \cdot 10^4$ $2.49 \cdot 10^3$ |
| 50 ab^{-1} | $\sim 5.35 \cdot 10^{10}$ | $4.99 \cdot 10^5$ $1.16 \cdot 10^5$ | $1.45 \cdot 10^6$ $3.43 \cdot 10^5$ |

coefficients, assuming all Wilson coefficients are real-valued. For an illustrative example, this assumption is applied here. The kinematic shape of the differential branching ratio, due to the vector, scalar and tensor interactions, is illustrated in [Figure 2.3](#).

The hadronic matrix elements of the [WET](#) operators in this decay are expressed in terms of seven independent form factors $V(q^2)$, $A_0(q^2)$, $A_1(q^2)$, $A_{12}(q^2)$, $T_1(q^2)$, $T_2(q^2)$ and $T_{23}(q^2)$, as discussed in [Section 2.3](#) and [Appendix A.3](#). In this work, these form factors are parametrized following the BSZ parametrization [[41](#)], which is truncated at order $K = 2$. The values for the corresponding 19 hadronic parameters arise from the Gaussian likelihood provided in Reference [[49](#)]; see [Table A.1](#). Correlations between the hadronic parameters are taken into account through their covariance matrix and implemented as discussed in [Section 7.2](#).

For the [SM](#) prediction, only the vector left-handed Wilson coefficient C_{VL} is non-zero (see [Equation \(2.55\)](#)), leading to a predicted branching ratio of [[37](#), [47](#)]

$$\mathcal{B}(B^+ \rightarrow K^{*+} \nu \bar{\nu})_{\text{SM}} = (9.3 \pm 0.9) \cdot 10^{-6}. \quad (7.14)$$

The theoretical uncertainty is dominated by the form factor uncertainties from lattice [QCD](#), with smaller contributions from [CKM](#) matrix elements and higher-order [QCD](#) corrections.

7.3.2.2 Datasets

To produce the $B \rightarrow K^* \nu \bar{\nu}$ datasets, the same procedure as in [Section 7.3.1.2](#) is adapted.

The [MC](#) data are produced according to the [SM](#) prediction (*null* hypothesis). The number of samples is calculated by multiplying the estimated number of $B\bar{B}$ events in a 50 ab^{-1} Belle II dataset with the predicted [SM](#) branching ratio of [Equation \(7.14\)](#).

The real data are produced according to the [BSM](#) prediction at the benchmark point in [Equation \(7.10\)](#) (*alternative* hypothesis). The number of data samples is calculated by multiplying the estimated number of $B\bar{B}$ events in a 50 ab^{-1} Belle II dataset by the predicted [BSM](#) branching ratio, $\mathcal{B}(B \rightarrow K^* \nu \bar{\nu}) \approx 2.72 \cdot 10^{-5}$ [[47](#), [115](#)].

The number of samples is listed in [Table 7.3](#). MC samples of the decay's probability distribution for both the null and the alternative hypothesis are produced using the EOS software in version 1.0.11 [[115](#)].

The efficiency map in this case is chosen to be

$$\varepsilon(q^2) = 0.3 \left(1 - 0.08 \exp \left(-2.5 q^2 / M_B^2 \right) \right), \quad (7.15)$$

which is an approximate expectation for an inclusive $B \rightarrow K^* \nu \bar{\nu}$ analysis. For the sake of simplicity, it is assumed that the efficiency is independent of the helicity angle θ_K and the $K\pi$ invariant mass.

Ten bins in the reconstruction variable (q_{rec}^2) are chosen, and it is found that 25 bins in the kinematic variable (q^2) provide sufficient accuracy using the procedure described in [Appendix C.2](#) and [Section 7.3.1.2](#).

Both datasets, according to the null ([SM](#)) and alternative ([BSM](#)) hypotheses, and their corresponding changes after detector resolution smearing and efficiency correction, are shown in [Figure 7.6](#).

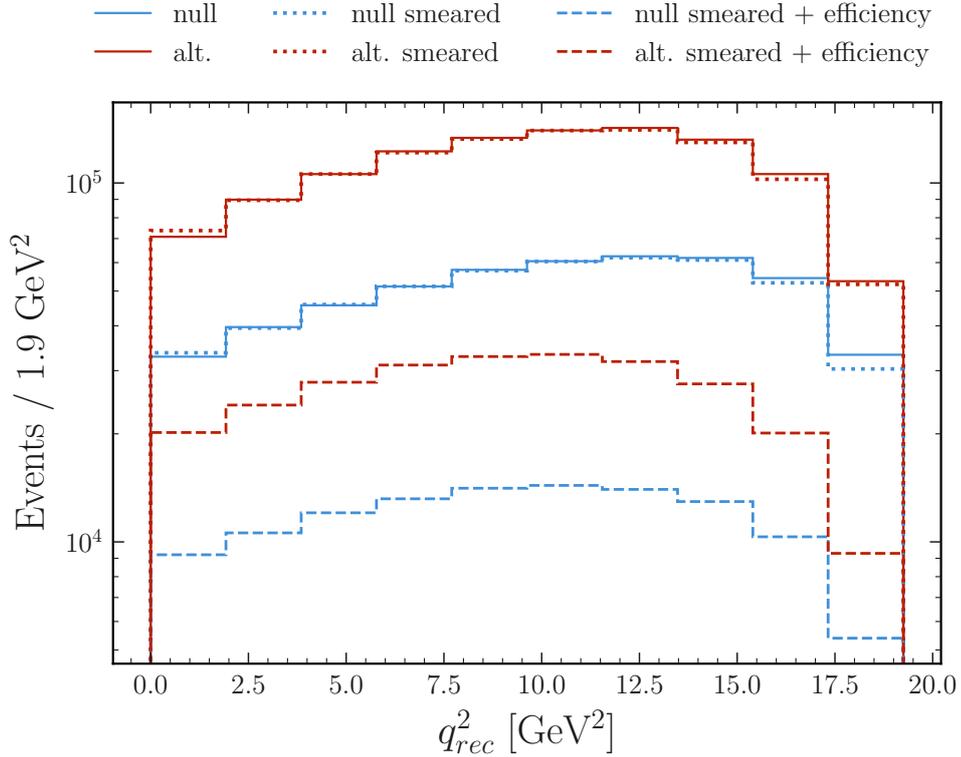


Figure 7.6: Both the null/[SM](#) (blue lines) and alternative/[BSM](#) (red lines) $B \rightarrow K^* \nu \bar{\nu}$ datasets, according to the pure theoretical prediction, after detector resolution smearing and efficiency correction.

7.3.2.3 Full statistical model

To derive the statistical model encompassing $B \rightarrow K\nu\bar{\nu}$ and $B \rightarrow K^*\nu\bar{\nu}$, individual posteriors for each channel are constructed following the methodology outlined in [Section 7.3.1.3](#). The combined posterior arises from their product, with the [WET](#) theory parameters being the only parameters shared by the individual posteriors.

These parameters correspond to the five magnitudes of Wilson coefficients that enter the theoretical description of $B \rightarrow K\nu\bar{\nu}$ and $B \rightarrow K^*\nu\bar{\nu}$ decays (see [Equations \(2.62\)](#) and [\(2.65\)](#)). They are C_{VL} , C_{VR} , C_{SL} , C_{SR} , and C_{TL} . Their prior is chosen as the uncorrelated product of uniform priors with support

$$\begin{aligned} 5 &< |C_{VL}| < 15, \\ 0 &< |C_{VR}| < 10, \\ 0 &< |C_{SL}| < 10, \\ 0 &< |C_{SR}| < 10, \\ 0 &< |C_{TL}| < 10. \end{aligned} \tag{7.16}$$

These ranges are motivated with the benchmark point, [Equation \(7.10\)](#), in mind and wide enough to ensure that the posterior can explore the parameter space sufficiently.

The hadronic parameters, describing the $B \rightarrow K$ and $B \rightarrow K^*$ form factors are discussed in [Sections 7.3.1.1](#) and [7.3.2.1](#). Their prior is a multivariate normal distribution, which is implemented as a sequence of independent univariate normal distributions, as discussed at the end of [Section 7.2](#).

In the context of `HistFactory` models, the combined likelihood of $B \rightarrow K\nu\bar{\nu}$ and $B \rightarrow K^*\nu\bar{\nu}$ is a combination at the *channel* level, as discussed in [Section 7.2](#). One custom modifier is added to each channel, which are functions of the same [WET](#) parameters but different hadronic parameters.

7.3.2.4 Reinterpretation results

From the constructed model-agnostic likelihood function, I investigate the power of constraining the full set of Wilson coefficients appearing in [Equations \(2.62\)](#) and [\(2.65\)](#), under the assumption that they are real-valued. The decay rates in [Equations \(2.62\)](#) and [\(2.65\)](#) exhibit two discrete symmetries: one under the exchange $C_{VL} \leftrightarrow C_{VR}$ and another under the exchange $C_{SL} \leftrightarrow C_{SR}$. The combination of both symmetries leads to a four-fold ambiguity in the extraction of the Wilson coefficients from data and therefore a multimodal posterior density. To avoid computational issues when sampling from the posterior, one of the four fully-equivalent modes is selected for sampling. This is achieved by imposing the additional constraints $C_{VL} > C_{VR}$ and $C_{SL} > C_{SR}$. [MCMC](#) sampling is used, and the chains are initialized with the mode of the full posterior. The posterior is sampled using 8 independent chains, each with 100 000 samples. The values at the mode of the full posterior align with those of the benchmark point outlined in [Equation \(7.10\)](#). To obtain the full multimodal posterior, the original symmetry is restored manually. From

the symmetrized samples, the 5-dimensional marginal posterior distribution of the Wilson coefficients is obtained. This is illustrated in [Figure 7.7](#) by showing the full set of 1- and 2-dimensional marginalizations and the resulting regions at 68% and 95% probability.³ The posterior mode and the 95% HDI (see [Section 5.4.1](#)) for the magnitudes of the Wilson coefficients are shown in [Table 7.4](#).

Table 7.4: The mode of the posterior, and HDI at 95% for the (sums of the) WET Wilson coefficients in [Equations \(2.62\)](#) and [\(2.65\)](#), derived from the posterior in [Figure 7.7](#).

| Parameters | 362 fb ⁻¹ | | 50 ab ⁻¹ | |
|------------|----------------------|-------------|---------------------|-------------|
| | Mode | 95% HDI | Mode | 95% HDI |
| $ C_{VL} $ | 10.0 | [6.9, 13.2] | 10.0 | [8.4, 12.7] |
| $ C_{VR} $ | 4.0 | [0.6, 6.8] | 4.0 | [1.3, 5.6] |
| $ C_{SL} $ | 3.2 | [0.8, 5.5] | 3.0 | [2.1, 4.2] |
| $ C_{SR} $ | 0.9 | [0.0, 2.2] | 1.0 | [0.0, 1.8] |
| $ C_{TL} $ | 1.0 | [0.2, 1.5] | 1.0 | [0.4, 1.3] |

Combining results for $B \rightarrow K\nu\bar{\nu}$ and $B \rightarrow K^*\nu\bar{\nu}$ enables probing of all five Wilson coefficient magnitudes. The posterior modes align well with the benchmark point in [Equation \(7.10\)](#) for both datasets, with the 50 ab⁻¹ dataset showing improved precision. The 95% HDIs contain the benchmark point in all cases, with substantial narrowing at higher luminosity.

The combination significantly improves constraints on $|C_{TL}|$ compared to $B \rightarrow K\nu\bar{\nu}$ alone (see [Figure 7.5](#) and [Table 7.2](#)), while the scalar sector shows the most dramatic precision gain between datasets. For the smaller 362 fb⁻¹ dataset, overlapping peaks and magnitude sampling lead to asymmetric scalar distributions with broader tails. The tail of $|C_{TL}|$ toward lower values persists, consistent with the previous example in [Figure 7.5](#).

This study provides a successful closure test for the reinterpretation method and demonstrates the power of combining measurements within this framework.

The combined analysis demonstrates the enhanced constraining power achievable through the reinterpretation method when applied to multiple complementary measurements. However, to fully appreciate the value of this approach, it is essential to understand the limitations of conventional reinterpretation strategies and the biases they introduce.

³The binning of the posterior samples is not aligned with the symmetry axis, leading to small visual asymmetries.

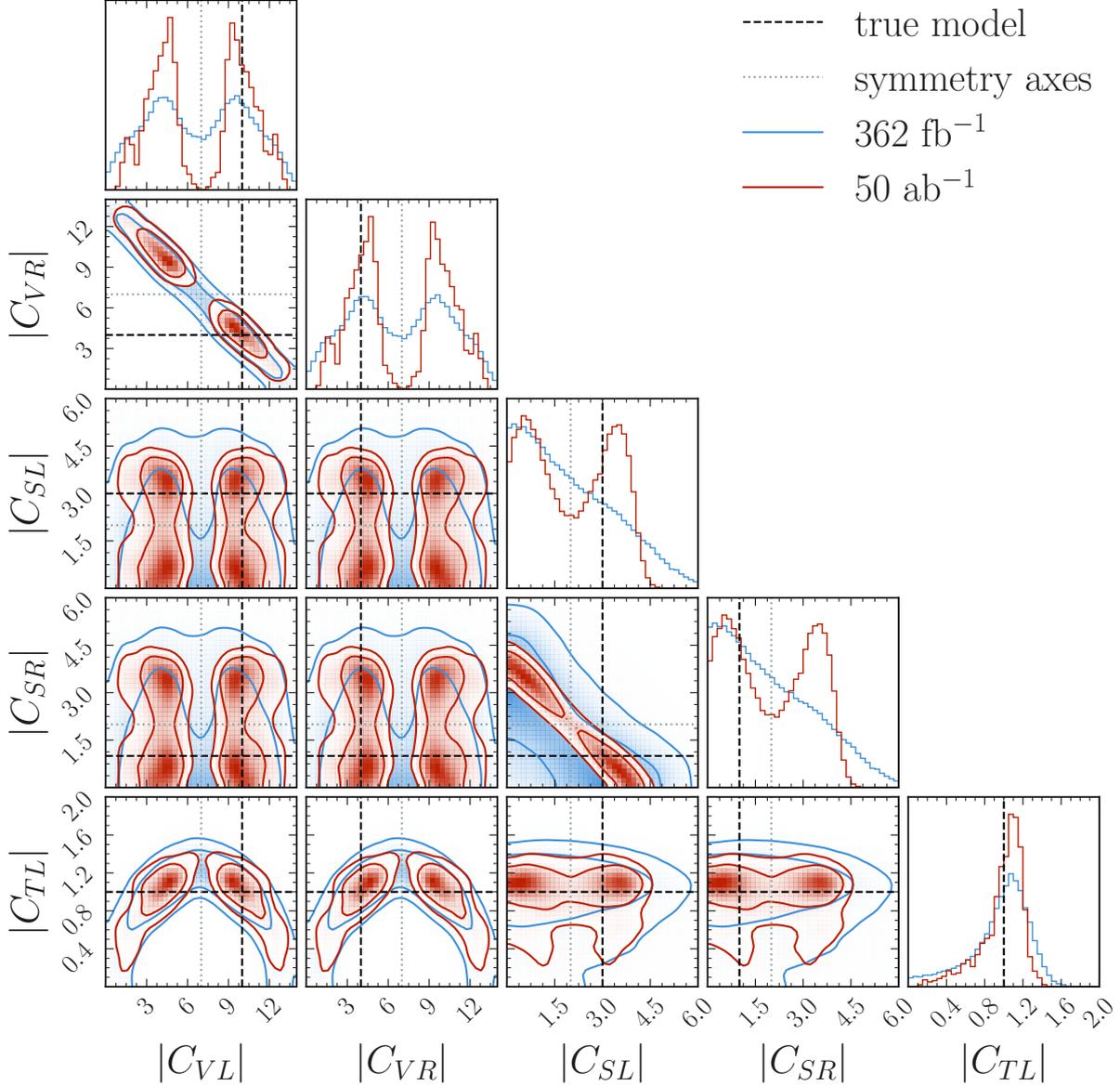


Figure 7.7: The marginal posterior distributions, obtained by **MCMC** sampling from the combined $B \rightarrow K\nu\bar{\nu}$ and $B \rightarrow K^*\nu\bar{\nu}$ likelihood. On the diagonal, the 1-dimensional marginal distributions of the Wilson coefficients appearing in [Equations \(2.62\) and \(2.65\)](#) are shown. The contours on the 2-dimensional plots correspond to 68% (inner) and 95% (outer) probability. The dashed lines indicate the true underlying model. The dotted lines indicate the symmetry axes of the global likelihood.

7.3.3 The necessity for reinterpretation

The availability of open datasets for most particle physics results is currently very limited, although improving thanks to the popularity of statistical approaches such as `HistFactory` [70] and tools such as `pyhf` [71, 72]. These current limitations regularly hinder theorists from fully interpreting existing experimental results in their **BSM** analyses. In particular, **BSM** changes to the distribution of the reconstruction variable are routinely neglected. In fact, the most common approach in a **BSM** analysis is to constrain the ratio of the **BSM** prediction to the **SM** prediction from branching ratio measurements or upper limits. This approach is valid only if **BSM** changes to the shape of the kinematic variable distribution can be accounted for by systematic experimental uncertainties in the reconstruction space. As shown in the following, this does not hold for measurements of the $B \rightarrow K\nu\bar{\nu}$ branching ratio.

To illustrate the issue, the results in [Section 7.3.1](#) based on simulated data are compared with those obtained from naive rescaling of the branching ratio. In the language of the presented reinterpretation method, the latter corresponds to using only a single bin in the kinematic **d.o.f.**, covering the full kinematic range. This translates to a single weight applied to all bins of the reconstruction variable, corresponding to the ratio of the alternative to the null prediction integrated over the full kinematic range. Therefore, an additional “naive” $B \rightarrow K\nu\bar{\nu}$ posterior is constructed that deviates from the setup in [Section 7.3.1](#) only by using a single bin in the kinematic range.

After sampling from this “naive” posterior, the marginal distributions for the Wilson coefficients are compared to those presented in [Figure 7.5](#). This comparison is shown in [Figure 7.8](#). A striking difference in the overall shape of the distributions and the central intervals at 68% (95%) probability is found. Clearly, the “naive” procedure fails to validate, yielding large deviations from the benchmark point in [Equation \(7.11\)](#) in all three sectors. The results illustrate that the approach is essential for faithful reinterpretation of $B \rightarrow K\nu\bar{\nu}$ experimental results.

It should be emphasized that the method provides a means to ensure *accurate* interpretation of the existing likelihood beyond the assumptions of the underlying signal model. This does not imply, however, that this interpretation is more *precise* than a naive **BSM** interpretation. Put differently, the approach eliminates bias introduced by using an incorrect template for the decay’s kinematic distribution but at the expense of potentially larger uncertainties on the theory parameters.

7.4 Discussion and significance of the method

I present a novel reinterpretation method for particle physics results that is simple in its application and requires only minimal information in addition to published likelihoods.

The proposed method avoids biases that are introduced in naive reinterpretation of data with a negligible increase in compute time. As such, it provides most of the benefits of reinterpretation using full analysis preservation. Therefore, this method provides a good

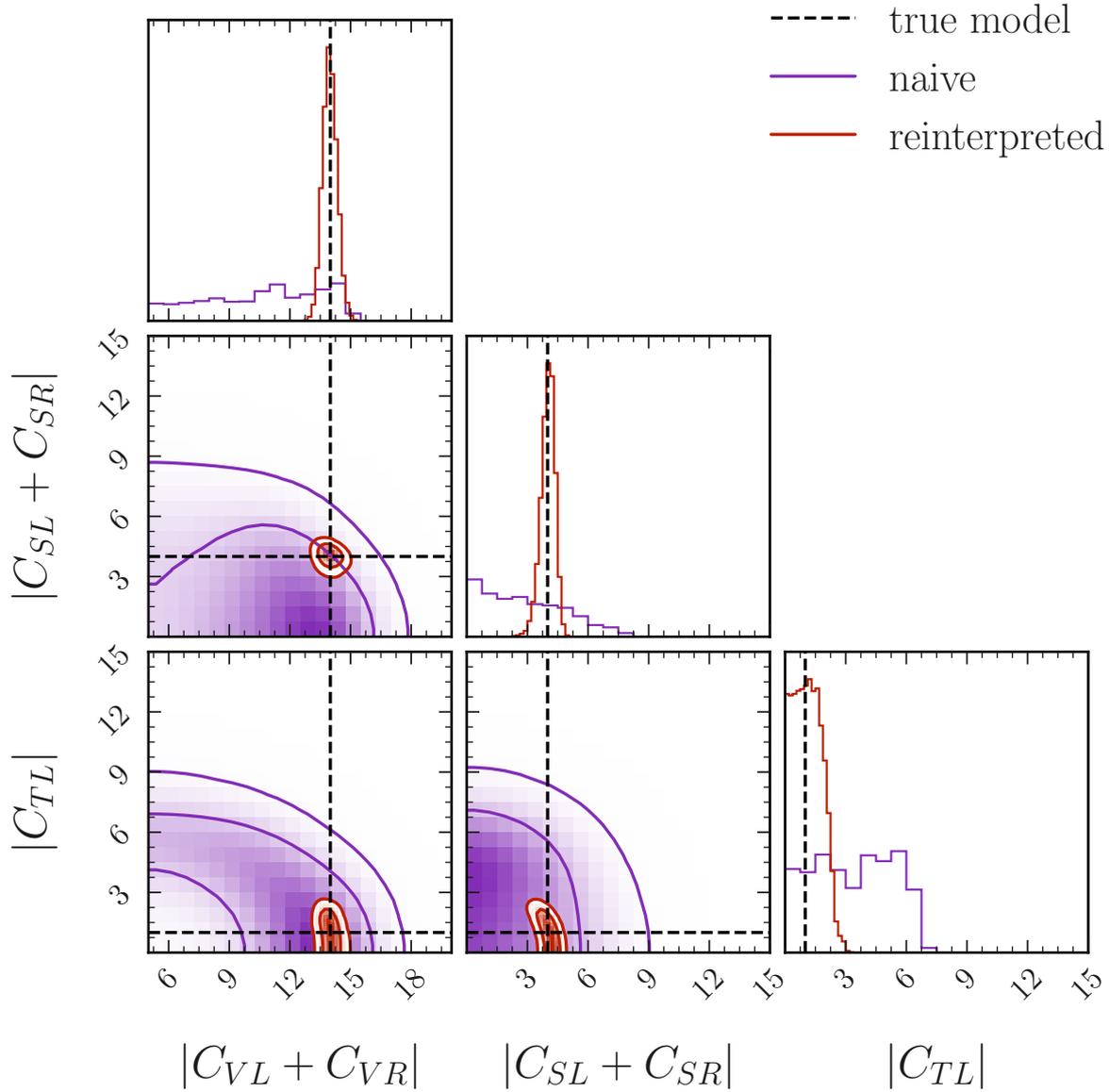


Figure 7.8: The comparison of the posterior distribution resulting from a model with only one bin in the kinematic **d.o.f.** to the proposed reinterpretation method, respecting shape changes in the kinematic distribution (see [Figure 7.5](#)).

trade-off between accuracy and speed, and has the potential to improve the accuracy of global effective field theory fits to many analysis results.

To showcase the method, it is applied to a simulated dataset of the $B \rightarrow K\nu\bar{\nu}$ decay, inspired by the recent Belle II analysis [2] but without resorting to using any public or private Belle II data.

Using the two examples discussed in Section 7.3, I validate the method by successfully recovering the benchmark theory point from the underlying synthetic data. This outcome underscores the accuracy and self-consistency of the approach. The bias introduced by naive rescaling of the $B \rightarrow K\nu\bar{\nu}$ branching ratio is further investigated. For the benchmark point, a sizable bias is found when determining the WET Wilson coefficients without applying the reinterpretation method.

In conclusion, the work presented in this chapter illustrates both the ease of applicability and the urgent necessity of shape-respecting reinterpretation over the traditional approach. It is hoped that this work will motivate experimental collaborations and analysts to consider the future reinterpretability of their results, and to publish the necessary material (for further details, see Appendix C.3). This will, in turn, enable the entire community to use these analysis results accurately.

Chapter 8

A model-agnostic likelihood for the Belle II $B^+ \rightarrow K^+ \nu \bar{\nu}$ measurement

¹ Having established the theoretical framework for model-agnostic reinterpretation in the previous chapter, I now proceed to its first practical application to real experimental data. This and the following chapters present the first-ever implementation of the reinterpretation method presented in [Chapter 7](#) to actual Belle II data, specifically applied to the $B^+ \rightarrow K^+ \nu \bar{\nu}$ analysis² [2].

The Belle II collaboration reported a 2.7σ excess over the [SM](#) expectation in the $B^+ \rightarrow K^+ \nu \bar{\nu}$ decay channel [2], which could indicate contributions from [BSM](#) physics. The original analysis quantified this excess through a signal strength parameter μ relative to the [SM](#) prediction. A reinterpretation within the [WET](#) framework enables a more detailed characterization of the potential new physics by resolving contributions from different operator structures.

This work not only demonstrates the real-world feasibility of the approach but also establishes important precedents for future applications of Belle II data. The choice of the Belle II $B^+ \rightarrow K^+ \nu \bar{\nu}$ analysis lies in its recent experimental prominence and the theoretical interest it has generated in the particle physics community.

The [SM](#) branching ratio of the $B^+ \rightarrow K^+ \nu \bar{\nu}$ decay used in the Belle II analysis [2] is predicted to be [36]

$$\mathcal{B}(B^+ \rightarrow K^+ \nu \bar{\nu}) = (5.58 \pm 0.37) \cdot 10^{-6}, \quad (8.1)$$

which includes a contribution of $(0.61 \pm 0.06) \cdot 10^{-6}$ from the long-distance double-charged-current $B^+ \rightarrow \tau^+ (\rightarrow K^+ \bar{\nu}_\tau) \nu_\tau$ decay. In this measurement, only the short-

¹The material presented in this chapter has been submitted to PRD [116]. While the content largely follows that of the referenced paper, editorial modifications have been made, including restructuring of the presentation, rephrasing of portions of the text, and expansion of selected explanations. As this project is the result of collaborative work of the Belle II collaboration, individual contributions of the proponents of this work, and further collaboration members are documented in [Appendix D.1](#).

²Charge-conjugate channels are implied.

distance contribution was considered as signal, with a corresponding branching ratio of $(4.97 \pm 0.37) \cdot 10^{-6}$ [36]. These can be distinguished due to a different kinematic signature.

The $B^+ \rightarrow K^+\nu\bar{\nu}$ decay has been studied by the CLEO, *BABAR*, Belle, and Belle II collaborations [2, 111, 117–121], with the latest measurement by the Belle II collaboration finding the first evidence for this decay at 3.5 standard deviations. This result, based on the SM prediction from Ref. [36] and hadronic parameters from the HPQCD collaboration [46] as a model ansatz, exceeds the SM expectation by 2.7 standard deviations. This enhanced branching ratio triggered the HEP community to interpret this result under different new physics scenarios [112–114]. Such reinterpretations, however, are only approximate as not all relevant information on the measurement was previously accessible to the public.

Generally, two distinct classes of new physics models can replicate the signature of a $B^+ \rightarrow K^+\nu\bar{\nu}$ decay. These models naturally mimic the neutrino pair in the final state, as the neutrinos remain undetected at Belle II. New physics scenarios can proceed via either three-body or two-body decays. In three-body decays, the properties of new physics can be studied within the framework of the WET (see Section 2.2.1), where new physics particles and force carriers have masses at or above the electroweak symmetry-breaking scale. Examples of such models include leptoquarks [6–8, 10] and heavy Z' bosons [9–12]. Two-body decays would signal the presence of light new physics, which lies below the scale of electroweak symmetry breaking and is not described by the WET. Explicit examples include axions [10, 13–15], axion-like particles [10, 16–18] or other dark-sector mediators [19–22].

The primary goals of this work are threefold: to publish the model-agnostic likelihood for the $B^+ \rightarrow K^+\nu\bar{\nu}$ analysis, which consists of the full likelihood from the $B^+ \rightarrow K^+\nu\bar{\nu}$ result [2] and the joint number densities used in this reinterpretation [122, 123] (for details see Appendix E.1); to reinterpret the result in the framework of the WET, providing constraints on relevant Wilson coefficients; and to reinterpret the result in the context of a two-body light new physics model. The published model-agnostic likelihood allows the broader scientific community to test alternative theoretical models in a statistically rigorous way. Importantly, this work sets a template for future Belle II measurements that are suitable for reinterpretation and reflects the collaboration’s commitment to publishing model-agnostic likelihoods as a means to maximize the scientific impact and reusability of its results.

This chapter is structured as follows: In Section 8.1 I give an overview of the $B^+ \rightarrow K^+\nu\bar{\nu}$ analysis [2], including the data used and the analysis strategy. How the reinterpretation method of Chapter 7 is applied in the context of the $B^+ \rightarrow K^+\nu\bar{\nu}$ analysis is described in Section 8.2.

8.1 The Belle II $B^+ \rightarrow K^+ \nu \bar{\nu}$ analysis

In this section I provide an overview of the Belle II analysis of the $B^+ \rightarrow K^+ \nu \bar{\nu}$ decay [2], which serves as the basis for the reinterpretation presented in this work.

8.1.1 Data and simulated samples

The Belle II analysis of $B^+ \rightarrow K^+ \nu \bar{\nu}$ decays [2], originating from an $e^+e^- \rightarrow \Upsilon(4S) \rightarrow B^+B^-$ process, was conducted using an integrated luminosity of 362 fb^{-1} of on-resonance data [124], recorded between 2019 and 2022 at a centre-of-mass energy of $\sqrt{s} \approx 10.58 \text{ GeV}$, corresponding to the $\Upsilon(4S)$ resonance. This dataset contains $(387 \pm 6) \cdot 10^6$ B^+B^- pairs [125]. An off-resonance sample of 42 fb^{-1} was also recorded at $\sqrt{s} \approx 10.52 \text{ GeV}$, which is used to constrain the background contributions from continuum processes.

The analysis was designed using MC simulated data for the signal and seven background categories: decays of charged and neutral $B\bar{B}$ mesons; continuum processes ($u\bar{u}$, $d\bar{d}$, $s\bar{s}$, $c\bar{c}$); and low-multiplicity $\tau^+\tau^-$ backgrounds. These simulated data were used to estimate the efficiency of the selection criteria, train multivariate classifiers, and build the statistical model. Detector simulation and reconstruction was performed using the `basf2` framework [60] interfaced with `GEANT4` [64], as described in Section 3.4.

8.1.2 Background suppression

This measurement was performed using two different reconstruction methods: the more sensitive Inclusive Tagging Analysis (ITA) and the more conventional Hadronic Tagging Analysis (HTA), targeting nearly orthogonal datasets. In the HTA, the non-signal B meson of the B^+B^- pair was reconstructed in specific hadronic decay modes. Subsequently, the signal B meson is reconstructed from the remaining particles in the event. Requiring a full reconstruction of one B meson ensures a low background, at the cost of lower signal efficiency. In contrast, in the ITA, the signal was identified by reconstructing the K^+ track, summing all remaining particles in the event, and then inferring the missing energy. The relaxed requirement on the reconstruction of the non-signal B meson allows for a higher signal efficiency at the cost of increased background contamination.

In addition to a basic preselection, the ITA used two Boosted Decision Trees (BDTs), BDT_1 and BDT_2 , to separate the signal from the background.

A first binary classifier, BDT_1 , is designed as a first-level filter after event selection. The most powerful discriminant is the difference between the Rest-Of-Event (ROE) energy in the centre-of-mass frame and half the total collision energy, $\Delta E_{\text{ROE}} = E_{\text{ROE}} - \sqrt{s}/2$, which is typically negative for signal events due to neutrinos but positive for background events containing additional reconstructed particles [2].

The second classifier, BDT_2 , is used for the final event selection, which is trained on events with $\text{BDT}_1 > 0.9$. This corresponds to a signal (background) selection efficiency of

34% (1.5%), using 35 input variables. For BDT_2 , the most discriminating variable is the cosine of the angle between the momentum of the signal-kaon candidate and the thrust axis of the ROE energy in the centre-of-mass frame. This has a uniform distribution for the signal and a peaking shape for the jet-like continuum background [2].

The HTA used one BDT s, BDTh , to suppress the background. The BDTh input variable providing the highest discriminating power is E_{extra} , which is the sum of the energy deposited in the ECL that cannot be directly associated with the reconstructed B -mesons. It is primarily expected to be zero or small for correctly reconstructed signal events, but can be larger for background events [2].

8.1.3 Signal region definition

For both the ITA and HTA methods, a `HistFactory` statistical model (see Section 4.4 and Reference [70]) was used. This was implemented in the `pyhf` framework (see Section 4.4.3 and References [71, 72]), incorporating all systematic uncertainties. The ITA analysis used 4×3 bins in the $\eta(\text{BDT}_2) \times q_{\text{rec}}^2$ space, where $\eta(\text{BDT}_2)$ is the transformed BDT_2 output,

$$\eta(\text{BDT}_2) = 1 - \int_{\text{BDT}_2}^1 db \xi(b), \quad (8.2)$$

such that each bin in BDT_2 represents a 2% efficiency quantile for the selection. In this way, the distribution of $\eta(\text{BDT}_2)$ for simulated signal events is uniform. Here, $\xi(b)$ is the signal selection efficiency for the BDT_2 output b . q_{rec}^2 is the reconstructed q^2 , defined as

$$q_{\text{rec}}^2 = \frac{s}{4} + M_K^2 - \sqrt{s} E_K^*, \quad (8.3)$$

assuming the signal B meson to be at rest in the e^+e^- centre-of-mass frame. Here, M_K is the nominal kaon mass and E_K^* is the reconstructed energy of the kaon in the collision centre-of-mass frame. The signal regions for the ITA were defined as $\text{BDT}_1 > 0.9$ and $\eta(\text{BDT}_2) > 0.92$, which maximizes the signal significance as expected by simulation. The q_{rec}^2 bin boundaries are given by $[-1.0, 4.0, 8.0, 25.0]$ GeV^2 , and the bin boundaries in $\eta(\text{BDT}_2)$ are $[0.92, 0.94, 0.96, 0.98, 1.00]$. The bin with $\eta(\text{BDT}_2) > 0.98$ provides the largest signal sensitivity. The signal region for the HTA was defined as $\eta(\text{BDTh}) > 0.4$, where $\eta(\text{BDTh})$ is defined analogously to Equation (8.2) but using the BDTh output. The HTA used 6 equally spaced bins in the $\eta(\text{BDTh})$ space [2].

The ITA and HTA q^2 resolutions are approximately 1 GeV^2 and 0.3 GeV^2 , respectively, leading to differences in the q_{rec}^2 and q^2 distributions. Final-state radiation has a negligible effect on the presented results compared to other sources of uncertainties. This is discussed further in Appendix D.2. The ITA includes both a signal and a control region, resulting in a total of 24 reconstruction bins. Signal region bins are populated with events that pass all selection criteria in the on-resonance data. Control region bins are filled with events that pass the same selection in the off-resonance data. The control region improves constraints on the background contributions from continuum processes. The HTA signal

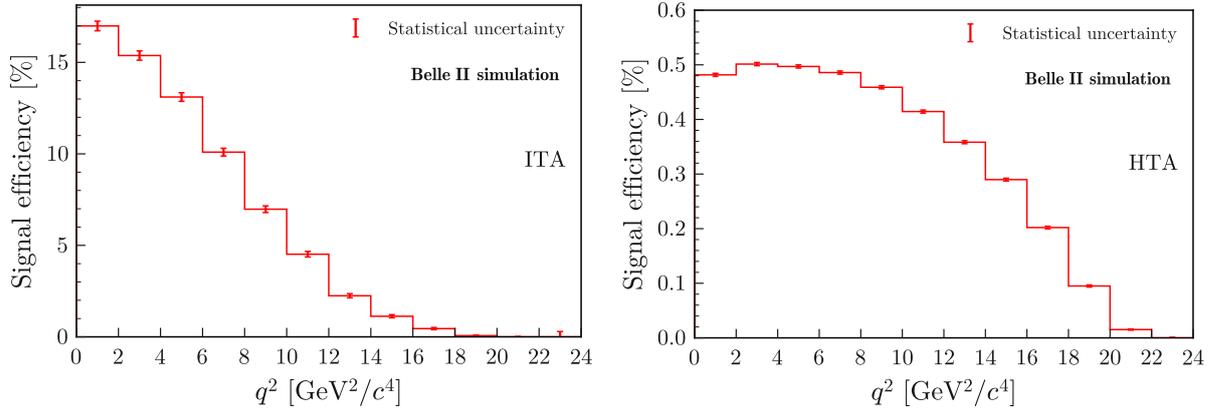


Figure 8.1: The signal-selection efficiencies for the **ITA** (*left*) and **HTA** (*right*) as a function of q^2 . Error bars indicate the statistical uncertainty. Figures taken from Reference [2].

region contains only events that pass all selection criteria in the on-resonance data. Signal-selection efficiencies for simulated events, as a function of q^2 are shown in [Figure 8.1](#) for the **ITA** and **HTA** methods. The **ITA** has a higher signal efficiency than the **HTA**, but the **HTA** has a lower background contamination. In both cases, the signal efficiency decreases with increasing q^2 .

8.1.4 Signal extraction result

The best-fit yields of the **ITA** analysis are shown in [Figure 8.2](#). A maximum likelihood fit yields a signal strength relative to the **SM** expectation of [2]

$$\mu_{\text{ITA}} = 5.4 \pm 1.0 \text{ (stat.)} \pm 1.1 \text{ (syst.)} = 5.4 \pm 1.5. \quad (8.4)$$

This translates into a branching ratio of [2]

$$\mathcal{B}(B^+ \rightarrow K^+ \nu \bar{\nu})_{\text{ITA}} = (2.7 \pm 0.5 \text{ (stat.)} \pm 0.5 \text{ (syst.)}) \cdot 10^{-5}. \quad (8.5)$$

A projection of the best-fit point onto the q_{rec}^2 variable is shown in [Figure 8.3](#). This is done by computing a ratio of post- to pre-fit yields in the reconstruction binning for each of the samples and applying these weights to all events in the respective bins. Interestingly, one can see an excess in the data in the region of $3 < q_{\text{rec}}^2 < 7 \text{ GeV}^2$. I will discuss this further in [Chapter 10](#).

The best-fit yields of the **HTA** analysis are shown in [Figure 8.4](#). A maximum likelihood fit yields a signal strength relative to the **SM** expectation of [2]

$$\mu_{\text{HTA}} = 2.2_{-1.7}^{+1.8} \text{ (stat.)}_{-1.1}^{+1.6} \text{ (syst.)} = 2.2_{-2.0}^{+2.4}. \quad (8.6)$$

This translates into a branching ratio of [2]

$$\mathcal{B}(B^+ \rightarrow K^+ \nu \bar{\nu})_{\text{HTA}} = (1.1_{-0.8}^{+0.9} \text{ (stat.)}_{-0.5}^{+0.8} \text{ (syst.)}) \cdot 10^{-5}. \quad (8.7)$$

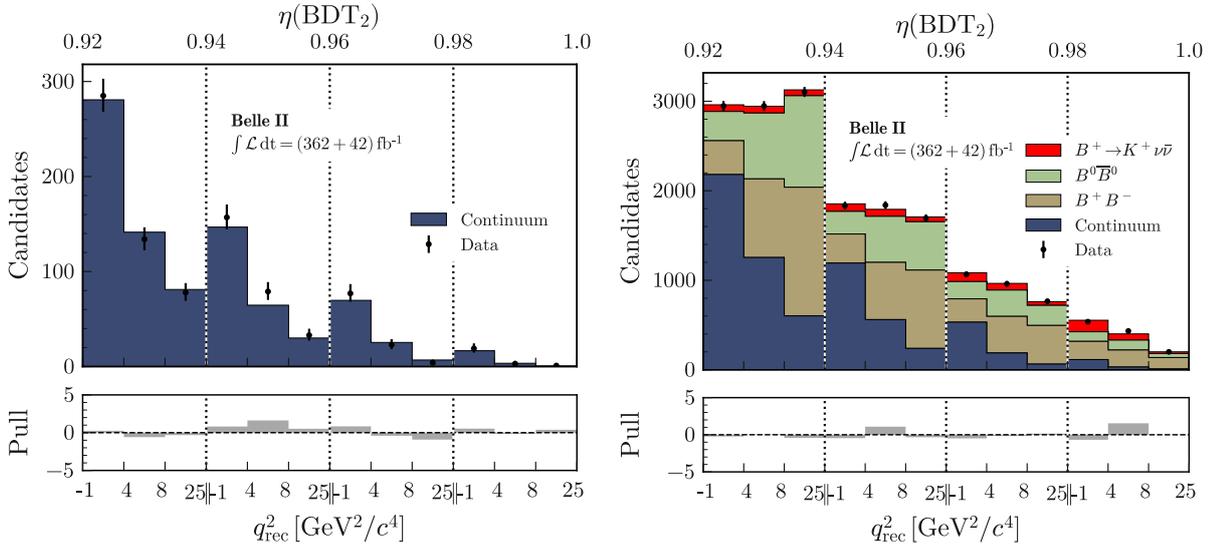


Figure 8.2: The best-fit yields of the **ITA** analysis as a function of $\eta(\text{BDT}_2) \times q_{\text{rec}}^2$. The off-resonance fit is shown on the left, the on-resonance fit on the right. Pull distributions are shown in the lower panels. Figure taken from Reference [2].

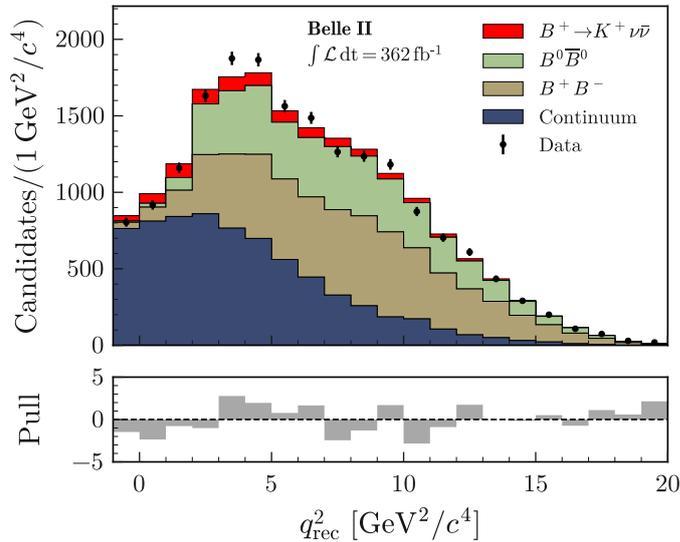


Figure 8.3: The best-fit projection of the **ITA** analysis onto the q_{rec}^2 variable. Pull distributions are shown in the lower panels. Figure taken from Reference [2].

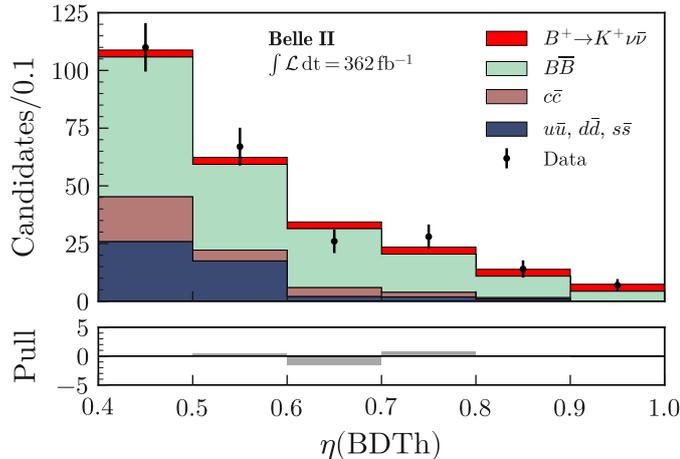


Figure 8.4: The best-fit yields of the HTA analysis as a function of $\eta(\text{BDTh})$. Pull distributions are shown in the lower panels. Figure taken from Reference [2].

The ITA and HTA likelihoods were combined into one likelihood, accounting for the correlations between the systematic uncertainties in the two methods (for details, see section XIV of Reference [2]). The *combined* likelihood includes 231 nuisance parameters χ in addition to one POI, the signal strength μ . A maximum likelihood fit yields a signal strength of [2]

$$\mu = 4.6 \pm 1.0 \text{ (stat.)} \pm 0.9 \text{ (syst.)} = 4.6 \pm 1.3, \quad (8.8)$$

which corresponds to a branching ratio of [2]

$$\mathcal{B}(B^+ \rightarrow K^+\nu\bar{\nu}) = (2.3 \pm 0.5(\text{stat})_{-0.4}^{+0.5}(\text{syst})) \cdot 10^{-5}. \quad (8.9)$$

This result has a significance of 3.5 standard deviations over the background-only hypothesis and a significance of 2.7 standard deviations over the SM hypothesis. The combined likelihood serves as the basis for reinterpretation in this study.

8.2 Joint number densities for the Belle II $B^+ \rightarrow K^+\nu\bar{\nu}$ analysis

To construct a model-agnostic likelihood for the $B^+ \rightarrow K^+\nu\bar{\nu}$ analysis [2], I follow the method introduced in Chapter 7. More concretely, since the $B^+ \rightarrow K^+\nu\bar{\nu}$ analysis is performed on binned data, I use the discrete reinterpretation method of Section 7.1.1.

Recall that the method requires the definition of the kinematic variables z , the reconstruction variables x , the null and alternative distributions σ_0 and σ_1 , and the joint number density $\nu_{0,xz}$, which is the basis of the model-agnostic likelihood.

Since the decay of the pseudoscalar B meson is isotropic in its rest frame, the only experimentally accessible kinematic d.o.f. is the squared dineutrino invariant mass, q^2 , as mentioned in Section 2.4.1. Hence, the only kinematic variable in this case is $z = q^2$.

The reconstruction variables are different for the **ITA** and **HTA** methods, as discussed in [Section 8.1](#). In the **ITA**, the reconstruction variables are $x = \eta(\text{BDT}_2) \times q_{\text{rec}}^2$. In the **HTA**, the reconstruction variable is $x = \eta(\text{BDTh})$.

The null distribution $\sigma_0(q^2)$ is the **SM** prediction for the $B^+ \rightarrow K^+\nu\bar{\nu}$ decay, which is presented in Reference [\[36\]](#), with form factors taken from Reference [\[46\]](#).

In this work, I will discuss two different cases for the alternative distribution $\sigma_1(q^2)$: one two-body and one three-body final-state scenario. In [Chapter 9](#), I will discuss the reinterpretation of the $B^+ \rightarrow K^+\nu\bar{\nu}$ measurement in the framework of the **WET**. In [Chapter 10](#), I will discuss the reinterpretation of the $B^+ \rightarrow K^+\nu\bar{\nu}$ measurement in a light new physics scenario, where the alternative distribution is constructed with the contribution of a two-body decay $B^+ \rightarrow K^+X$, with X being a new light particle that decays invisibly. This manifests itself as a narrow resonance peak in the q^2 distribution, which is not present in the **SM** prediction.

To obtain the joint number density ν_{0,xq^2} for both **ITA** and **HTA**, 10^7 (**ITA**) and $5 \cdot 10^7$ (**HTA**) simulated **SM** signal events from the $B^+ \rightarrow K^+\nu\bar{\nu}$ analysis [\[2\]](#) are used, to which all selection criteria are applied. These samples include information on the generated and reconstructed squared momenta, q^2 and q_{rec}^2 , as well as the classifier responses $\eta(\text{BDT}_2)$ and $\eta(\text{BDTh})$.

The granularity of the q^2 binning for ν_{0,xq^2} is determined by the differences between the null and the anticipated alternative distributions. The usual guideline of ensuring sufficiently many events in each bin does not apply here, since the binning serves only as a temporary partitioning for the application of kinematic weights. In principle, finer binning is preferable; however, beyond the intrinsic resolution of the analysis, no improvement is gained from extremely fine binning. The **WET** predicts a broad distribution in q^2 and hence would not require fine binning. However, the $B^+ \rightarrow K^+X$ scenario predicts a narrow resonance peak in the q^2 distribution, indicating a significant shape alteration. This requires finer binning to capture the shape of the distribution. Consequently, one should adopt very fine binning in q^2 , which will also facilitate future reinterpretations. I have therefore chosen to divide the q^2 range into 100 bins. In [Appendix D.3](#), I present a comparative study of different binning choices, which shows that the choice of 100 bins is sufficient to capture the shape of the q^2 distribution in both cases.

I add an additional bin in the range $-1 < q^2 < 0 \text{ GeV}^2$ to account for all events with unphysical or unassigned values. For the **ITA** analysis, no such events are present, while in the **HTA** analysis, 0.16% of the events fall into this category. The origin of these unphysical values is a problem in the **MC** truth matching, where an incorrect four-momentum is used to calculate q^2 . Excluding these events has no impact on the inferred signal strength and only minimal effect on the best-fit values of the new physics parameters considered in this study (see [Appendix D.4](#)). I therefore keep these events in the analysis for consistency. The corresponding bin remains unweighted in the reinterpretation process.

The joint number densities, ν_{0,xq^2} , for the **ITA** and **HTA** are shown in [Figure 8.5](#). Together with the $B^+ \rightarrow K^+\nu\bar{\nu}$ likelihood and the corresponding null distribution, one can perform

a reinterpretation of the $B^+ \rightarrow K^+ \nu \bar{\nu}$ measurement in terms of alternative theoretical models, as described in the following chapters.

A key difference between the **ITA** and **HTA** number densities is the correlation between the q^2 distributions in the kinematic bins and the reconstruction bins. In the **ITA** analysis, which uses binning in q_{rec}^2 , this correlation is significantly higher than in the **HTA** case, where it remains low. A strong correlation, as in the **ITA** analysis, improves sensitivity to variations in the kinematic shape, since changes in the q^2 distribution affect the individual reconstruction bins in a more distinctive way. I will discuss the implications of this in **Section 9.2**.

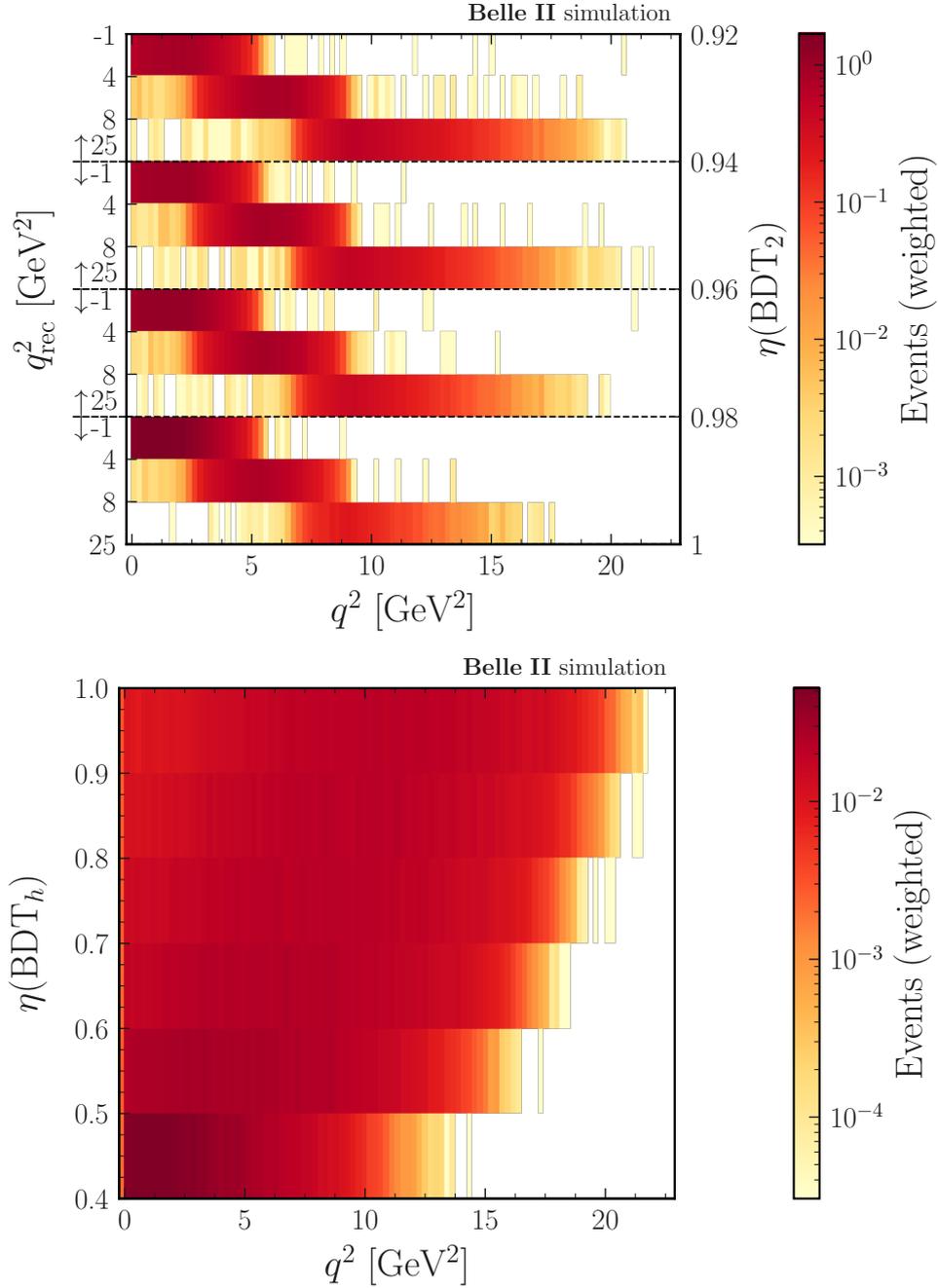


Figure 8.5: The *ITA* (top) and *HTA* (bottom) binned joint number densities. The horizontal axis corresponds to the generated q^2 . The vertical axis represents the binning used in the $B^+ \rightarrow K^+ \nu \bar{\nu}$ analysis [2]. The heatmap shows the weighted signal events.

Chapter 9

A Weak Effective Theory reinterpretation of the Belle II $B^+ \rightarrow K^+ \nu \bar{\nu}$ measurement

¹ This chapter presents the first-ever reinterpretation of the Belle II $B^+ \rightarrow K^+ \nu \bar{\nu}$ measurement in terms of the **WET**. It demonstrates how the model-agnostic likelihood method developed in **Chapter 7** can be used to extract new physics insights from experimental data beyond the original analysis scope.

The **WET** provides a systematic, model-independent framework for parametrizing potential new physics contributions to the $b \rightarrow s \nu \bar{\nu}$ transition through Wilson coefficients, which quantify deviations from the **SM** prediction. By reinterpreting the Belle II data within this framework, this work presents the first-ever measurement of the $b \rightarrow s \nu \bar{\nu}$ **WET** Wilson coefficients from experimental data with proper treatment of experimental uncertainties and kinematic shape information.²

A central question addressed in this chapter is whether the observed excess can be attributed solely to an enhancement in the vector sector (**SM**-like kinematics), or whether additional contributions from scalar or tensor operators are required to optimally describe the data. This distinction is crucial, as different patterns of Wilson coefficients point to different classes of new physics models.

The analysis employs Bayesian inference to derive posterior distributions and credible intervals for the Wilson coefficients. To validate the results, goodness-of-fit tests and model comparison studies assess the compatibility of theoretical hypotheses with the

¹The material presented in this chapter has been submitted to PRD [116]. While the content largely follows that of the referenced paper, editorial modifications have been made, including restructuring of the presentation, rephrasing of portions of the text, and expansion of selected explanations. As this project is the result of collaborative work of the Belle II collaboration, individual contributions of the proponents of this work, and further collaboration members are documented in **Appendix D.1**.

²Excluding “naive” reinterpretation approaches that neglect kinematic shape information.

observed data.

This chapter is structured as follows: The theoretical framework and parametrization details for the **WET** predictions are summarized in [Section 9.1](#). The main results, including the marginal posterior distributions and credible intervals for the Wilson coefficients, are presented in [Section 9.2](#). Statistical validation through goodness-of-fit tests is provided in [Section 9.3](#), followed by a model comparison study in [Section 9.4](#). Finally, [Section 9.5](#) summarizes the main results.

9.1 Weak Effective Theory as alternative hypothesis

This reinterpretation study is performed within the framework of the **WET**, discussed in detail in [Section 2.2.1](#). The prediction for the differential decay width of the $B^+ \rightarrow K^+\nu\bar{\nu}$ decay is given by [Equation \(2.62\)](#), which extends the **SM** prediction by including additional operator contributions beyond the **SM** vector operator.

The differential decay width is parametrized in terms of three combinations of Wilson coefficients ([Equation \(2.64\)](#)):

$$|C_{VL} + C_{VR}|, \quad |C_{SL} + C_{SR}|, \quad |C_{TL}|, \quad (9.1)$$

corresponding to vector, scalar, and tensor operator contributions, respectively. Due to the sensitivity to only the absolute values of these sums, this analysis treats each combination as a non-negative real-valued parameter.

In this reinterpretation study, the form factors are parametrized following the **BSZ** parametrization [\[48\]](#). The parametrization is truncated at second order in the series expansion, yielding a total of 8 hadronic parameters that fully describe the relevant form factors $f_+(q^2)$, $f_0(q^2)$, and $f_T(q^2)$ (see [Section 2.3](#)).

The values and uncertainties of these 8 hadronic parameters are extracted from a joint theoretical prior **PDF** comprised of the 2021 lattice world average based on results by the Fermilab/MILC and HPQCD collaborations [\[45\]](#) and recent results by the HPQCD collaboration [\[46\]](#); see [Table 2.3](#). This joint prior accounts for correlations between the form factor parameters, which are taken into account through their covariance matrices. Notably, the **SM** prediction used in the original $B^+ \rightarrow K^+\nu\bar{\nu}$ analysis [\[2\]](#) was based solely on hadronic parameters from the HPQCD collaboration [\[46\]](#), and only the vector form factor $f_+(q^2)$ contributes to the **SM** decay width (see [Sections 2.4.1](#) and [8.1](#)).

All theoretical predictions for the differential distributions are computed using the **EOS** software package in version 1.0.16 [\[37, 47\]](#).

9.2 Marginal posterior and credible intervals

This section presents the Bayesian analysis of the Belle II $B^+ \rightarrow K^+\nu\bar{\nu}$ data under the **WET** hypothesis, deriving posterior distributions and credible intervals for the Wilson

coefficients. The analysis employs the model-agnostic likelihood described in [Chapter 8](#), and the Bayesian inference framework described in [Section 4.4.2](#), implemented using the `bayesian pyhf` package [74]. The posterior distribution is sampled using `MCMC` methods implemented in the `PyMC` software [80].

9.2.1 Statistical model and prior choices

To obtain a marginal posterior for the `WET` Wilson coefficients, the $B^+ \rightarrow K^+ \nu \bar{\nu}$ statistical model is extended by introducing 11 additional parameters beyond those already present in the original analysis. These comprise three unconstrained `POIs`,

$$\boldsymbol{\eta} = [C_{\text{VL}} + C_{\text{VR}}, C_{\text{SL}} + C_{\text{SR}}, C_{\text{TL}}], \quad (9.2)$$

representing the Wilson coefficient combinations of interest, along with 8 nuisance parameters that encode the hadronic form factor uncertainties.

The hadronic parameters are correlated according to their joint prior distribution from lattice `QCD`. These 8 correlated parameters are decorrelated via eigendecomposition of their covariance matrix, transforming them into 8 independent Gaussian-distributed nuisance parameters (see [Sections 4.4.4.1](#) and [7.2](#)).

Importantly, the original $B^+ \rightarrow K^+ \nu \bar{\nu}$ statistical model already contained 3 nuisance parameters for the hadronic parameters entering the `SM` prediction (corresponding to the vector form factor $f_+(q^2)$). These are removed from the model to avoid double counting when the full set of 8 form factor parameters is introduced for the `WET` interpretation. After this modification, the complete statistical model contains 239 parameters: 3 Wilson coefficients of interest, 8 hadronic nuisance parameters, and 228 systematic nuisance parameters from the original $B^+ \rightarrow K^+ \nu \bar{\nu}$ analysis.

For the Wilson coefficient parameters, uniform priors are adopted over the range $[0, 20]$. Uniform priors are chosen to avoid assigning preference to any part of the parameter space and because no non-linear transformation of the Wilson coefficients is anticipated in this study. The upper bound of 20 is chosen to be sufficiently large to encompass the full posterior distribution, as verified a posteriori. The robustness of the results with respect to alternative prior choices is examined in detail in [Section 9.2.5](#).

9.2.2 Posterior sampling and marginalization

The marginal posterior distribution for the Wilson coefficients is obtained as follows. First, the joint posterior of all 239 parameters is sampled using `MCMC` methods. Subsequently, this joint posterior is marginalized over the 236 nuisance parameters (8 hadronic and 228 systematic parameters), leaving only the 3-dimensional posterior distribution over the Wilson coefficients.

The `MCMC` sampling employs 8 independent Markov chains, each generating 50 000 samples, for a total of 400 000 posterior samples. Multiple independent chains enable convergence diagnostics through inter-chain comparison (see [Appendix B.10](#)). Convergence

diagnostics are reported in [Appendix E.2](#), confirming that the chains have converged to the target posterior distribution and that the sampling is reliable.

To improve computational efficiency, the symmetry of the differential decay width expression in [Equation \(2.62\)](#) under sign changes of the Wilson coefficients is exploited. The sampling is performed only in the octant where all three Wilson coefficients are positive ($C_{VL} + C_{VR} \geq 0$, $C_{SL} + C_{SR} \geq 0$, $C_{TL} \geq 0$), and the resulting samples are then reflected across the symmetry axes to populate all octants.

The full set of 1- and 2-dimensional marginal posterior distributions and the resulting credible intervals at 68% and 95% probability is shown in [Figure 9.1](#).

9.2.3 Credible intervals and posterior mode

From the 1-dimensional marginal posterior distributions, I calculate the HDIs (see [Section 5.4.1](#)) at 68% and 95% probability for the Wilson coefficient combinations. The posterior mode and the credible intervals are summarized in [Table 9.1](#).

Table 9.1: The mode of the posterior, and HDI at 68% and 95% for the (sums of the) WET Wilson coefficients in [Equation \(2.62\)](#), derived from the posterior in [Figure 9.1](#).

| Parameters | Mode | 68% HDI | 95% HDI |
|---------------------|------|-------------|-------------|
| $ C_{VL} + C_{VR} $ | 11.3 | [7.8, 14.6] | [1.9, 16.2] |
| $ C_{SL} + C_{SR} $ | 0.0 | [0.0, 9.6] | [0.0, 15.4] |
| $ C_{TL} $ | 8.2 | [2.3, 9.6] | [0.0, 11.2] |

Direct comparison of the observed data with the predicted yields at the posterior mode (see [Table 9.1](#)) for the unconstrained $B^+ \rightarrow K^+ \nu \bar{\nu}$ SM (μ unconstrained) and the WET is shown in [Figure 9.2](#), for the Highest-Sensitivity Region (HSR) of the analysis (ITA, $\eta(\text{BDT}_2) > 0.98$). The WET model provides a better fit to the data than the unconstrained SM prediction, as indicated by the smaller pull values.

9.2.4 Physical interpretation of the posterior

The posterior distribution reveals a clear deviation from the SM prediction in the vector sector, consistent with the original analysis result of a signal strength $\mu = 4.6 \pm 1.3$ [[2](#)]. This connection can be understood quantitatively as follows: Since the Wilson coefficients enter quadratically into the differential branching ratio (see [Equation \(2.62\)](#)), the decay rate scales as $|C_{\gamma\delta}|^2$. For a pure vector contribution, the signal strength μ should satisfy

$$\mu = \frac{|C_{VL} + C_{VR}|^2}{|C_{VL}^{SM}|^2}, \quad (9.3)$$

where $C_{VL}^{SM} = 6.6$ is the SM value. This implies an expected vector coefficient of

$$|C_{VL} + C_{VR}| \simeq \sqrt{4.6} \cdot 6.6 \simeq 14.2, \quad (9.4)$$

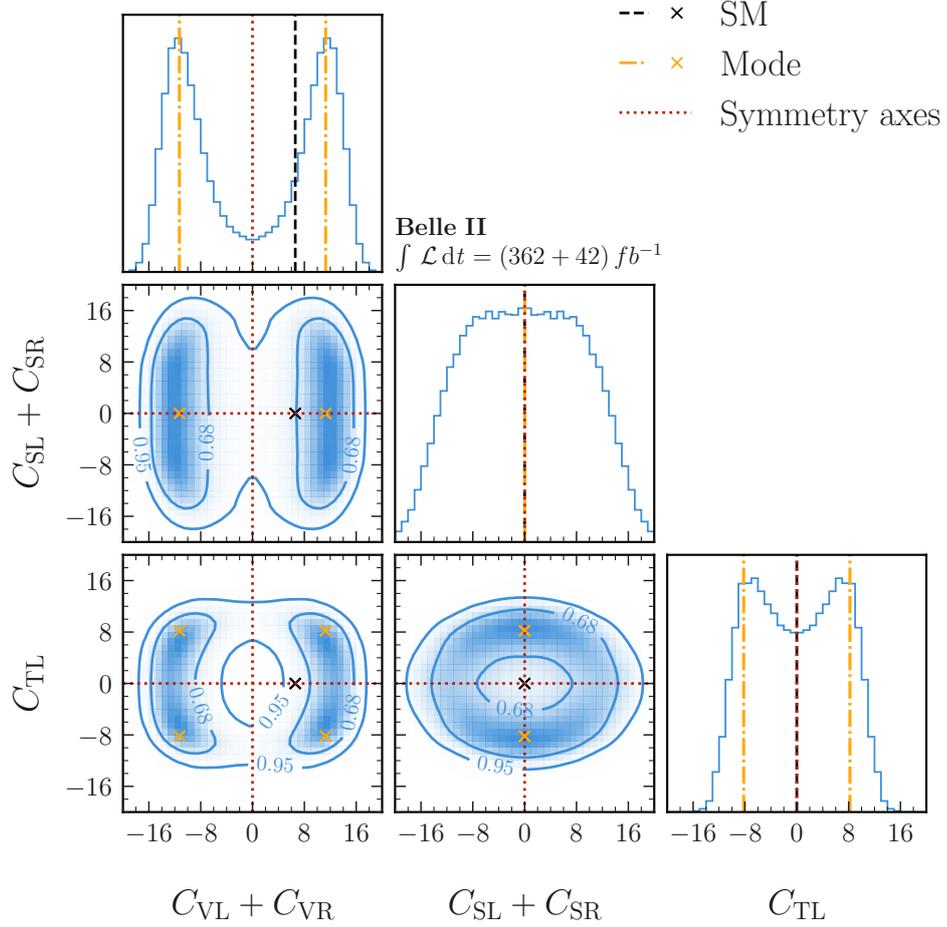


Figure 9.1: The marginalized posterior for the Wilson coefficients in Equation (2.62). The convention adopted is that $C_{VL} + C_{VR}$, $C_{SL} + C_{SR}$, and C_{TL} are real-valued. Diagonal and off-diagonal panels show the 1-dimensional and 2-dimensional sample density PDFs on a linear scale, respectively. The overall scale is omitted, as all relevant information is contained in the shape of the distribution. The contours indicate 68% and 95% credible intervals. The dashed black lines and cross mark the SM point; the dash-dotted yellow lines and cross indicate the posterior mode; dotted red lines mark the symmetry axes used for sample symmetrization.

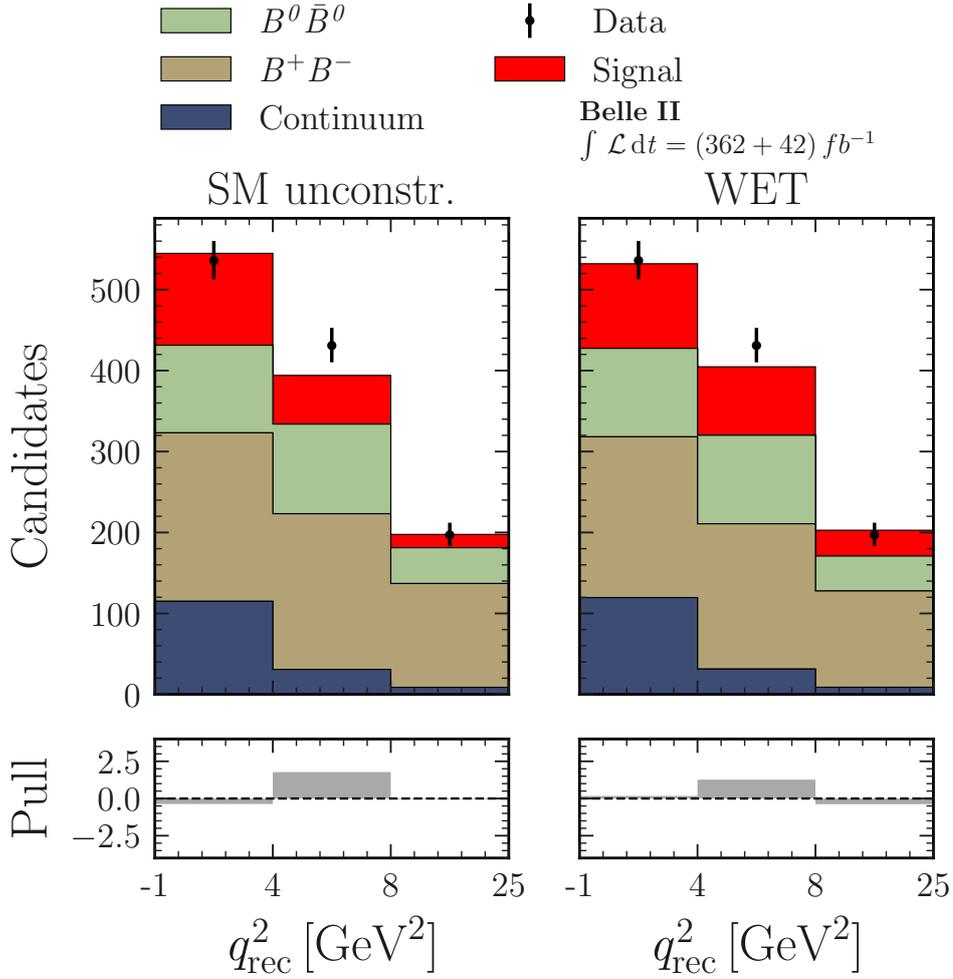


Figure 9.2: Observed and predicted best-fit yields in the **HSR** of the analysis. The signal is shown for the unconstrained $B^+ \rightarrow K^+ \nu \bar{\nu}$ **SM** (*left*) and the **WET** (*right*) predictions. The predicted background yields are shown individually for the neutral and charged B -meson decays, and the summed five continuum categories. Pulls are shown in the lower panels.

assuming negligible contributions from scalar and tensor operators. The observed posterior mode of 11.3 (see [Table 9.1](#)) is slightly lower than this naive expectation, indicating that the data prefer a model in which additional operator contributions share some of the excess rate.

To verify this interpretation, I perform a constrained fit fixing $C_{\text{SL}} + C_{\text{SR}} = 0$ and $C_{\text{TL}} = 0$, allowing only the vector coefficient to vary. This yields

$$C_{\text{VL}} + C_{\text{VR}} = 14.1 \pm 2.1, \quad (9.5)$$

which is indeed close to the naive expectation. To quantify the consistency between this vector-only interpretation and the signal strength measurement, I evaluate the ratio

$$\frac{|C_{\text{VL}} + C_{\text{VR}}|^2}{\mu |C_{\text{VL}}^{\text{SM}}|^2} = 1.0 \pm 0.4, \quad (9.6)$$

which equals unity within uncertainties, confirming the expected relationship.

Importantly, the posterior distribution reveals a preference for a non-zero tensor contribution, with the posterior mode at $C_{\text{TL}} = 8.2$. This finding indicates that a pure **SM**-like signal template does not provide the optimal description of the observed kinematic distributions. Instead, the data favour a linear combination of vector and tensor contributions, each with a distinct kinematic shape. This preference is detectable only because kinematic shape information is preserved in the reinterpretation procedure.

The marginal posterior and corresponding credible intervals derived independently from the **ITA** and **HTA** likelihoods are presented in [Appendix E.3](#). This decomposition provides insight into which analysis drives the observed deviations from the **SM**. The comparison reveals that the deviations in the vector and tensor sectors are features of the **ITA**-based posterior. The **HTA** plays a crucial complementary role by constraining the overall normalization. This normalization constraint is particularly important for the scalar and tensor sectors.

9.2.5 Prior sensitivity

Bayesian inference inherently depends on the choice of prior distributions. To assess the robustness of the presented results and quantify the influence of prior choices, I perform a prior sensitivity analysis by comparing the baseline uniform priors to two physically motivated alternative prior specifications.

The first alternative employs *truncated-normal priors* centred on the **SM** expectation, representing a scenario in which deviations from the **SM** are disfavoured a priori:

$$p(\eta_i) = \begin{cases} \mathcal{N}(\eta_i | \mu = C_i^{\text{SM}}, \sigma = 20) & \eta_i \geq 0 \\ 0 & \eta_i < 0 \end{cases}, \quad (9.7)$$

where η_i correspond to the Wilson coefficient combinations in [Equation \(9.2\)](#), and $C_i^{\text{SM}} \in [6.6, 0.0, 0.0]$ represents the **SM** point. The standard deviation $\sigma = 20$ is chosen to allow substantial deviations while maintaining a preference for **SM**-like values.

The second alternative adopts *uniform priors in the squared Wilson coefficients*, which translates to triangular priors for the Wilson coefficients themselves:

$$p(\eta_i) \propto \begin{cases} \eta_i & \eta_i \leq 30 \\ 0 & \eta_i > 30 \end{cases}. \quad (9.8)$$

This choice is motivated by the fact that Wilson coefficients enter the decay rate quadratically (see Equation (2.62)), so uniform priors on $|C_{\gamma\delta}|^2$ correspond to uniform priors on the decay rate enhancement. These priors favour larger values of the Wilson coefficients relative to the uniform priors.

The resulting posterior modes and credible intervals for both alternative prior choices are summarized in Table 9.2. The comparison reveals that the vector and tensor Wilson coefficients are the most robustly determined, with the posterior modes and credible intervals showing minimal sensitivity to the prior choice.

In contrast, the scalar Wilson coefficients exhibit the largest prior dependence, with the posterior mode shifting from 0.0 (uniform and truncated-normal priors) to 8.9 (triangular prior). This heightened sensitivity reflects the weak experimental constraints on the scalar sector due to the kinematic distribution peaking at high q^2 (see Figure 2.2), where the experimental efficiency is low (see Figure 8.1). When the data provide weak constraints, the posterior becomes more strongly influenced by the prior choice, as seen here.

Table 9.2: The posterior modes, and HDIs at 68% and 95% for the (sums of the) WET Wilson coefficients in Equation (2.62), for alternative prior choices (cf. Table 9.1).

| Priors | Parameters | Mode | 68% HDI | 95% HDI |
|----------------|---------------------|------|-------------|-------------|
| Equation (9.7) | $ C_{VL} + C_{VR} $ | 11.4 | [9.0, 14.6] | [2.2, 16.4] |
| | $ C_{SL} + C_{SR} $ | 0.0 | [0.0, 9.2] | [0.0, 14.7] |
| | $ C_{TL} $ | 7.7 | [1.5, 8.8] | [0.0, 11.0] |
| Equation (9.8) | $ C_{VL} + C_{VR} $ | 11.6 | [8.2, 14.0] | [4.2, 16.0] |
| | $ C_{SL} + C_{SR} $ | 8.9 | [4.6, 12.6] | [1.3, 15.6] |
| | $ C_{TL} $ | 7.2 | [3.9, 9.6] | [1.4, 11.7] |

9.2.6 Naive reinterpretation

The key advantage of the model-agnostic likelihood approach is that kinematic shape information is utilized for reinterpretation. To quantify the impact of this feature and demonstrate its necessity, I perform a comparative “naive” reinterpretation that discards all kinematic shape information.

Technically, this naive approach is implemented by integrating the joint number densities over the entire kinematic range. This produces a single overall normalization constraint,

discarding all differential information. This procedure is equivalent to performing inference using only the total number of observed signal events, without knowledge of how those events are distributed in q^2 .

The resulting marginalized posterior distribution from this naive approach is shown in [Figure 9.3](#).

Within this naive framework, the resulting 95% credible intervals are $|C_{VL} + C_{VR}| < 14.4$, $|C_{SL} + C_{SR}| < 8.5$, and $|C_{TL}| < 7.1$. Comparing these naive bounds to the full kinematic reinterpretation results in [Table 9.1](#) (95% HDI: [1.9, 16.2], [0.0, 15.4], [0.0, 11.2], respectively) reveals systematic biases introduced by neglecting shape information.

The most significant difference can be observed in the vector sector, where the naive upper limit of 14.4 underestimates the full result's upper bound of 16.2. In addition, the naive approach yields only an upper limit, failing to establish a lower bound, whereas the full analysis provides a 95% credible interval starting at 1.9.

For the scalar and tensor sectors, the naive bounds appear tighter than the full analysis bounds. However, this apparent improvement is illusory — it reflects biased central values rather than genuine enhanced precision. Any combination of vector, scalar, and tensor contributions that yields the observed total excess rate is equally plausible in the naive framework.

These systematic biases and loss of discriminating power highlight the importance of preserving kinematic shape information in reinterpretation studies.

9.3 Goodness-of-fit

Beyond obtaining credible intervals for the Wilson coefficients, it is important to validate that the [WET](#) model provides a good description of the observed data.

I perform goodness-of-fit tests for several competing hypotheses: the [WET](#) model with three free Wilson coefficients, the unconstrained [SM](#) where only the signal strength μ is allowed to vary, and a background-only hypothesis that assumes no $B^+ \rightarrow K^+ \nu \bar{\nu}$ signal contribution. The unconstrained [SM](#) considers a signal which exhibits [SM](#)-like kinematics, but with an overall free normalization (equivalent to a vector-only contribution of the [WET](#)). A detailed summary of these models and their defining characteristics is provided in [Table 9.3](#). The statistical framework for the goodness-of-fit test is described in [Section 5.3](#).

The goodness-of-fit P -value is calculated from (see [Section 5.3](#))

$$P_{\text{gof}} = \int_{t_{\text{obs}}}^{\infty} dt_{\text{gof}} p(t_{\text{gof}}), \quad t_{\text{gof}} = -2 \ln \frac{p(\mathbf{n}, \mathbf{a} | \hat{\boldsymbol{\eta}}, \hat{\boldsymbol{\chi}})}{p_{\text{sat}}(\mathbf{n}, \mathbf{a} | \bar{\boldsymbol{\chi}})}, \quad (9.9)$$

where $p(\mathbf{n}, \mathbf{a} | \hat{\boldsymbol{\eta}}, \hat{\boldsymbol{\chi}})$ is the likelihood evaluated at the best-fit point for the model under consideration, with $\hat{\boldsymbol{\eta}}$ denoting the fitted [POIs](#) (Wilson coefficients or signal strength) and $\hat{\boldsymbol{\chi}}$ the fitted nuisance parameters. The saturated likelihood $p_{\text{sat}}(\mathbf{n}, \mathbf{a} | \bar{\boldsymbol{\chi}})$ represents a

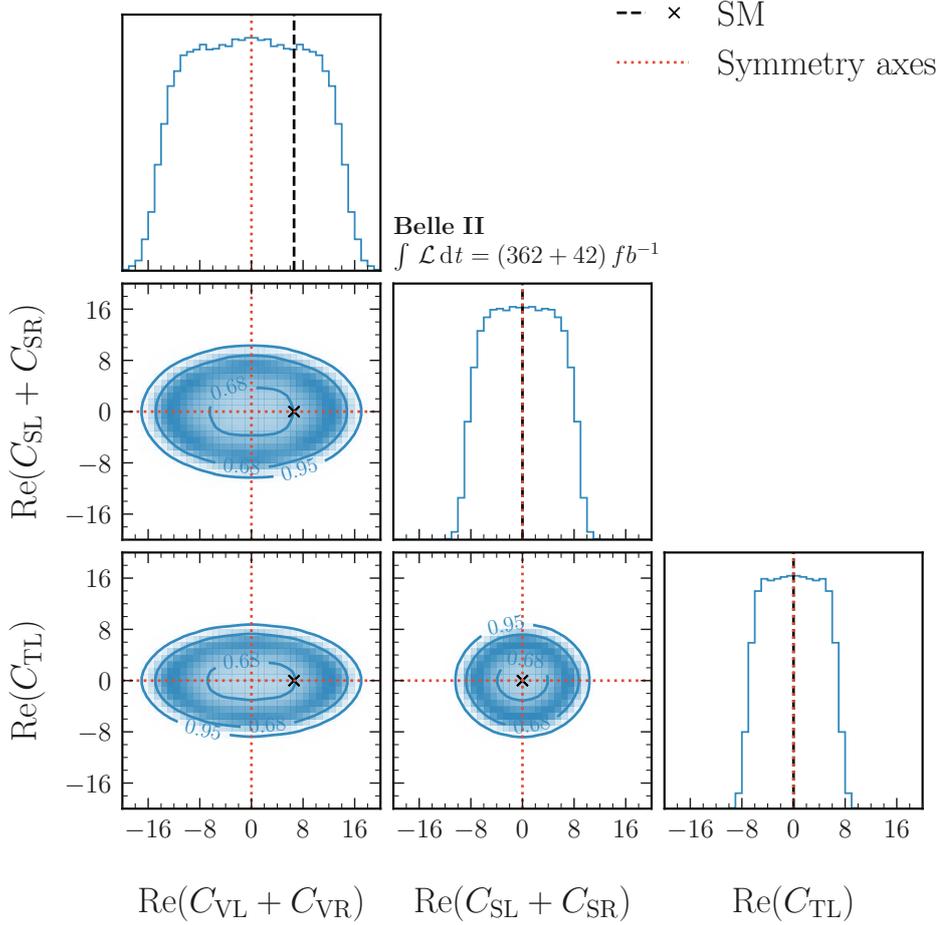


Figure 9.3: The marginalized posterior for the Wilson coefficients in Equation (2.62), based on a naive reinterpretation approach. The convention adopted is that $C_{VL} + C_{VR}$, $C_{SL} + C_{SR}$, and C_{TL} are real-valued. Diagonal and off-diagonal panels show the 1-dimensional and 2-dimensional sample density PDFs on a linear scale, respectively. The overall scale is omitted, as all relevant information is contained in the shape of the distribution. The contours indicate 68% and 95% credible intervals. The dashed black lines and cross mark the SM point; dotted red lines mark the symmetry axes used for sample symmetrization.

Table 9.3: A summary of the models under consideration, their corresponding theoretical predictions, references for the hadronic parameters, and the number of hadronic parameters included in each model is provided.

| Model | Prediction | Free param. | Hadronic param. | |
|-----------------------|--|---|-----------------|-----|
| | | | Ref. | Nr. |
| WET | Equation (2.62) | $C_{\text{VL}} + C_{\text{VR}},$ $C_{\text{SL}} + C_{\text{SR}},$ C_{TL} | [45, 46] | 8 |
| SM unconstrained | Equation (2.62), $C_{\text{VL}} \geq 0,$ $C_{\gamma\delta} = 0$ otherwise | μ | [46] | 3 |
| SM constrained | Equation (2.62), $C_{\text{VL}} = 6.6,$ $C_{\gamma\delta} = 0$ otherwise, 5% norm. unc. | – | [46] | 3 |
| background-only (BKG) | No signal | – | – | – |

perfect fit where the expected event rates in each bin of the `HistFactory` likelihood [70] are set exactly equal to the observed data, with all constraint terms maximized at $\bar{\chi}$.

The PDF of the test statistic $p(t_{\text{gof}})$ is obtained from fits to toy data sampled from $p(\mathbf{n}, \mathbf{a} \mid \hat{\boldsymbol{\eta}}, \hat{\boldsymbol{\chi}})$ [126]. The goodness-of-fit P -values are calculated as the fraction of toys with $t > t_{\text{obs}}$, where t_{obs} is the value of the test statistic for the observed data.

For the WET model, the distribution of the goodness-of-fit test statistic obtained from 10 000 toy experiments is shown in Figure 9.4. Under standard asymptotic assumptions, one expects this distribution to follow a χ^2 -distribution with $N_{\text{dof}} = 30 - 3 = 27$ d.o.f. (30 bins in the analysis minus 3 free Wilson coefficients). However, a fit to the distribution from toys yields $N_{\text{dof}} = 27.7 \pm 0.1$, showing a small but statistically significant deviation.

This deviation from the asymptotic expectation can be understood from the structure of the WET model: the Wilson coefficients can only contribute positively to the decay rate (they enter quadratically in Equation (2.62)), and the accessible parameter space is therefore bounded at zero. Such boundary effects violate the conditions required for Wilks' theorem, which guarantees asymptotic χ^2 behaviour only when parameters are unconstrained. The presence of such boundaries modifies the test statistic distribution, as discussed in Reference [127]. For this reason, I use the distribution from toys rather than the asymptotic approximation to calculate P -values.

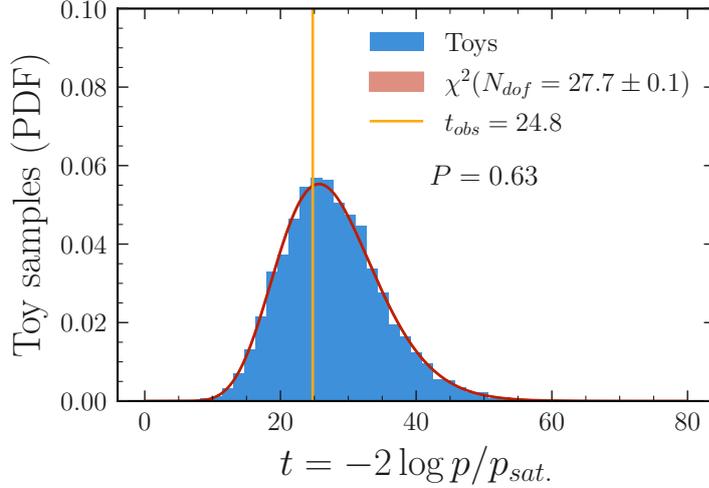


Figure 9.4: The distribution of the goodness-of-fit test statistic for the **WET** model, obtained from 10 000 toys. The value obtained from a fit to the data is shown as vertical yellow line. The asymptotic χ^2 -distribution (red band) corresponds to the best-fit N_{dof} , which is shown with the corresponding uncertainty band.

Test statistic distributions for the unconstrained **SM** and the background-only model are reported in [Appendix E.4](#).

The goodness-of-fit P -values for all models are summarized in [Table 9.4](#). Both the **WET** ($P_{\text{gof}} = 0.63$) and the unconstrained **SM** ($P_{\text{gof}} = 0.58$) show good fits to the data. The background-only model yields a lower P -value of 0.13, indicating a less favourable fit to the data.

9.4 Model comparison

Model comparison quantifies the *relative* support that the data provide for competing hypotheses. I perform both Bayesian and frequentist model comparisons to assess the relative performance of the **WET** and unconstrained **SM** hypotheses against two reference models: the background-only hypothesis (BKG), and the constrained **SM** (see [Table 9.3](#)).

The constrained **SM** model (see [Table 9.3](#)) fixes $\mu = 1$ to the **SM** prediction. To account for theoretical normalization uncertainties in the **SM** prediction, this model includes a 5% relative uncertainty, which combines in quadrature a 4.4% uncertainty from the **CKM** matrix elements $|V_{ts}^* V_{tb}|^2$ and a 2.3% uncertainty from the Wilson coefficient calculation $|C_{\text{VL}}^{\text{SM}}|^2$ [[36](#)].

9.4.1 Bayes factors

Bayesian model comparison employs the Bayes factor, defined as the ratio of marginal likelihoods between two competing models (see [Section 5.2.2](#)).

The Bayes factors are computed using sequential **MC** sampling methods [80, 128], which provide numerical estimates of the marginal likelihoods for each model. The results are presented in [Table 9.4](#).

Both the **WET** ($\log_{10} B_{\text{BKG}} = 1.78$) and the unconstrained **SM** ($\log_{10} B_{\text{BKG}} = 2.02$) exceed Jeffreys’s threshold [87] of $\log_{10} B > 1.5$ for a *very strong* model preference (see [Table 5.1](#)).

Comparing both models to the constrained **SM**, I find *substantial* model preference ($\log_{10} B > 0.5$) for the unconstrained scenarios: $\log_{10} B_{\text{SM}}^{\text{constr.}} = 0.68$ for the **WET** and $\log_{10} B_{\text{SM}}^{\text{constr.}} = 0.92$ for the unconstrained **SM**.

Interestingly, the unconstrained **SM** shows slightly stronger evidence than the **WET**, suggesting that the data do not strongly require the additional **d.o.f.** provided by the scalar and tensor Wilson coefficients — the vector enhancement alone largely suffices to explain the observations.

Table 9.4: Bayes factors and goodness-of-fit P -values for the **WET** and unconstrained **SM**. The Bayes factors are reported relative to the background-only hypothesis (B_{BKG}) and the constrained $B^+ \rightarrow K^+ \nu \bar{\nu}$ **SM** hypothesis ($B_{\text{SM}}^{\text{constr.}}$).

| Model | $\log_{10} B_{\text{BKG}}$ | $\log_{10} B_{\text{SM}}^{\text{constr.}}$ | P_{gof} |
|-------------------------|----------------------------|--|------------------|
| WET | 1.78 | 0.68 | 0.63 |
| SM unconstrained | 2.02 | 0.92 | 0.58 |
| background-only | 0.00 | -1.08 | 0.13 |

9.4.2 Frequentist hypothesis test

To complement the Bayesian model comparison and provide a frequentist perspective, I perform hypothesis tests that quantify the significance of the signal hypotheses relative to the background-only hypothesis.

The tests are performed using the P -value defined as

$$P = \int_{t_{\text{obs}}}^{\infty} dt p(t), \quad t = -2 \ln \frac{p(\mathbf{n}, \mathbf{a} \mid \boldsymbol{\eta} = \mathbf{0}, \hat{\boldsymbol{\chi}})}{p(\mathbf{n}, \mathbf{a} \mid \hat{\boldsymbol{\eta}}, \hat{\boldsymbol{\chi}})}, \quad (9.10)$$

where $p(\mathbf{n}, \mathbf{a} \mid \hat{\boldsymbol{\eta}}, \hat{\boldsymbol{\chi}})$ is the likelihood at the best-fit point with all parameters free to vary, and $p(\mathbf{n}, \mathbf{a} \mid \boldsymbol{\eta} = \mathbf{0}, \hat{\boldsymbol{\chi}})$ is the likelihood with Wilson coefficients or signal strength fixed to zero (background-only hypothesis) but nuisance parameters still optimized at $\hat{\boldsymbol{\chi}}$.

The distribution of the test statistic $p(t)$ is determined by generating toy datasets from the background-only hypothesis $p(\mathbf{n}, \mathbf{a} \mid \boldsymbol{\eta} = \mathbf{0}, \hat{\boldsymbol{\chi}})$ and computing the test statistic for each toy. The P -value is the fraction of toys yielding test statistics larger than the one observed, $t > t_{\text{obs}}$.

For the **WET** hypothesis, the hypothesis test yields a P -value of $P = 4.6 \cdot 10^{-4}$, corresponding to a significance of $Z = 3.3\sigma$ with respect to the background-only hypothesis. The distribution of the test statistic obtained from 100 000 toy experiments is shown in **Figure 9.5**, where the observed value $t_{\text{obs}} = 13.9$ falls in the far tail of the distribution.

As discussed in **Section 9.3**, the test statistic distribution for the **WET** does not follow the asymptotic χ^2 -distribution due to the positivity constraints on Wilson coefficients. The distribution from toys shows that the test statistic value corresponding to $Z = 3.0\sigma$ significance is $t(Z = 3) = 11.5_{-0.2}^{+0.2}$, which is notably smaller than the observed $t_{\text{obs}} = 13.9$. This confirms that the observed excess represents a significant deviation from the background-only hypothesis.

For comparison, the unconstrained **SM** hypothesis yields a P -value of $P = 5.6 \cdot 10^{-4}$, corresponding to a significance of $Z = 3.3\sigma$. The test statistic distribution from 10 000 toy experiments is shown in **Figure 9.6**. In this case, the distribution closely follows a χ^2 -distribution with $N_{\text{dof}} = 1.02 \pm 0.01$ **d.o.f.**, as expected for a single free parameter. The agreement with the asymptotic expectation validates the use of asymptotic formulae for this simpler hypothesis. The observed test statistic is $t_{\text{obs}} = 12.0$.

The significances of $Z = 3.3\sigma$ for both the **WET** and unconstrained **SM** hypotheses indicate that both models provide comparable descriptions of the deviation from the background-only hypothesis. The slightly smaller P -value for the **WET** suggests a marginally better fit, consistent with the posterior preference for tensor contributions discussed earlier, though the difference is not substantial.

One can compare these results to the original Belle II publication [2], which reported a significance of $Z = 3.5\sigma$. That value was obtained using the approximation $Z = \sqrt{t_{\text{obs}}}$, valid for an asymptotic one-sided test with one **d.o.f.** (corresponding to an upper-limit test statistic as defined in Reference [79]). The difference between 3.5σ and the present 3.3σ arises mainly from the fact that the present analysis uses the likelihood ratio test statistic, rather than an upper limit test statistic. The updated form-factor parameters also play a minor role. Nonetheless, the conclusions remain consistent.

9.5 Discussion

This chapter presents the first comprehensive reinterpretation of the Belle II $B^+ \rightarrow K^+\nu\bar{\nu}$ measurement [2] in terms of the **WET** framework, demonstrating the power of the model-agnostic likelihood method developed in **Chapter 7**. Starting from the published analysis likelihood and joint number densities, I have constructed full Bayesian posterior distributions and derived credible intervals for the $b \rightarrow s\nu\bar{\nu}$ **WET** Wilson coefficients. This represents the first time that experimental data have been used to constrain this set of coefficients with proper treatment of uncertainties and kinematic shape information.

The principal findings are summarized in **Table 9.1**. The posterior mode for the Wilson coefficient combinations is located at $(|C_{\text{VL}} + C_{\text{VR}}|, |C_{\text{SL}} + C_{\text{SR}}|, |C_{\text{TL}}|) = (11.3, 0.0, 8.2)$, with 95% credible intervals of [1.9, 16.2], [0.0, 15.4], and [0.0, 11.2], respectively.

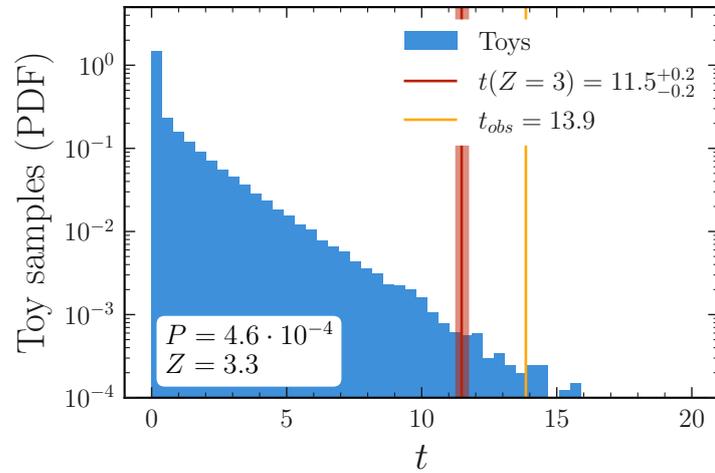


Figure 9.5: Distribution of the test statistic for the **WET** hypothesis over the background-only hypothesis, obtained from 100 000 toy experiments. The value from the fit to data is shown as a yellow line. The vertical red line indicates the test statistic corresponding to a significance of $Z = 3$, with the shaded band denoting its uncertainty.

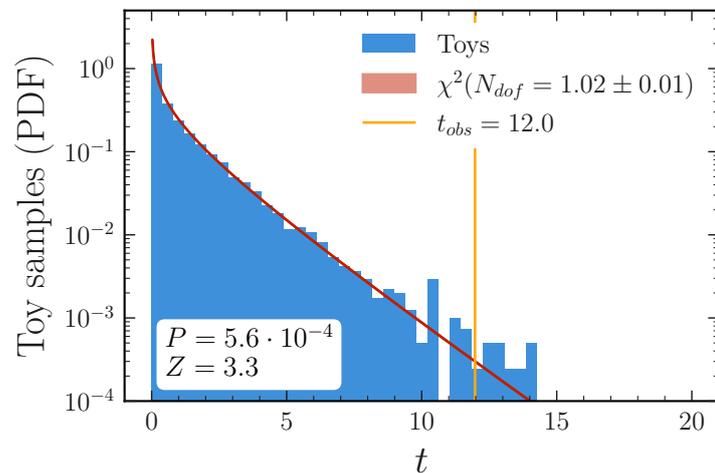


Figure 9.6: Distribution of the test statistic for the unconstrained **SM** hypothesis over the background-only hypothesis, obtained from 10 000 toy experiments. The value from the fit to data is shown as a vertical yellow line. The asymptotic χ^2 -distribution (red band) corresponds to the best-fit N_{dof} , which is shown with the corresponding uncertainty band.

The data require an enhancement in the vector sector beyond the **SM** expectation of $C_{\text{VL}}^{\text{SM}} = 6.6$. The 95% credible interval excludes the **SM** value, and the posterior sharply peaks at $C_{\text{VL}} + C_{\text{VR}} \approx 11.3$, consistent with the large signal strength $\mu = 4.6 \pm 1.3$ reported in the original analysis.

The data exhibit a preference for an additional tensor contribution, with the posterior mode at $C_{\text{TL}} = 8.2$. The preserved kinematic shape information allows for the identification of this preferred admixture of vector and tensor contributions.

The scalar Wilson coefficient remains weakly constrained, with the posterior consistent with zero but allowing values up to $|C_{\text{SL}} + C_{\text{SR}}| \sim 15$. This weak constraint stems from the unfavourable overlap of a kinematic distribution that peaks at high q^2 (see [Figure 2.2](#)) with the analysis efficiency, which is low for high- q^2 signals (see [Figure 8.1](#)).

Goodness-of-fit tests and model comparisons support the reliability of these conclusions. The **WET** model provides a good absolute fit to the data ($P_{\text{gof}} = 0.63$), comparable to the unconstrained **SM** ($P_{\text{gof}} = 0.58$). The background-only hypothesis provides a less favourable fit to the data ($P_{\text{gof}} = 0.13$). This confirms that the **WET** model adequately describes the data.

Bayesian model comparison via Bayes factors shows *very strong* preference ($\log_{10} B_{\text{BKG}} = 1.78$ for **WET**, 2.02 for unconstrained **SM**) for signal-plus-background models over the background-only hypothesis, and a *substantial* preference ($\log_{10} B_{\text{SM}}^{\text{constr.}} = 0.68$ for **WET**, 0.92 for unconstrained **SM**) over the constrained **SM** prediction. Interestingly, the unconstrained **SM** shows slightly stronger preference than the **WET**, suggesting that the additional complexity of scalar and tensor operators is not strongly required by the data.

Frequentist hypothesis tests support these findings, yielding a 3.3σ significance for both models against the background-only hypothesis. The significances confirm that both the **WET** and unconstrained **SM** provide strong evidence for new physics contributions. The **WET**'s marginally smaller P -value ($4.6 \cdot 10^{-4}$ vs. $5.6 \cdot 10^{-4}$) reflects its slightly improved fit from the tensor contribution.

The comparison to the naive reinterpretation in [Section 9.2.6](#) illustrates the value of preserving kinematic shape information. The naive approach, which discards all differential distribution information, yields biased credible intervals and fails to identify the preference for a tensor contribution. This comparison validates, that inference results can depend critically on the kinematic shape information, which reinterpretation approaches based on published summary results cannot utilize.

Chapter 10

A light new physics reinterpretation of the Belle II $B^+ \rightarrow K^+ \nu \bar{\nu}$ measurement

Having demonstrated the successful application of the model-agnostic reinterpretation method to extract **WET** Wilson coefficients in the previous chapter, this chapter explores an alternative theoretical scenario that addresses a specific feature in the experimental data. The **WET** reinterpretation in **Chapter 9** assumed heavy new physics integrated out at the electroweak scale. Here I investigate the possibility of light new physics contributing to the $B^+ \rightarrow K^+ \nu \bar{\nu}$ decay.

In the $B^+ \rightarrow K^+ \nu \bar{\nu}$ decay, the final-state neutrinos are not directly detected but inferred from missing energy and momentum in the detector. This does not only hold for **SM** neutrinos — any light, weakly-interacting particle X that escapes detection would produce such an experimental signature [10, 13–22].

The recent Belle II $B^+ \rightarrow K^+ \nu \bar{\nu}$ measurement [2] observed an excess of events compared to **SM** expectations, with a non-**SM**-like distribution in q_{rec}^2 . The enhancement was observed in the region $3 \text{ GeV}^2 < q_{\text{rec}}^2 < 7 \text{ GeV}^2$ (see **Figure 8.3** and **Section 8.1**). This would correspond to a resonance near $m_X \sim 2 \text{ GeV}$.

Simplified studies of light new physics scenarios in the context of the Belle II result have been published in References [112–114]. These analyses provide initial insights, but they rely on simplified background templates and neglect the full experimental systematic uncertainties.

This chapter addresses these limitations by performing the first reinterpretation of the Belle II $B^+ \rightarrow K^+ \nu \bar{\nu}$ measurement that includes the full experimental statistical model. Using the model-agnostic likelihood approach, I derive Bayesian credible intervals and frequentist confidence limits on both the new physics branching ratio $\mathcal{B}(B^+ \rightarrow K^+ X)$ and the mass m_X of the hypothetical light particle. The analysis is performed for several choices of the resonance width Γ_X .

This chapter is structured as follows: In [Section 10.1](#) I introduce the $B^+ \rightarrow K^+X$ model as an alternative hypothesis, and give the parametrized model used for reinterpretation. In [Section 10.2](#) I present the marginal posterior and credible intervals for the model parameters. In [Section 10.3](#) I present a frequentist mass scan, deriving upper limits on the branching ratio $\mathcal{B}(B^+ \rightarrow K^+X)$ as a function of the mass m_X . The goodness-of-fit test is shown in [Section 10.4](#), followed by the model comparison in [Section 10.5](#). Finally, I conclude with a discussion of the results in [Section 10.6](#).

10.1 The $B^+ \rightarrow K^+X$ model as alternative hypothesis

To test the hypothesis of a light invisible resonance contributing to the $B^+ \rightarrow K^+\nu\bar{\nu}$ signal, I construct a theoretical model that combines the [SM](#) prediction with a resonant two-body contribution.

The presence of an unknown two-body final state $B^+ \rightarrow K^+X$ would manifest itself as a sharp resonant peak in the q^2 spectrum at the squared mass of the new particle, $q^2 = m_X^2$. To model this resonance in a general way, I employ a relativistic Breit-Wigner [PDF](#),

$$p_{\text{BW}}(q^2|m_X, \Gamma_X) = \frac{k}{(q^2 - m_X^2)^2 + m_X^2\Gamma_X^2}, \quad (10.1)$$

$$k = \frac{m_X\Gamma_X}{\pi/2 + \arctan(m_X/\Gamma_X)},$$

where m_X is the mass of the resonance, Γ_X is its decay width, and k is a normalization constant.

Including both the [SM](#) contribution from the $b \rightarrow s\nu\bar{\nu}$ transitions and the new physics resonance, the total differential branching ratio is modelled as

$$\frac{d\mathcal{B}}{dq^2} = \mu \frac{d\mathcal{B}_{\text{SM}}}{dq^2} + \mu_X \sigma_X p_{\text{BW}}(q^2|m_X, \Gamma_X), \quad (10.2)$$

where μ is a scale factor for the [SM](#) contribution, and μ_X controls the strength of the new physics contribution. The factor $\sigma_X = 10^{-6}$ is introduced for numerical stability. The new physics branching ratio is given by $\mathcal{B}(B^+ \rightarrow K^+X) = \mu_X \cdot 10^{-6}$.

The [SM](#) prediction $d\mathcal{B}_{\text{SM}}/dq^2$ is computed using the theoretical framework described in [Section 2.4.1](#), with hadronic form factor parameters from the lattice [QCD](#) calculation by the HPQCD collaboration [46]. This is the same [SM](#) baseline used in the original Belle II analysis [2].

The model defined in [Equation \(10.2\)](#) contains five free parameters:

- μ : [SM](#) scale factor
- μ_X : New physics scale (related to $\mathcal{B}(B^+ \rightarrow K^+X)$ by a factor $\sigma_X = 10^{-6}$)
- m_X : Resonance mass

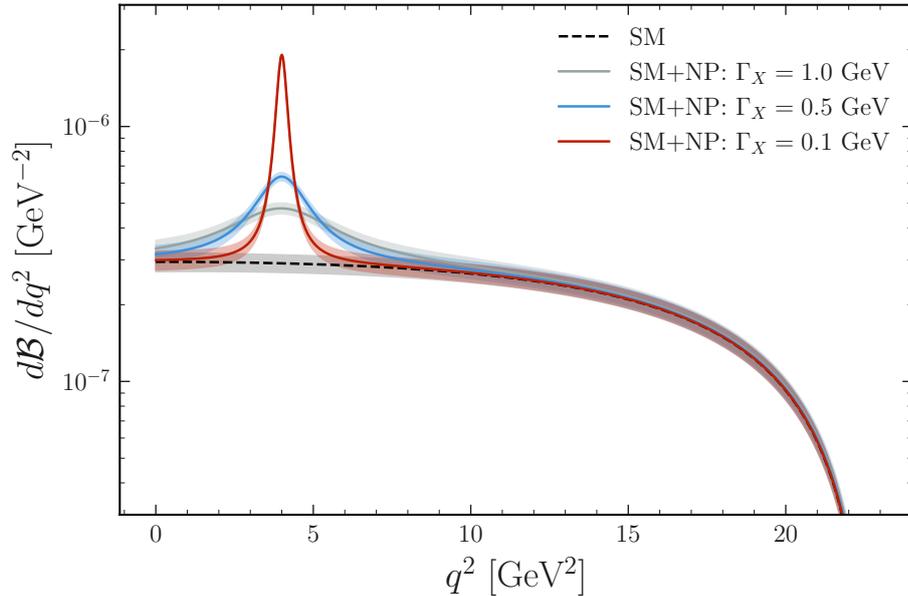


Figure 10.1: The predicted differential branching ratio from Equation (10.2) for three choices of resonance width, shown for the exemplary parameter point $\mu_X = 1$ and $m_X = 2$ GeV. The bands represent the combined theoretical uncertainty from hadronic form factor parameters and the 5% normalization uncertainty on the SM contribution.

- Γ_X : Resonance width

In this study, the SM contribution is treated as a background. Hence, the SM scale factor is constrained around unity, $\mu = 1.0$, with a 5% normalization uncertainty added to all other systematic uncertainties. This 5% uncertainty accounts for a 4.4% uncertainty on the CKM matrix elements $|V_{ts}^* V_{tb}|^2$, and a 2.3% uncertainty on the Wilson coefficient $|C_{VL}^{SM}|^2$ [36].

The primary POIs for this analysis are the new physics scale μ_X and the resonance mass m_X . Rather than treating the width Γ_X as a free parameter, I fix it to three representative values: $\Gamma_X = 0.1, 0.5, 1.0$ GeV. Different theoretical models for the particle X predict different widths. By considering multiple widths, I remain relatively model-agnostic. In addition, the analysis has a limited momentum transfer resolution, which makes widths below $\Gamma_X \lesssim 0.1$ GeV indistinguishable. Therefore, $\Gamma_X = 0.1$ GeV represents the narrow-width limit accessible to this analysis. All chosen widths lead to extremely short lifetimes of X , which is consequently assumed to decay invisibly.

The predicted differential branching ratio for these three width choices is illustrated in Figure 10.1 for an exemplary parameter point with $\mu_X = 1$ and $m_X = 2$ GeV.

Figure 10.2 shows the binned distributions for the null hypothesis and the alternative hypothesis for the same exemplary parameter point.

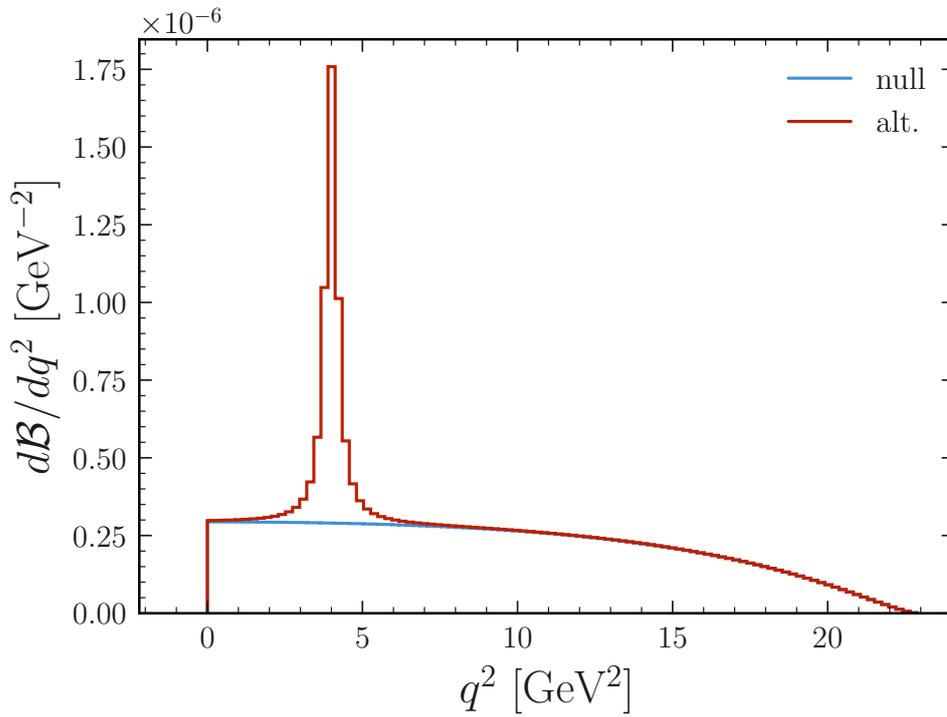


Figure 10.2: Comparison of bin-integrated distributions for the null hypothesis (blue, SM only [36]) and alternative hypothesis (red, SM plus $B^+ \rightarrow K^+ X$ resonance) according to Equation (10.2). The alternative distribution is shown for $\mu_X = 1$, $m_X = 2$ GeV, and $\Gamma_X = 0.1$ GeV.

10.2 Marginal posterior and credible intervals

This section presents the Bayesian analysis of the Belle II $B^+ \rightarrow K^+ \nu \bar{\nu}$ data under the $B^+ \rightarrow K^+ X$ hypothesis, deriving posterior distributions and credible intervals for the model parameters. The analysis employs the model-agnostic likelihood described in [Chapter 8](#), and the Bayesian inference framework described in [Section 4.4.2](#), implemented using the `bayesian pyhf` package [74]. The posterior distribution is sampled using [MCMC](#) methods implemented in the `PyMC` software [80].

10.2.1 Statistical model and prior choices

The objective is to obtain the marginal posterior distributions for the [POIs](#),

$$\boldsymbol{\eta} = [\mu_X, m_X], \quad (10.3)$$

for each of the three fixed resonance widths $\Gamma_X = 0.1, 0.5, 1.0$ GeV. To achieve this, I extend the baseline $B^+ \rightarrow K^+ \nu \bar{\nu}$ statistical model (which contains 232 parameters) with the two [POIs](#) μ_X and m_X . The complete statistical model contains 234 parameters: the 2 [POIs](#), 1 constrained [SM](#) scale parameter, and 231 additional nuisance parameters (the [SM](#) scale is considered a nuisance parameter here).

Prior distributions for the [POIs](#) are chosen to be uniform over the ranges specified in [Table 10.1](#). Uniform priors are chosen to avoid assigning preference to any part of the parameter space and because no non-linear transformation of the parameters is anticipated in this study. Ranges are chosen to cover the full posterior, as verified a posteriori. The robustness of the results with respect to alternative prior choices is examined in detail in [Section 10.2.5](#).

Table 10.1: Prior ranges for the new physics scale μ_X and resonance mass m_X , specified for each of the three resonance widths considered.

| $\Gamma_X =$ | 0.1 GeV | 0.5 GeV | 1.0 GeV |
|--------------|-------------|-------------|-------------|
| μ_X | [0.0, 24.0] | [0.0, 32.0] | [0.0, 42.0] |
| m_X [GeV] | [1.5, 3.0] | [1.5, 3.2] | [1.5, 3.4] |

10.2.2 Posterior sampling and marginalization

The marginal posterior distributions for μ_X and m_X are obtained as follows. First, the joint posterior of all 234 parameters is sampled using [MCMC](#) methods. Subsequently, this joint posterior is marginalized over the 232 nuisance parameters, leaving only the 2-dimensional posterior distribution over the [POIs](#), (μ_X, m_X) .

Each analysis (one per value of Γ_X) employs 8 independent Markov chains, each generating 50 000 samples, for a total of 400 000 posterior samples per width. Multiple

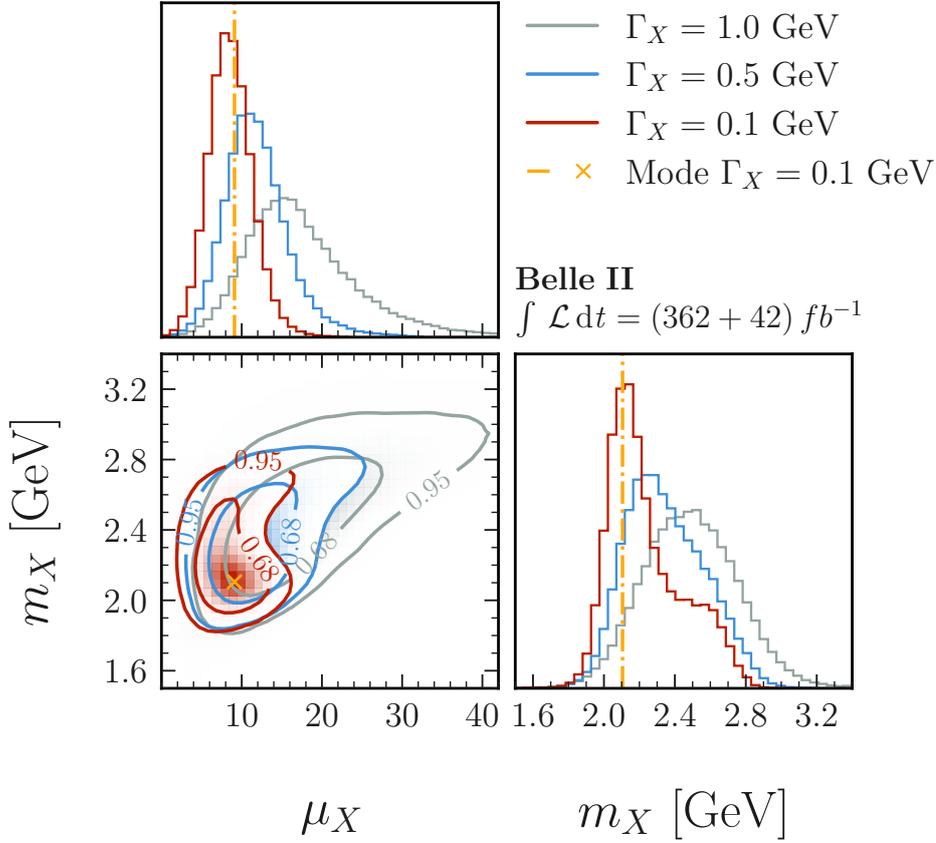


Figure 10.3: Marginalized posterior distributions for the $B^+ \rightarrow K^+ X$ model parameters from Equation (10.2), shown for the three resonance widths $\Gamma_X = 0.1, 0.5, 1.0$ GeV. Diagonal and off-diagonal panels show the 1-dimensional and 2-dimensional sample density PDFs on a linear scale, respectively. The overall scale is omitted, as all relevant information is contained in the shape of the distribution. The contours indicate 68% and 95% credible intervals. The dash-dotted yellow lines and yellow cross indicate the posterior mode.

independent chains enable convergence diagnostics through inter-chain comparison (see Appendix B.10). Convergence diagnostics are reported in Appendix F.1, confirming that the chains have converged to the target posterior distribution and that the sampling is reliable.

The full set of 1- and 2-dimensional marginal posterior distributions is shown in Figure 10.3 for all three widths. The contours enclose 68% and 95% credible regions.

10.2.3 Credible intervals and posterior mode

From the 1-dimensional marginal posterior distributions, I calculate the HDIs (see Section 5.4.1) at 68% and 95% probability for the model parameters. The posterior mode and the credible intervals are summarized in Table 10.2.

Table 10.2: Posterior mode and HDIs at 68% and 95% credible levels for the $B^+ \rightarrow K^+ X$ model parameters from Equation (10.2), derived from the posterior distributions in Figure 10.3. The new physics branching ratio is $\mathcal{B}(B^+ \rightarrow K^+ X) = \mu_X \cdot 10^{-6}$. Results are shown for three choices of resonance width.

| Param. | Γ_X [GeV] | Mode | 68% HDI | 95% HDI |
|-------------|------------------|------|-------------|-------------|
| μ_X | 0.1 | 9.2 | [5.8, 11.0] | [3.4, 14.0] |
| | 0.5 | 11.1 | [7.8, 15.0] | [4.2, 20.2] |
| | 1.0 | 15.3 | [9.6, 21.8] | [5.5, 32.1] |
| m_X [GeV] | 0.1 | 2.1 | [2.0, 2.3] | [1.9, 2.7] |
| | 0.5 | 2.2 | [2.1, 2.5] | [2.0, 2.8] |
| | 1.0 | 2.4 | [2.2, 2.7] | [2.0, 3.0] |

Direct comparison of the observed data with the predicted yields at the posterior mode (see Table 10.2) for the unconstrained $B^+ \rightarrow K^+ \nu \bar{\nu}$ SM and $B^+ \rightarrow K^+ X$ ($\Gamma_X = 0.1$ GeV) is shown in Figure 9.2, for the HSR of the analysis (ITA, $\eta(\text{BDT}_2) > 0.98$). The $B^+ \rightarrow K^+ X$ model provides a significantly better fit to the data than the unconstrained SM prediction, as indicated by the smaller pull values.

10.2.4 Physical interpretation of the posterior

For all three widths, the mass parameter m_X has a sharp peak in the range 2.1 – 2.4 GeV, with an extended tail toward higher masses. The peak location is consistent across the different width models, indicating that the data prefer a resonance in this mass range regardless of the assumed width. This result agrees well with the region where an excess in data were observed in Reference [2] and Figure 8.3.

These findings are consistent with the analyses in References [112, 113], which identified best-fit masses around $m_X \approx 2$ GeV using simplified treatments of the experimental data.

The new physics scale parameter μ_X shows a clear deviation from zero for all widths, indicating that the data prefer a model with a non-zero $B^+ \rightarrow K^+ X$ contribution. The posterior mode increases with resonance width: $\mu_X = 9.2$ for $\Gamma_X = 0.1$ GeV, $\mu_X = 11.1$ for $\Gamma_X = 0.5$ GeV, and $\mu_X = 15.3$ for $\Gamma_X = 1.0$ GeV. This is expected because broader resonances distribute the same total rate over a larger q^2 range, requiring a larger overall normalization to match the observed excess in any given bin.

Translating to branching ratios, the narrowest resonance ($\Gamma_X = 0.1$ GeV) yields $\mathcal{B}(B^+ \rightarrow K^+ X) = 9.2 \cdot 10^{-6}$ at the posterior mode, with a 95% credible interval of $[3.4, 14.0] \cdot 10^{-6}$. For comparison, the SM prediction for the total $B^+ \rightarrow K^+ \nu \bar{\nu}$ branching ratio is $\mathcal{B}_{\text{SM}} = (4.97 \pm 0.37) \cdot 10^{-6}$ [36]. The preferred $B^+ \rightarrow K^+ X$ branching ratio is thus roughly twice the SM expectation.

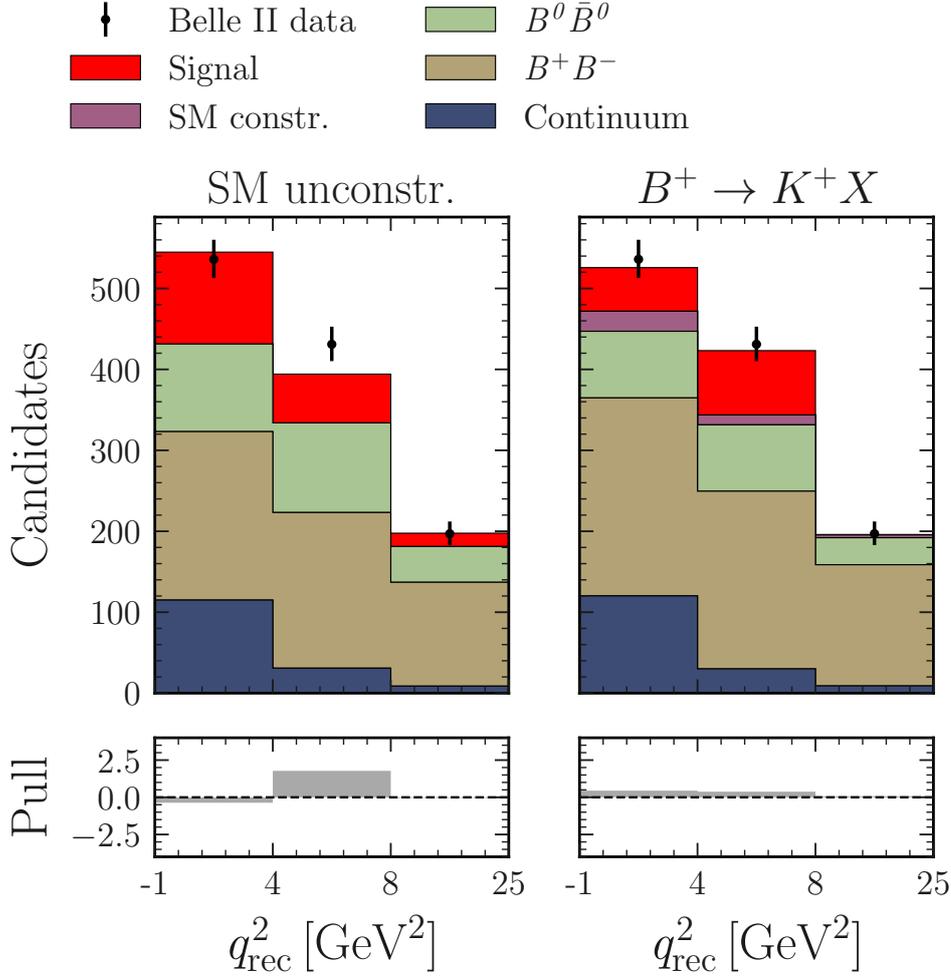


Figure 10.4: Observed and predicted best-fit yields in the HSR of the analysis. The signal is shown for the unconstrained $B^+ \rightarrow K^+\nu\bar{\nu}$ SM (left) and the $B^+ \rightarrow K^+ X$ ($\Gamma_X = 0.1$ GeV) (right) predictions. In the latter case, the constrained $B^+ \rightarrow K^+\nu\bar{\nu}$ SM is separated from the $B^+ \rightarrow K^+ X$ contribution. The predicted background yields are shown individually for the neutral and charged B -meson decays, and the summed five continuum categories. Pulls are shown in the lower panels.

10.2.5 Prior sensitivity

Bayesian inference inherently depends on the choice of prior distributions. To assess the robustness of the presented results and quantify the influence of prior choices, I perform a prior sensitivity analysis by comparing the baseline uniform priors to two physically motivated alternative prior specifications.

For simplicity, I perform the sensitivity study for the $B \rightarrow KX$ model in Equation (10.2) for a width of $\Gamma_X = 0.1$ GeV only.

The first alternative employs a *truncated-normal prior* for μ_X and a uniform prior for m_X (as in Table 10.1),

$$p(\mu_X) = \begin{cases} \mathcal{N}(\mu_X | \mu = 0, \sigma = 20) & \mu_X \geq 0 \\ 0 & \mu_X < 0 \end{cases}, \quad (10.4)$$

$$p(m_X) = \mathcal{U}([1.5, 3.0] \text{ GeV}).$$

Here \mathcal{U} corresponds to a uniform distribution.

The second alternative adopts a uniform prior for μ_X (as in Table 10.1) and a *uniform prior in the squared mass* m_X^2 , translating to a triangular prior in the mass,

$$p(\mu_X) = \mathcal{U}([0, 24]),$$

$$p(m_X) \propto \begin{cases} m_X & m_X \leq 4.8 \text{ GeV} \\ 0 & m_X > 4.8 \text{ GeV} \end{cases}. \quad (10.5)$$

The resulting posterior modes and credible intervals for both alternative prior choices are summarized in Table 10.3. The comparison reveals that the mass m_X is most robustly determined, with the posterior mode ($m_X = 2.1$ GeV for all priors) and credible intervals showing minimal sensitivity to the prior choice. The new physics scale μ_X exhibits a small shift in the posterior mode from 9.2 (uniform prior) to 9.0 (truncated-normal and triangular priors), and the credible intervals demonstrate similar robustness against changes in the priors.

10.3 Frequentist mass scan

This section presents a frequentist scan over the resonance mass, deriving 95% confidence-level upper limits on the branching ratio $\mathcal{B}(B^+ \rightarrow K^+ X)$ as a function of m_X .

For each fixed value of m_X , I determine the 95% confidence-level upper limit on μ_X using the modified frequentist test statistic for setting upper limits [79]:

$$q(\mu_X) = \begin{cases} t(\mu_X) & \hat{\mu}_X \leq \mu_X, \\ 0 & \hat{\mu}_X > \mu_X, \end{cases} \quad (10.6)$$

Table 10.3: The posterior modes, and HDIs at 68% and 95% for the $B^+ \rightarrow K^+ X$ model in Equation (10.2), for alternative priors (cf. Table 10.2). These results are based on the $\Gamma_X = 0.1$ GeV model.

| Priors | Param. | Mode | 68% HDI | 95% HDI |
|-----------------|-------------|------|-------------|-------------|
| Equation (10.4) | μ_X | 9.0 | [5.7, 10.8] | [3.4, 13.8] |
| | m_X [GeV] | 2.1 | [2.0, 2.4] | [1.9, 2.7] |
| Equation (10.5) | μ_X | 9.0 | [5.7, 11.0] | [3.0, 14.4] |
| | m_X [GeV] | 2.1 | [2.0, 2.4] | [1.9, 2.7] |

where $t(\mu_X)$ is the likelihood ratio test statistic defined in Equation (5.20). This test statistic excludes values of μ_X below the best-fit value $\hat{\mu}_X$ from contributing. This modification is appropriate for upper limit setting because observing a signal strength larger than the hypothesized value should not be considered evidence against the hypothesis [79].

The confidence limit is determined using the CL_s criterion of Equation (5.14) [84],

$$CL_s = \frac{P(q_{\text{obs}} | s + b)}{P(q_{\text{obs}} | b)}. \quad (10.7)$$

The CL_s method is more conservative than using $P(q_{\text{obs}} | s + b)$ alone. The 95% confidence-level upper limit is defined as the value of μ_X for which $CL_s = 0.05$.

The resulting 95% confidence-level upper limit on $\mathcal{B}(B^+ \rightarrow K^+ X)$ as a function of m_X is shown in Figure 10.5 for the narrow resonance case $\Gamma_X = 0.1$ GeV.

A pronounced excess of the observed limit relative to the expected limit is found in the mass range $m_X \in [1.5, 2.9]$ GeV. In this region, the observed limit lies significantly above the expected limit. The excess is most prominent around $m_X \sim 2.1$ GeV. The location and extent of the excess region are consistent with the 95% credible intervals on m_X reported in Table 10.2.

For $m_X > 3.5$ GeV, both the observed and expected limits rise steeply, indicating reduced sensitivity. This loss of sensitivity arises because the analysis efficiency decreases rapidly at high q^2 (see Figure 8.1).

The observed limit exhibits a localized increase near $m_X \sim 2$ GeV and a change in slope around $m_X \sim 3$ GeV. The location of these features correspond to the analysis bin boundaries in q_{rec}^2 . This behaviour is expected, as the sensitivity is reduced when the new physics contribution is distributed across two bins rather than fully contained within a single bin.

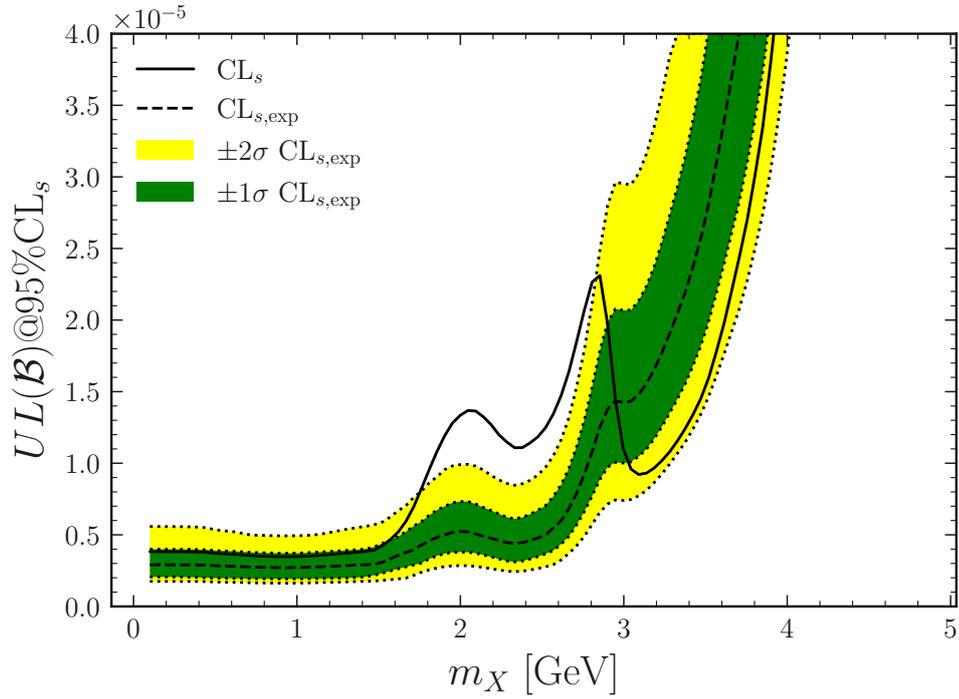


Figure 10.5: Frequentist 95% confidence-level upper limit on $\mathcal{B}(B^+ \rightarrow K^+ X)$ as a function of resonance mass m_X for $\Gamma_X = 0.1$ GeV, determined using the CL_s criterion with $CL_s \leq 0.05$. The solid black line shows the observed upper limit from data. The dashed black line shows the expected limit under the **SM** hypothesis, with green and yellow bands indicating the $\pm 1\sigma$ and $\pm 2\sigma$ ranges, respectively.

10.3.1 Injection study validation

To validate the reliability of the mass scan results and quantify the expected performance in different regions of parameter space, I perform an injection study. In this study, toy data are generated from the $B^+ \rightarrow K^+ X$ model PDF at specific parameter points (μ_X, m_X) , and fitting is applied to recover these injected values. The results are presented in detail in [Appendix F.2](#).

The injection study reveals that parameter recovery is good in the well-constrained mass range $1.0 \text{ GeV} < m_X < 2.5 \text{ GeV}$. Here the fitted values recover the injected truth within statistical uncertainties.

For $m_X > 3 \text{ GeV}$, the reduced sensitivity makes parameter determination challenging. In addition, the analysis employs coarse binning in q_{rec}^2 with boundaries at $[-1, 4, 8, 25] \text{ GeV}^2$. This binning is less ideal for resolving narrow resonances.

10.4 Goodness-of-fit

Beyond determining credible intervals and confidence limits, it is important to validate that the $B^+ \rightarrow K^+ X$ hypothesis provides a good description of the observed data. The goodness-of-fit P -values are evaluated analogously to the WET model in [Equation \(9.9\)](#), using a toy study with fluctuated data.

The distribution of the goodness-of-fit test statistic under the $B^+ \rightarrow K^+ X$ hypothesis is determined from 10 000 toy experiments. The distributions of t_{gof} for the three resonance widths are shown in [Figure 10.6](#). Under standard asymptotic assumptions, t_{gof} should follow a χ^2 -distribution with $N_{\text{dof}} = 30 - 2 = 28 \text{ d.o.f.}$, where $N_{\text{bins}} = 30$ is the total number of bins in the analysis and $N_{\text{POI}} = 2$ is the number of free POIs (μ_X and m_X). This yields an expectation of $N_{\text{dof}} = 28$. Fits to the distributions from toys yield $N_{\text{dof}} = 27.8 \pm 0.1$ for all three widths, showing a small deviation but overall good agreement with the expectation. The present deviation arises likely due to slightly non-Gaussian distributions of the POIs.

The goodness-of-fit P -values for all three resonance widths are summarized in [Table 10.4](#). These high P -values indicate excellent agreement between the $B^+ \rightarrow K^+ X$ models and the observed data — about 80% of toy experiments under the $B^+ \rightarrow K^+ X$ hypothesis would yield worse fits than what is observed in the real data.

The consistency of the P -values across different widths demonstrates that the fit quality is robust. This insensitivity to Γ_X arises because of the coarse binning of the analysis (q_{rec}^2 bins of $[-1, 4, 8, 25] \text{ GeV}^2$). As long as resonances fall in the same bin, they produce a similar distribution in q_{rec}^2 .

Comparing the $B^+ \rightarrow K^+ X$ goodness-of-fit results to those from the WET analysis in [Section 9.3](#), the WET model yielded $P_{\text{gof}} = 0.63$, and the unconstrained SM yielded $P_{\text{gof}} = 0.58$. The superior goodness-of-fit for the $B^+ \rightarrow K^+ X$ models indicates that the resonance hypothesis provides a better explanation for the observed data. This advantage

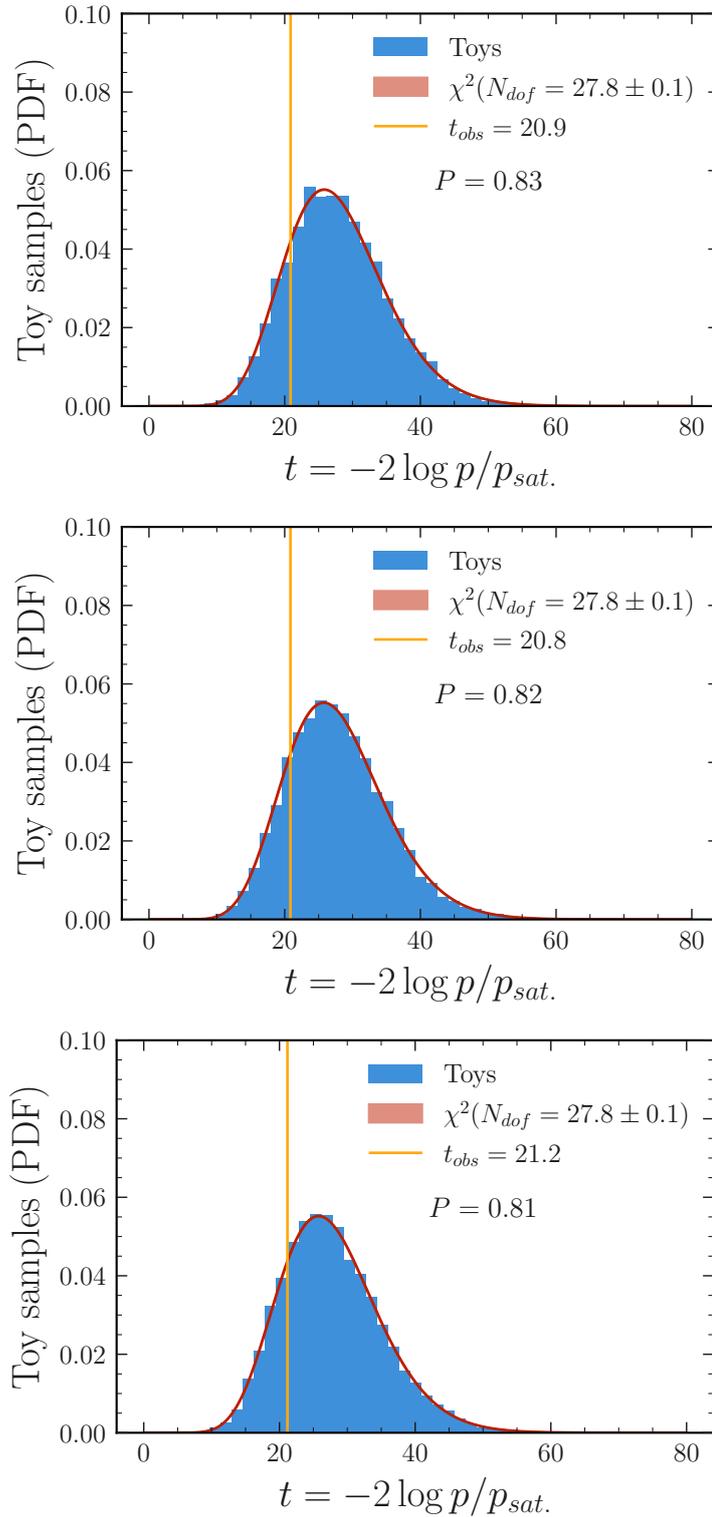


Figure 10.6: The distribution of the goodness-of-fit test statistic for the $B^+ \rightarrow K^+ X$ model with resonance widths $\Gamma_X = 0.1, 0.5, 1.0$ GeV (*top/middle/bottom*), each obtained from 10 000 toys. The values obtained from a fit to the data are shown as vertical yellow lines. The asymptotic χ^2 -distributions (red band) correspond to the best-fit N_{dof} , which are shown with the corresponding uncertainty band.

arises because the $B^+ \rightarrow K^+ X$ model fits well to the localized data excess in the q_{rec}^2 distribution around $3 - 7 \text{ GeV}^2$ by construction.

10.5 Model comparison

Model comparison quantifies the *relative* support that the data provide for competing hypotheses. I perform both Bayesian and frequentist model comparisons to assess the relative performance of the $B^+ \rightarrow K^+ X$ hypothesis against two reference models: the background-only hypothesis (BKG), the constrained $B^+ \rightarrow K^+ \nu \bar{\nu}$ SM, where μ is fixed to the SM prediction.

10.5.1 Bayes factors

Bayesian model comparison employs the Bayes factor, defined as the ratio of marginal likelihoods between two competing models (see Section 5.2.2).

The Bayes factors are computed using sequential MC sampling methods [80, 128], which provide numerical estimates of the marginal likelihoods for each model. The Bayes factors for all three resonance widths are summarized in Table 10.4.

Comparing to the constrained SM, the Bayes factors are $\log_{10} B_{\text{SM}}^{\text{constr.}} = 1.71$ for $\Gamma_X = 0.1 \text{ GeV}$, 1.83 for $\Gamma_X = 0.5 \text{ GeV}$, and 1.82 for $\Gamma_X = 1.0 \text{ GeV}$. These values all exceed Jeffreys's threshold [87] of $\log_{10} B > 1.5$ for a *very strong* model preference (see Table 5.1).

Comparing to the background-only hypothesis, the Bayes factors are $\log_{10} B_{\text{BKG}} = 2.82$ for $\Gamma_X = 0.1 \text{ GeV}$, 2.93 for $\Gamma_X = 0.5 \text{ GeV}$, and 2.92 for $\Gamma_X = 1.0 \text{ GeV}$. These values all exceed the threshold of $\log_{10} B > 2.0$ for a *decisive* model preference.

Comparing the Bayes factors across different new physics hypotheses (see Table 9.4 for WET and unconstrained SM results), the $B^+ \rightarrow K^+ X$ models show the strongest overall preference. The WET model yields $\log_{10} B_{\text{BKG}} = 1.78$, and the unconstrained SM yields $\log_{10} B_{\text{BKG}} = 2.02$, both lower than the $B^+ \rightarrow K^+ X$ results. This hierarchy reflects, that the excess in data around $q_{\text{rec}}^2 \sim 3 - 7 \text{ GeV}^2$ is better described by a resonance than by smooth modifications to the kinematic distribution.

The Bayes factor inherently accounts for the look-elsewhere effect through the integration over the full parameter space in the marginal likelihood (see Equation (5.27)). The look-elsewhere effect is the reduction of global significance compared to local significance, arising because testing many possible mass values increases the chance of observing a random fluctuation that looks signal-like [129]. The strong Bayes factors obtained here thus represent global evidence for the $B^+ \rightarrow K^+ X$ hypothesis, not just local significance at the best-fit mass.

Table 10.4: Bayes factors and goodness-of-fit P -values for the $B^+ \rightarrow K^+ X$ models with three different resonance widths. The Bayes factors are reported relative to the background-only hypothesis (B_{BKG}) and the constrained $B^+ \rightarrow K^+ \nu \bar{\nu}$ SM hypothesis ($B_{\text{SM}}^{\text{constr.}}$).

| Model | Γ_X | $\log_{10} B_{\text{BKG}}$ | $\log_{10} B_{\text{SM}}^{\text{constr.}}$ | P_{gof} |
|-------------------------|------------|----------------------------|--|------------------|
| $B^+ \rightarrow K^+ X$ | 0.1 GeV | 2.82 | 1.71 | 0.83 |
| $B^+ \rightarrow K^+ X$ | 0.5 GeV | 2.93 | 1.83 | 0.82 |
| $B^+ \rightarrow K^+ X$ | 1.0 GeV | 2.92 | 1.82 | 0.81 |

10.5.2 Frequentist hypothesis test

To complement the Bayesian model comparison and provide a frequentist perspective, I perform hypothesis tests that quantify the significance of the signal hypotheses relative to the background-only hypothesis.

The tests were conducted using the P -value defined in Equation (9.10). The distribution of the test statistic $p(t)$ is determined through toy experiments.

For the $B^+ \rightarrow K^+ X$ model with $\Gamma_X = 0.1$ GeV, the hypothesis test yields a P -value of $P = 1.4 \cdot 10^{-3}$, corresponding to a significance of $Z = 3.0\sigma$ with respect to the constrained SM hypothesis. The distribution of the test statistic from 100 000 toy experiments is shown in Figure 10.7, where the observed value $t_{\text{obs}} = 13.1$ falls in the far tail of the distribution.

This significance is notably higher than the $Z = 2.7\sigma$ quoted in the original Belle II publication [2] for the excess relative to the SM. The increased significance when testing the $B^+ \rightarrow K^+ X$ hypothesis reflects the fact that the shape of the excess in data matches the $B^+ \rightarrow K^+ X$ prediction better than a simple overall enhancement.

For a model with two free parameters (μ_X and m_X), the asymptotic expectation is a χ^2 distribution with $N_{\text{dof}} = 2$. Fitting to the toy distribution yields $N_{\text{dof}} = 1.646 \pm 0.004$, showing a statistically significant deviation from the asymptotic expectation. This potentially arises from boundary effects (m_X must be positive) introducing non-Gaussian features in the test statistic distribution. The obtained P -value is based on the toy distribution and hence reliable.

The hypothesis test as performed here does account for the look-elsewhere effect. The test statistic distribution is generated from toys sampled under the constrained $B^+ \rightarrow K^+ \nu \bar{\nu}$ SM hypothesis. Each toy is fitted allowing both μ_X and m_X to vary freely. Therefore, this procedure incorporates the penalty for searching over different mass values. In the toys, random fluctuations can mimic resonances at various masses. The test statistic distribution reflects how often such fluctuations occur when fitting with a free mass parameter.

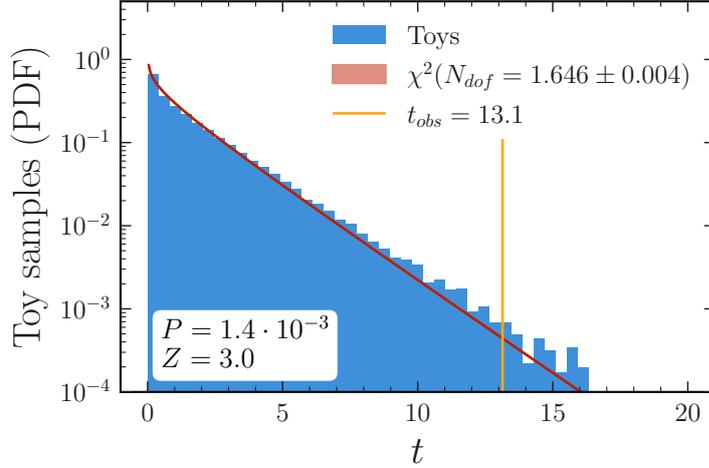


Figure 10.7: Distribution of the test statistic for the $B^+ \rightarrow K^+ X$ hypothesis ($\Gamma_X = 0.1$ GeV) versus the constrained **SM** hypothesis, obtained from 100 000 toy experiments. The value from the fit to data is shown as vertical yellow line. The asymptotic χ^2 -distribution (red band) corresponds to the best-fit N_{dof} , which is shown with the corresponding uncertainty band.

10.6 Discussion

This chapter presents a comprehensive reinterpretation of the Belle II $B^+ \rightarrow K^+ \nu \bar{\nu}$ measurement under the hypothesis of a light invisible resonance $B^+ \rightarrow K^+ X$. Using the model-agnostic likelihood method of [Chapter 7](#), I have derived Bayesian credible intervals and frequentist confidence limits for the mass and branching ratio of this potential contribution.

The principal findings are summarized in [Table 10.2](#). The data exhibit a strong preference for a resonance mass in the range of 2.1 – 2.4 GeV. This finding is approximately independent of the three tested widths ($\Gamma_X = 0.1, 0.5, 1.0$ GeV), with posterior modes at $m_X = 2.1, 2.2, 2.4$ GeV, respectively. The 95% credible intervals are [1.9, 2.7] GeV for $\Gamma_X = 0.1$ GeV, broadening slightly to [2.0, 3.0] GeV for $\Gamma_X = 1.0$ GeV. This preferred mass range corresponds to $q^2 = m_X^2 \sim 4 - 6$ GeV², in agreement with the localized excess observed in the Belle II data around $3 \text{ GeV}^2 < q_{\text{rec}}^2 < 7 \text{ GeV}^2$ [[2](#)].

For the narrowest resonance ($\Gamma_X = 0.1$ GeV), the branching ratio posterior mode is $\mathcal{B}(B^+ \rightarrow K^+ X) = 9.2 \cdot 10^{-6}$ with a 95% credible interval of [3.4, 14.0] $\cdot 10^{-6}$. This branching ratio is approximately twice the **SM** prediction for the total $B^+ \rightarrow K^+ \nu \bar{\nu}$ rate, $\mathcal{B}_{\text{SM}} = (4.97 \pm 0.37) \cdot 10^{-6}$ [[36](#)].

Both the mass and branching ratio results are consistent with the findings of [References \[112–114\]](#). These [References](#) identify best-fit masses around $m_X \approx 2$ GeV using simplified treatments of the Belle II data. The key advance of this analysis is the correct treatment of experimental uncertainties through the model-agnostic likelihood framework.

Previous studies necessarily rely on simplified assumptions, and neglected systematic uncertainties.

Goodness-of-fit tests and model comparisons support the reliability of these conclusions. All three $B^+ \rightarrow K^+ X$ models provide excellent fits to the data, with goodness-of-fit P -values of $P_{\text{gof}} \sim 0.81 - 0.83$. These values exceed those obtained for the **WET** model ($P_{\text{gof}} = 0.63$) and the unconstrained **SM** ($P_{\text{gof}} = 0.58$) presented in [Section 9.3](#). This indicates that the $B^+ \rightarrow K^+ X$ models fit the data significantly better.

Bayesian model comparison via Bayes factors shows *decisive* preference ($\log_{10} B_{\text{BKG}} = 2.82 - 2.93$) for the $B^+ \rightarrow K^+ X$ models over the background-only hypothesis, and a *very strong* preference ($\log_{10} B_{\text{SM}}^{\text{constr.}} = 1.71 - 1.83$) over the constrained **SM**. These Bayes factors are the largest observed among all tested hypotheses (**WET**, unconstrained **SM**, $B^+ \rightarrow K^+ X$).

The frequentist hypothesis test yields a significance of $Z = 3.0\sigma$ for the $B^+ \rightarrow K^+ X$ model ($\Gamma_X = 0.1$ GeV) against the constrained **SM** hypothesis. This represents a global significance that accounts for the look-elsewhere effect through the freedom to fit the resonance mass. The 3.0σ significance exceeds the 2.7σ reported in the original Belle II publication [2]. This reflects the improved fit quality when the kinematic structure of the excess is explicitly modelled.

However, despite the high statistical significance and the strong Bayes factor, caution is advised before claiming evidence for $B^+ \rightarrow K^+ X$. Two factors are the reason for this. First, the Belle II analysis was optimized for measuring the **SM** $B^+ \rightarrow K^+ \nu \bar{\nu}$ branching ratio, not for searching for narrow resonances. Second, the $B^+ \rightarrow K^+ X$ model was motivated by features observed in the unblinded data.¹ Therefore, the present findings should be regarded as motivation for dedicated resonance searches in future analyses rather than as evidence for new physics.

This comprehensive reinterpretation demonstrates that a light invisible resonance with $m_X \sim 2$ GeV and $\mathcal{B}(B^+ \rightarrow K^+ X) \sim 10^{-5}$ provides an excellent description of the Belle II $B^+ \rightarrow K^+ \nu \bar{\nu}$ data, significantly outperforming the **WET** model and the **SM**.

¹Particle physics analyses are designed using simulated data. Experimental data are examined only after the analysis procedure is finalized, to avoid bias.

Chapter 11

A combined analysis of $B \rightarrow K^{(*)}\nu\bar{\nu}$ decays at Belle II

¹ In the previous chapters, I have demonstrated how powerful reinterpretation can be. For reinterpretation in terms of the **WET**, the underlying theory only allows us to probe sums of Wilson coefficients rather than individual ones for $B \rightarrow K\nu\bar{\nu}$ decays. This is evident from [Equation \(2.62\)](#). I have also shown that the **WET** prediction for $B \rightarrow K^*\nu\bar{\nu}$ decays is sensitive to the same set of Wilson coefficients but has a different structure, as shown in [Equation \(2.65\)](#). Hence, a combined analysis of $B \rightarrow K\nu\bar{\nu}$ and $B \rightarrow K^*\nu\bar{\nu}$ enables probing the magnitude of all five contributing Wilson coefficients, as discussed in [Section 7.3.2](#).

In this chapter, I present the work-in-progress Belle II analysis of the four channels $B^+ \rightarrow K^+\nu\bar{\nu}$, $B^0 \rightarrow K_s^0\nu\bar{\nu}$, $B^0 \rightarrow K^{*0}\nu\bar{\nu}$, and $B^+ \rightarrow K^{*+}\nu\bar{\nu}$ in combination. This study is not only interesting as a comparison with the existing $B^+ \rightarrow K^+\nu\bar{\nu}$ analysis, but also enables very powerful reinterpretation studies, as shown in [Section 7.3.2](#).

This chapter is structured as follows: [Section 11.1](#) presents the theoretical predictions for all four decay channels. [Section 11.2](#) provides an overview of the analysis strategy, extending from the published $B^+ \rightarrow K^+\nu\bar{\nu}$ measurement to a simultaneous four-channel analysis. The background suppression techniques are described in [Section 11.3](#). The signal region definition is covered in [Section 11.4](#). [Section 11.5](#) describes the statistical framework for simultaneous signal extraction, including uncorrelated, isospin-averaged, and fully correlated scenarios. Finally, [Section 11.6](#) discusses the expected sensitivity improvements and implications for reinterpretation studies.

¹The material presented in this chapter is the result of collaborative work within the Belle II collaboration. Individual contributions of the proponents of this work and further collaboration members are documented in [Appendix G.1](#). My personal contributions are the theoretical predictions in [Section 11.1](#) and the statistical inference in [Section 11.5](#). Furthermore, this is a work-in-progress project, and the presented results are subject to change.

11.1 Theory predictions

As should be evident by now, theory predictions form the foundation of any analysis. This analysis is based on the SM prediction, which is obtained from Equations (2.62) and (2.65) at the SM Wilson coefficient point given in Equation (2.55).

The signal MC for the $B^+ \rightarrow K^+\nu\bar{\nu}$ and $B^0 \rightarrow K_s^0\nu\bar{\nu}$ channels is based on the hadronic parameters from Reference [42]. To base the analysis on the most recent findings, the samples are reweighted to hadronic parameters obtained from a joint theoretical prior PDF composed of the 2021 lattice world average, based on results by the Fermilab/MILC and HPQCD collaborations [45, 46], see Table 2.3. The reduction of uncertainties in the hadronic parameters is significant, as shown in Figure 11.1.

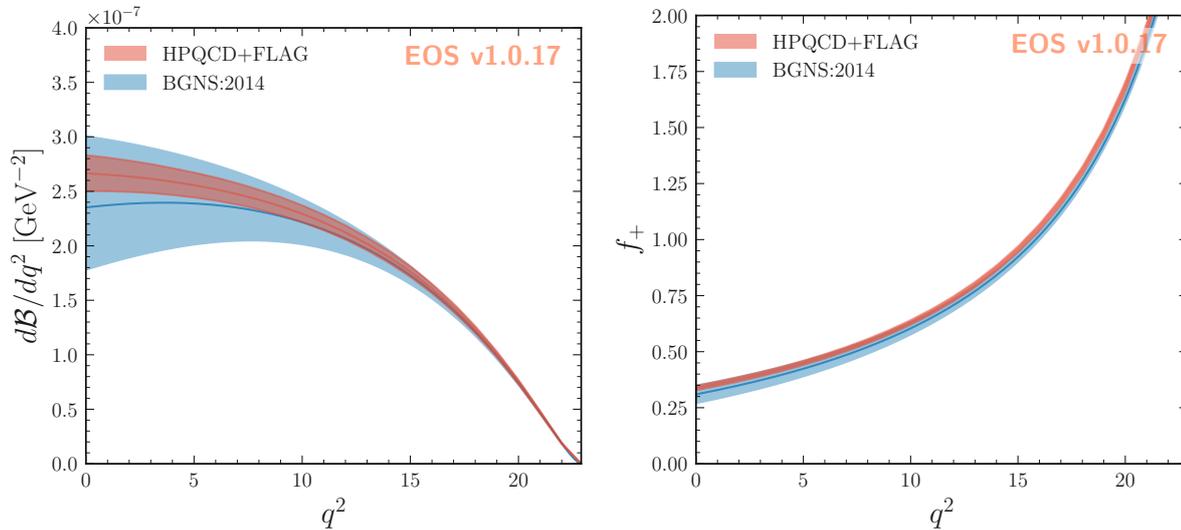


Figure 11.1: The $B \rightarrow K\nu\bar{\nu}$ SM differential branching ratio (left) and f_+ form factor (right), based on the hadronic parameters from Reference [42] (labelled as BGNS:2014) and from References [45, 46] (labelled as HPQCD+FLAG). The bands represent the uncertainty in the hadronic parameters.

The signal MC for the $B^0 \rightarrow K^{*0}\nu\bar{\nu}$ and $B^+ \rightarrow K^{*+}\nu\bar{\nu}$ channels is based on the hadronic parameters from Reference [48]. In this case, the samples are reweighted to hadronic parameters obtained from a joint theoretical prior PDF based on results from References [130, 131], see Table A.2. Here there is no improvement in the uncertainties, as shown in Figure 11.2. The result is nevertheless more reliable due to the combination of multiple sources.

The predictions are obtained using the EOS software, discussed in Section 2.4.3. The relevant input parameters are summarized in Table 2.2. The product $|V_{tb}V_{ts}^*|$ is evaluated using CKM unitarity from the value of $|V_{cb}|$, which can be determined using exclusive or inclusive methods [10]. The predictions reported here are based on the product

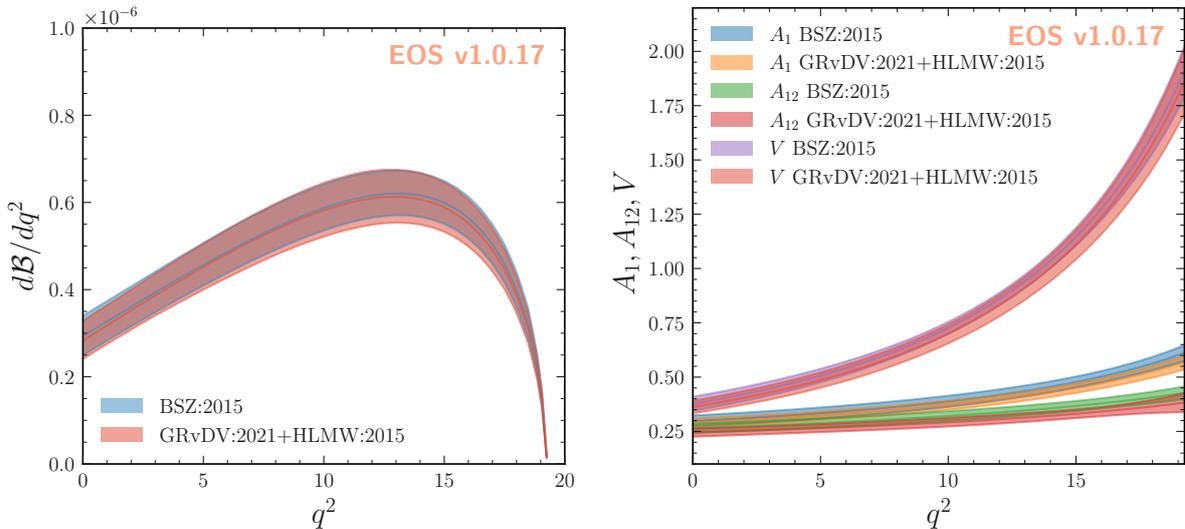


Figure 11.2: The $B \rightarrow K^* \nu \bar{\nu}$ SM differential branching ratio (left) and A_1, A_{12}, V form factors (right), based on the hadronic parameters from Reference [48] (labelled as BSZ:2015) and from References [130, 131] (labelled as GRvDV:2021+HLMW:2015). The bands represent the uncertainty in the hadronic parameters.

$$|V_{tb}V_{ts}^*| = 0.039 \pm 0.001 \text{ from exclusive modes [10, 43, 45].}$$

Theoretical uncertainties arise from two main sources: uncertainties in the CKM matrix elements and the hadronic form factors, with small additional contributions from other parametric dependencies and higher-order corrections. The discrepancies in the determinations of the matrix element $|V_{cb}|$ using exclusive or inclusive methods contribute a 6% uncertainty [10]. Predictions for decays involving a pseudoscalar in the final state contain a 4% form factor uncertainty, such that the CKM uncertainties dominate. Conversely, predictions for decays with a vector meson in the final state have a dominant 10% form factor uncertainty.

The SM predictions for the total branching ratios of the four decays are summarized in Table 11.1.

The long-distance double-charged current contribution $B \rightarrow (\tau \rightarrow K \nu_\tau) \bar{\nu}_\tau$ is treated as a background in this analysis. The contribution is normalized based on experimental measurements of $B^+ \rightarrow \tau^+ \nu_\tau$ and $\tau^+ \rightarrow K^{(*)+} \bar{\nu}_\tau$ decays [10].

11.2 Analysis outline

The analysis is designed as an extension of the published $B^+ \rightarrow K^+ \nu \bar{\nu}$ measurement described in Section 8.1, using the same Belle II data sample (see Section 8.1.1). The $B^+ \rightarrow K^+ \nu \bar{\nu}$ analysis itself is largely unchanged, with the selected data sample in the

Table 11.1: **SM** branching ratios for $B \rightarrow K\nu\bar{\nu}$ decays used in this work (as described in the text), compared to predictions from Reference [43]. The uncertainties are separated into contributions from hadronic form factors (FF) and **CKM** elements (CKM). Note that Reference [43] uses slightly different input parameters. In addition, it assumes that $B \rightarrow K^{*}\nu\bar{\nu}$ form factor parameters are uncorrelated, leading to larger form factor uncertainties.

| Decay | $\mathcal{B}_{\text{SM}} \cdot 10^6$ | |
|--------------------------------------|---|---|
| | This work | Reference [43] |
| $B^+ \rightarrow K^+\nu\bar{\nu}$ | $4.32 \pm 0.16(\text{FF}) \pm 0.26(\text{CKM})$ | $4.37 \pm 0.15(\text{FF}) \pm 0.26(\text{CKM})$ |
| $B^0 \rightarrow K_s^0\nu\bar{\nu}$ | $2.00 \pm 0.07(\text{FF}) \pm 0.12(\text{CKM})$ | $2.02 \pm 0.07(\text{FF}) \pm 0.12(\text{CKM})$ |
| $B^+ \rightarrow K^{*+}\nu\bar{\nu}$ | $9.35 \pm 0.94(\text{FF}) \pm 0.56(\text{CKM})$ | $9.73 \pm 1.37(\text{FF}) \pm 0.58(\text{CKM})$ |
| $B^0 \rightarrow K^{*0}\nu\bar{\nu}$ | $8.65 \pm 0.87(\text{FF}) \pm 0.57(\text{CKM})$ | $8.97 \pm 1.22(\text{FF}) \pm 0.54(\text{CKM})$ |

signal region left unmodified, and the simulated samples also remain unchanged. However, the relative contributions of the leading backgrounds are updated to reflect the latest knowledge, ensuring consistency with the other channels. In contrast to the published version, where only the uncertainties were reweighted, both the central values and uncertainties of the background contributions are now updated. This leads to an overall reduction in background and reduces the expected uncertainty for the $B^+ \rightarrow K^+\nu\bar{\nu}$ channel compared to the published result. The signal-extraction machinery is updated to determine the four channels of interest simultaneously, accounting for their mutual cross-influences. Additional updates include newly developed control channels.

Conversely, the measurements of $B^0 \rightarrow K_s^0\nu\bar{\nu}$, $B^+ \rightarrow K^{*+}\nu\bar{\nu}$, and $B^0 \rightarrow K^{*0}\nu\bar{\nu}$ are designed to follow the $B^+ \rightarrow K^+\nu\bar{\nu}$ analysis (see Section 8.1) as closely as possible, unless deviations provide significant benefits. The main update is the use of improved simulated samples that incorporate dedicated data to model beam-background activity. This enables the inclusion of additional variables to discriminate signal from background and to fine-tune the reconstruction. It was verified that these updates have only a minor impact on the $B^+ \rightarrow K^+\nu\bar{\nu}$ measurement.

The analysis begins with the reconstruction of charged and neutral particles, followed by the selection of signal kaon candidates in events with one or more K mesons. Subsequently, relevant quantities to discriminate between signal and background processes are computed using the kaon candidates together with the remaining particles in the event. These quantities are used in **BDTs**, which are optimized and trained on simulated samples. A signal region is then defined, and a single K -meson candidate is selected.

Up to this point, the analysis is performed for each channel individually, while accounting for correlated systematic uncertainties. With the signal-region samples prepared for all four channels, a simultaneous binned profile-likelihood sample-composition fit is performed on data using a **HistFactory** statistical model (Section 4.4). This fit uses simulated samples to provide predictions and determines the branching ratios of the four

decay modes along with the rates of background processes. The fit incorporates systematic uncertainties arising from detector and physics-modelling imperfections as nuisance parameters. To validate the modelling of signal and background processes in simulation, several control channels are employed.

11.3 Background suppression

Simulated signal and background events are used to train **BDTs** that suppress background. Several inputs are considered, including general event-shape variables as well as variables characterizing the signal candidates and the kinematic properties of the **ROE**.

This analysis uses two consecutive **BDTs**, as described in [Section 8.1.2](#). The BDT_1 classifier is constructed using the same input variables for the $B^0 \rightarrow K_s^0 \nu \bar{\nu}$ and $B^0 \rightarrow K^{*0} \nu \bar{\nu}$ channels as the published $B^+ \rightarrow K^+ \nu \bar{\nu}$ channel. For the $B^+ \rightarrow K^{*+} \nu \bar{\nu}$ channel, the classifier uses a different set of 11 variables, which additionally contains kinematic properties of the K^{*+} decay products. In the $B^0 \rightarrow K_s^0 \nu \bar{\nu}$ and $B^0 \rightarrow K^{*0} \nu \bar{\nu}$ channels, the most powerful discriminant is the difference between the **ROE** energy in the centre-of-mass frame and half the total collision energy, ΔE_{ROE} . For the $B^+ \rightarrow K^{*+} \nu \bar{\nu}$ channel, the modified Fox-Wolfram moments [[132](#)] provide the strongest separation power.

The second classifier, BDT_2 , is used for the final event selection, which is trained on events with $\text{BDT}_1 > 0.9$. For the $B^0 \rightarrow K_s^0 \nu \bar{\nu}$ channel, the most discriminating variable in BDT_2 is the cosine of the angle between the momentum of the signal K_s^0 candidate and its vertex displacement vector. For the $B^0 \rightarrow K^{*0} \nu \bar{\nu}$ channel, the most powerful variable is the invariant mass of the $K^+ \pi^-$ pair of the K^{*0} candidate. For the $B^+ \rightarrow K^{*+} \nu \bar{\nu}$ channel, the most discriminating variable is ΔE_{ROE} .

11.4 Signal region definition

Using the simulated signal sample, the BDT_2 variable is mapped to the complement of the integrated signal-selection efficiency, as in [Equation \(8.2\)](#). The Signal Region (**SR**) is defined by requiring $\text{BDT}_1 > 0.9$ for all channels. In addition, channel-dependent requirements are applied to $\eta(\text{BDT}_2)$: $\eta(\text{BDT}_2) > 0.92$ for $B^0 \rightarrow K_s^0 \nu \bar{\nu}$, $\eta(\text{BDT}_2) > 0.95$ for $B^0 \rightarrow K^{*0} \nu \bar{\nu}$, and $\eta(\text{BDT}_2) > 0.97$ for $B^+ \rightarrow K^{*+} \nu \bar{\nu}$.

The **SR** is further divided into 4×3 , 5×3 and 6×3 bins in the $\eta(\text{BDT}_2) \times q_{\text{rec}}^2$ space for the channels $B^0 \rightarrow K_s^0 \nu \bar{\nu}$, $B^0 \rightarrow K^{*0} \nu \bar{\nu}$ and $B^+ \rightarrow K^{*+} \nu \bar{\nu}$, respectively. The q_{rec}^2 bin boundaries are the same for all channels and are given by $[-1.0, 4.0, 8.0, 25.0]$ GeV^2 . The bin boundaries in q_{rec}^2 are chosen to follow the theoretical predictions from Reference [[42](#)] while ensuring a sufficient number of expected signal events in each bin.

For the $B^0 \rightarrow K_s^0 \nu \bar{\nu}$ channel, the bin boundaries in $\eta(\text{BDT}_2)$ are $[0.92, 0.94, 0.96, 0.98, 1.00]$. For the $B^0 \rightarrow K^{*0} \nu \bar{\nu}$ channel, the bin boundaries in $\eta(\text{BDT}_2)$ are $[0.95, 0.96, 0.97, 0.98, 0.99, 1.00]$. For the $B^+ \rightarrow K^{*+} \nu \bar{\nu}$ channel, the bin boundaries in $\eta(\text{BDT}_2)$ are $[0.97, 0.975, 0.98, 0.985, 0.99, 0.995, 1.00]$. The binning choice

and the lower boundary for $\eta(\text{BDT}_2)$ are optimized by minimizing the expected statistical uncertainty on the **POI**. The last bin of $\eta(\text{BDT}_2)$ provides the dominant sensitivity to the signal and is therefore referred to as the **HSR**.

The signal-selection efficiency in the **SR** for the $B^0 \rightarrow K_s^0\nu\bar{\nu}$, $B^+ \rightarrow K^{*+}\nu\bar{\nu}$, and $B^0 \rightarrow K^{*0}\nu\bar{\nu}$ decays is shown in **Figure 11.3**. The efficiency varies significantly as a function of the kinematic variables q^2 and the helicity angle $\cos\theta$, which is defined as the angle between the K^* boost vector and the kaon direction in the K^* rest frame.

11.5 Signal extraction

A **HistFactory** statistical model is constructed to extract the signal strengths of the four channels $B^+ \rightarrow K^+\nu\bar{\nu}$, $B^+ \rightarrow K^{*+}\nu\bar{\nu}$, $B^0 \rightarrow K_s^0\nu\bar{\nu}$, and $B^0 \rightarrow K^{*0}\nu\bar{\nu}$ simultaneously. Inference is performed with **pyhf** (**Section 4.4.3**), across the four channels, using both on-resonance $\Upsilon(4S)$ data and off-resonance data to constrain the continuum background.

The systematic uncertainties affect normalizations and may also influence the shape of the templates. The three dominant sources are:

- The yields of the seven individual background categories are allowed to vary independently in the fit. In each case, a Gaussian constraint (**normsys** modifier, see **Section 4.4.1**) is added to the fit, centred at the expectation based on simulation and with a standard deviation corresponding to 50% of the central value. The conservative 50% value is motivated by a global normalization difference between the off-resonance data and continuum simulation for the $B^+ \rightarrow K^+\nu\bar{\nu}$ channel. The uncertainty is modelled by using seven nuisance parameters for each of the four channels leading to in total 28 parameters in the fit.
- The systematic uncertainty due to the limited size of simulated samples (**statererror** modifier, see **Section 4.4.1**), which also includes small contributions of uncorrelated uncertainty arising from covariance matrix decomposition for **PID**, K_s^0 , and π^0 efficiency is modelled by an independent nuisance parameter for each analysis bin (114 nuisance parameters in total). This is the second-largest uncertainty of the analysis, ranging between 0.5 and 1.2.
- The branching ratios of decay modes contributing to more than 80% of the background B decays in the **SR** are allowed to vary according to their known uncertainties [10]. They are described by 113 nuisance parameters (**histosys** modifier, see **Section 4.4.1**).

To account for all systematic sources, including the ones above, but also smaller contributions, a total of 338 nuisance parameters, along with the four signal strengths μ , are varied in the fit.

The fitting proceeds in two steps: A first fit is performed using **scipy** [76] as the minimizer. Then, a second fit is performed using **Minuit (Migrad)** [77] as the minimizer with the result of the first fit as the starting point. The covariance matrices for the **POIs** and

nuisance parameters are obtained using the `Hesse` algorithm. This two-step approach significantly increases the robustness of the fit.

The fit has been performed on the Asimov dataset, defined as the dataset with the number of observed events equivalent to the number of expected events for every bin. The Asimov dataset is obtained as the sum of all process templates (signals and backgrounds). The information gained from this procedure consists of estimates of the uncertainties and correlations of the fit parameters. The post-fit distributions are shown in [Figures 11.4](#) and [11.5](#).

The fit results in terms of uncertainty on the signal strength multiplier, σ_μ , and uncertainty on the SM branching ratio ($\sigma_B = \sigma_\mu \mathcal{B}_{SM}$) are shown in [Table 11.2](#). The world's best results [[133](#), [134](#)] are also reported for reference (obtained by Belle with a semileptonic tag on the full dataset).² The expected sensitivity is comparable to the world's best results in all channels.

Table 11.2: Fit uncertainties on POIs in terms of signal strength μ and branching ratio \mathcal{B} . The world's best results are taken from Reference [[133](#), [134](#)] (The Belle II result from Reference [[2](#)] is excluded).

| Channel | σ_μ | $\sigma_B \cdot 10^{-6}$ | World's best $\sigma_B \cdot 10^{-6}$ | $\pm \sigma_\mu$ profiled |
|--|--------------|--------------------------|--|---------------------------|
| $B^+ \rightarrow K^+ \nu \bar{\nu}$ | 1.23 | 5.46 | 5.7 | + 1.24 - 1.24 |
| $B^0 \rightarrow K_s^0 \nu \bar{\nu}$ | 2.75 | 5.64 | 6.5 | + 2.90 - 2.82 |
| $B^+ \rightarrow K^{*+} \nu \bar{\nu}$ | 2.17 | 21.24 | 17 | + 2.26 - 2.18 |
| $B^0 \rightarrow K^{*0} \nu \bar{\nu}$ | 1.49 | 13.48 | 11 | + 1.55 - 1.48 |

As a test for sensitivity and symmetry of the parameter distributions, 1-dimensional profiled likelihood scans for each of the four signal strengths are performed on Asimov data. The results are shown in [Figure 11.6](#). The negative-log-likelihood curve is highly symmetric in all cases, which indicates expected asymptotic behaviour. From the profiled likelihood scan, the parameter uncertainties can be calculated through the relative change in the value of the negative log likelihood compared to the best-fit point [[65](#)],

$$-2 \ln L(\hat{\mu} \pm \sigma_\mu) + 2 \ln L(\hat{\mu}) = 1. \quad (11.1)$$

The resulting profiled uncertainties are shown in [Table 11.2](#).

In [Figure 11.7](#), the correlation matrix of the Asimov fit is shown. Most importantly, there is only a small correlation between the four signal strengths. The largest correlation between the signal and background components is found for the $B^+ B^-$ background in the $B^+ \rightarrow K^{*+} \nu \bar{\nu}$ channel.

²Numerical branching ratio uncertainties are only found in the Belle internal note [[134](#)].

To investigate by how much a particular systematic source increases the uncertainty on μ , this particular source is removed from the fit, and the uncertainty on μ of the modified fit is subtracted from that of the nominal fit in quadrature. The results are reported in [Table 11.3](#). The background normalization uncertainties have the largest contribution to σ_μ . Overall, the normalizations of $B\bar{B}$ backgrounds play a more important role. For the $B^0 \rightarrow K_S^0\nu\bar{\nu}$ decay, in contrast, the normalization uncertainty of the continuum background is more important. The systematic uncertainty due to the limited size of simulated samples is the second-largest uncertainty of the analysis. The branching ratio uncertainty of the leading B backgrounds is the third-largest uncertainty.

Table 11.3: The impact of systematic uncertainties, estimated using the Asimov dataset.

| Source | Uncertainty on μ | | | |
|-----------------------|-----------------------------------|-------------------------------------|--------------------------------------|--------------------------------------|
| | $B^+ \rightarrow K^+\nu\bar{\nu}$ | $B^0 \rightarrow K_S^0\nu\bar{\nu}$ | $B^+ \rightarrow K^{*+}\nu\bar{\nu}$ | $B^0 \rightarrow K^{*0}\nu\bar{\nu}$ |
| $B\bar{B}$ norm. | 0.64 | 1.38 | 1.20 | 0.60 |
| Continuum norm. | 0.22 | 1.64 | 0.58 | 0.53 |
| MC stats | 0.47 | 1.14 | 0.95 | 0.59 |
| Leading \mathcal{B} | 0.21 | 0.16 | 0.45 | 0.20 |

11.5.1 Isospin average

The isospin average combines the two K and the two K^* modes, respectively. In this case, a single **POI** $\mu_{B \rightarrow K\nu\bar{\nu}}$ is used for the K^+ and K_S^0 modes, and a second $\mu_{B \rightarrow K^*\nu\bar{\nu}}$ is used for the K^{*+} and K^{*0} modes. The resulting uncertainties are $\sigma(\mu_{B \rightarrow K\nu\bar{\nu}}) = 1.13$ and $\sigma(\mu_{B \rightarrow K^*\nu\bar{\nu}}) = 1.30$. This corresponds to a 9.0% and a 15.0% improvement in accuracy compared to using the $B^+ \rightarrow K^+\nu\bar{\nu}$ and $B^0 \rightarrow K^{*0}\nu\bar{\nu}$ decays alone. The resulting branching ratio uncertainties are reported in [Table 11.4](#). The correlation matrix is reported in [Figure 11.8](#). There is only a weak correlation between the two signal strengths. The largest correlation between the signal and background components is found between $\mu_{B \rightarrow K\nu\bar{\nu}}$ and the B^+B^- background in the $B^+ \rightarrow K^+\nu\bar{\nu}$ channel.

The results of 1-dimensional profiled likelihood scans for each of the two parameters, performed on Asimov data, are shown in [Figure 11.9](#). The negative-log-likelihood curve is highly symmetric in all cases, which indicates the expected parabolic behaviour in the asymptotic regime. The uncertainties resulting from the profiled likelihood scan are shown in [Table 11.4](#).

11.5.2 4-channel correlation

Lastly, an analysis is performed assuming full correlation between the four K modes; thus, a single **POI** $\mu_{B \rightarrow K\nu\bar{\nu}}$ is used for all modes. The resulting uncertainty is $\sigma(\mu_{B \rightarrow K\nu\bar{\nu}}) =$

Table 11.4: Fit uncertainties on POIs in terms of signal strength μ and branching ratio \mathcal{B} . The world’s best results are taken from Reference [133, 134] (The Belle II result from Reference [2] is excluded).

| Channel | σ_μ | $\sigma_{\mathcal{B}} \cdot 10^{-6}$ | World’s best $\sigma_{\mathcal{B}} \cdot 10^{-6}$ | $\pm\sigma_\mu$ profiled |
|--|--------------|--------------------------------------|--|--------------------------|
| $B^+ \rightarrow K^+ \nu \bar{\nu}$ | 1.13 | 5.02 | 5.7 | + 1.14 - 1.15 |
| $B^0 \rightarrow K_s^0 \nu \bar{\nu}$ | 1.13 | 2.32 | 6.5 | + 1.14 - 1.15 |
| $B^+ \rightarrow K^{*+} \nu \bar{\nu}$ | 1.30 | 12.73 | 17 | + 1.34 - 1.29 |
| $B^0 \rightarrow K^{*0} \nu \bar{\nu}$ | 1.30 | 11.77 | 11 | + 1.34 - 1.29 |

0.82. This corresponds to a 50.0% improvement in accuracy compared to using the $B^+ \rightarrow K^+ \nu \bar{\nu}$ decay alone. The resulting branching ratio uncertainties are reported in Table 11.5. The correlation matrix is reported in Figure 11.10. The largest correlation between the signal and background components is found for the charged $B^+ B^-$ background in the $B^+ \rightarrow K^+ \nu \bar{\nu}$ channel.

The results of 1-dimensional profiled likelihood scans for the single parameter, performed on Asimov data, are shown in Figure 11.11. The negative-log-likelihood curve is highly symmetric, which indicates the expected parabolic behaviour in the asymptotic regime. The uncertainties resulting from the profiled likelihood scan are shown in Table 11.5.

Table 11.5: Fit uncertainties on POIs in terms of signal strength μ and branching ratio \mathcal{B} . The world’s best results are taken from Reference [133, 134] (The Belle II result from Reference [2] is excluded).

| Channel | σ_μ | $\sigma_{\mathcal{B}} \cdot 10^{-6}$ | World’s best $\sigma_{\mathcal{B}} \cdot 10^{-6}$ | $\pm\sigma_\mu$ profiled |
|--|--------------|--------------------------------------|--|--------------------------|
| $B^+ \rightarrow K^+ \nu \bar{\nu}$ | 0.82 | 3.64 | 5.7 | + 0.83 - 0.83 |
| $B^0 \rightarrow K_s^0 \nu \bar{\nu}$ | 0.82 | 1.68 | 6.5 | + 0.83 - 0.83 |
| $B^+ \rightarrow K^{*+} \nu \bar{\nu}$ | 0.82 | 8.03 | 17 | + 0.83 - 0.83 |
| $B^0 \rightarrow K^{*0} \nu \bar{\nu}$ | 0.82 | 7.42 | 11 | + 0.83 - 0.83 |

11.6 Discussion

This work-in-progress analysis demonstrates the significant advantages of a simultaneous four-channel approach to $B \rightarrow K \nu \bar{\nu}$ measurements at Belle II.

The expected uncertainties presented in [Tables 11.2, 11.4 and 11.5](#) show that Belle II can achieve sensitivity competitive with the world's best results using a significantly smaller dataset than the one of the Belle experiment. This demonstrates the power of combining multiple decay modes and the superior performance of the Belle II detector and analysis techniques. The systematic uncertainties are well controlled, with background normalization being the dominant source, followed by [MC](#) statistics and branching ratio uncertainties of leading backgrounds.

The simultaneous measurement of four $B \rightarrow K^{(*)}\nu\bar{\nu}$ channels opens new opportunities for model-agnostic reinterpretation studies. As demonstrated in [Section 7.3.2](#), the combination of pseudoscalar ($B^+ \rightarrow K^+\nu\bar{\nu}$, $B^0 \rightarrow K_s^0\nu\bar{\nu}$) and vector ($B^+ \rightarrow K^{*+}\nu\bar{\nu}$, $B^0 \rightarrow K^{*0}\nu\bar{\nu}$) final states enables the determination of the magnitudes of individual Wilson coefficients rather than just their sums. The different kinematic structures of these decays provide complementary sensitivity to various [BSM](#) scenarios. The statistical framework constructed here provides an ideal foundation for reinterpreting these results in terms of new physics models.

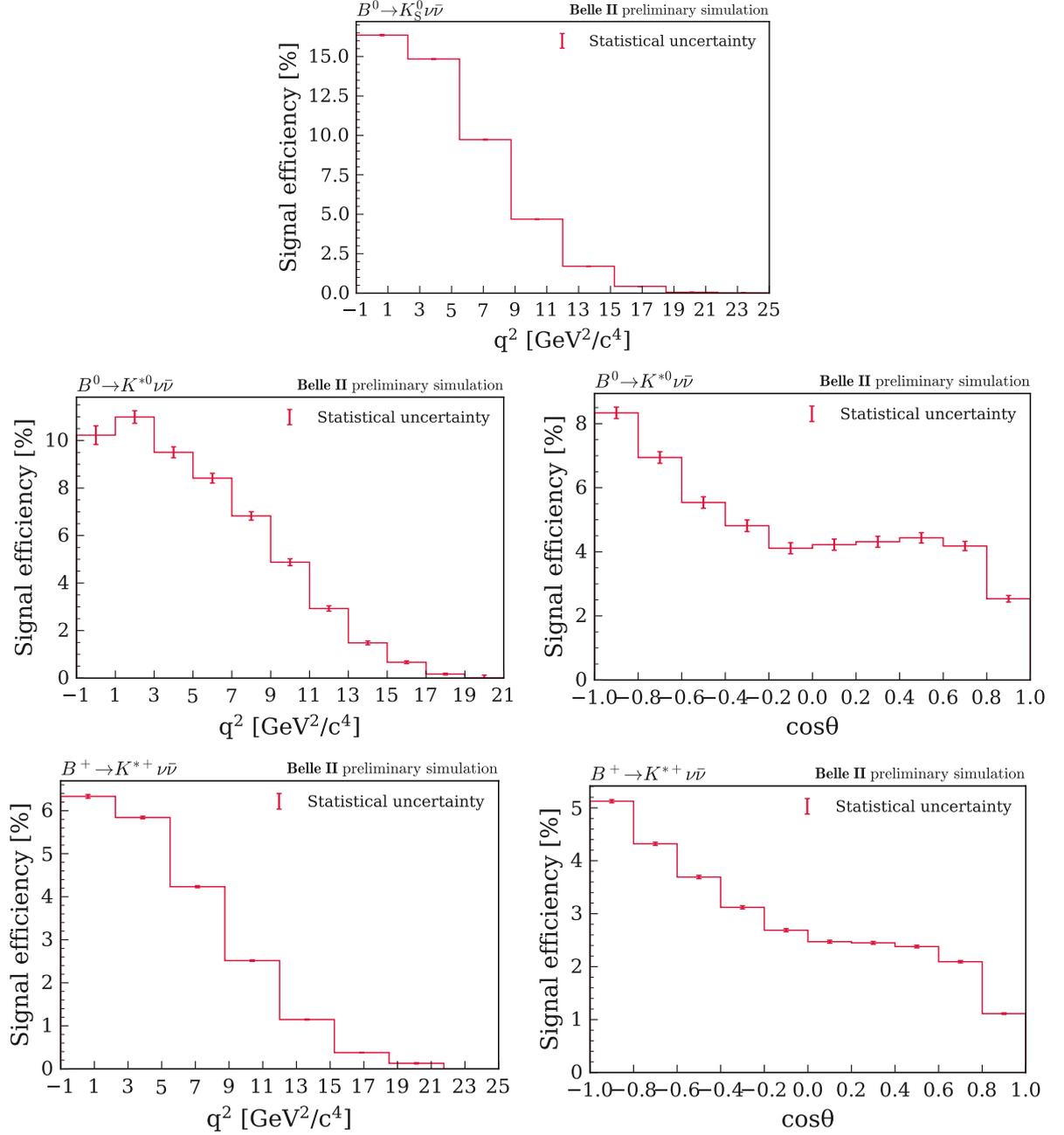


Figure 11.3: Signal efficiency as a function of q^2 and $\cos\theta$ for all channels, determined from the signal region selection. The $\cos\theta$ distributions are provided for vector mesons only.

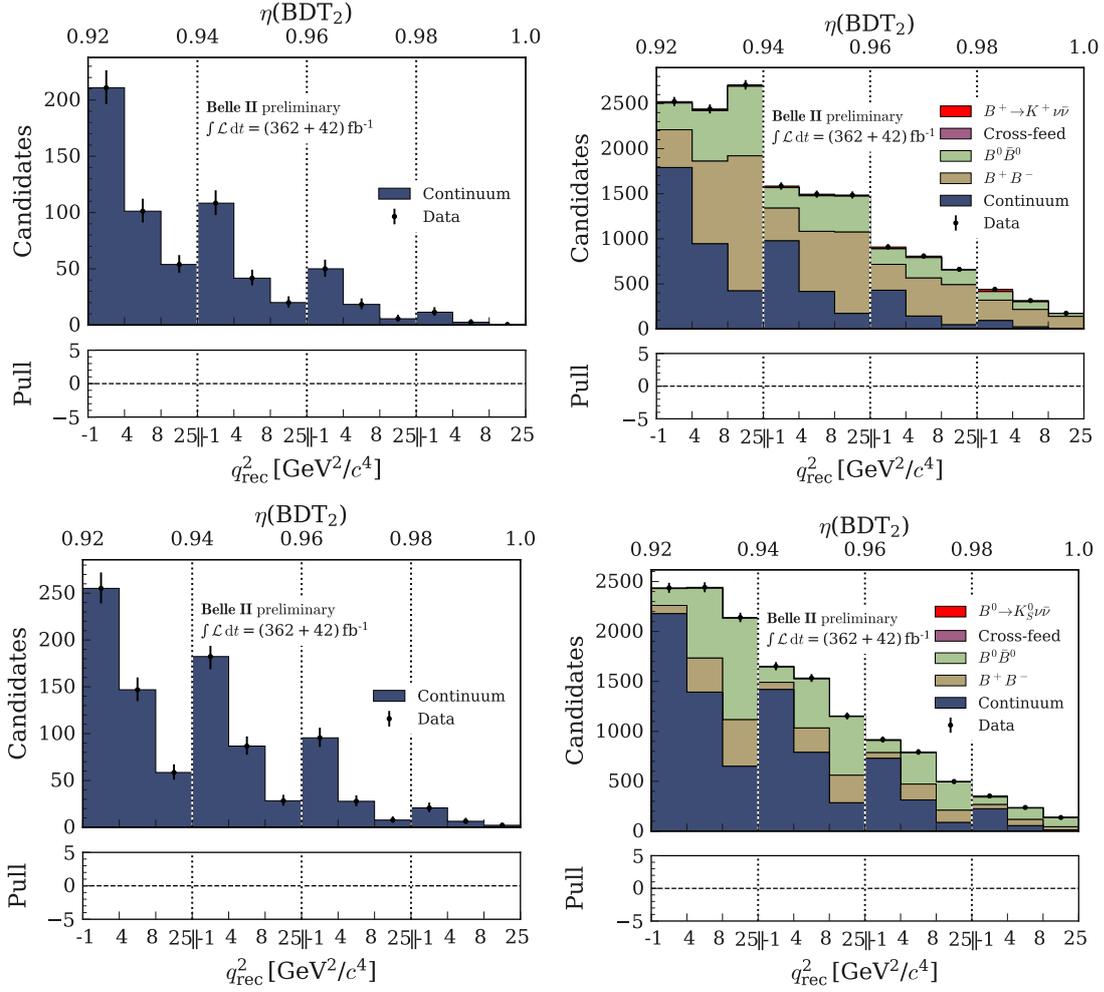


Figure 11.4: Post fit distribution from the Asimov fit of the $B^+ \rightarrow K^+\nu\bar{\nu}$ and $B^0 \rightarrow K_s^0\nu\bar{\nu}$ channels. The off-resonance region is shown on the *left* and the on-resonance region on the *right*. For each signal channel, the cross-feed indicates the events from the other three signal channels combined.

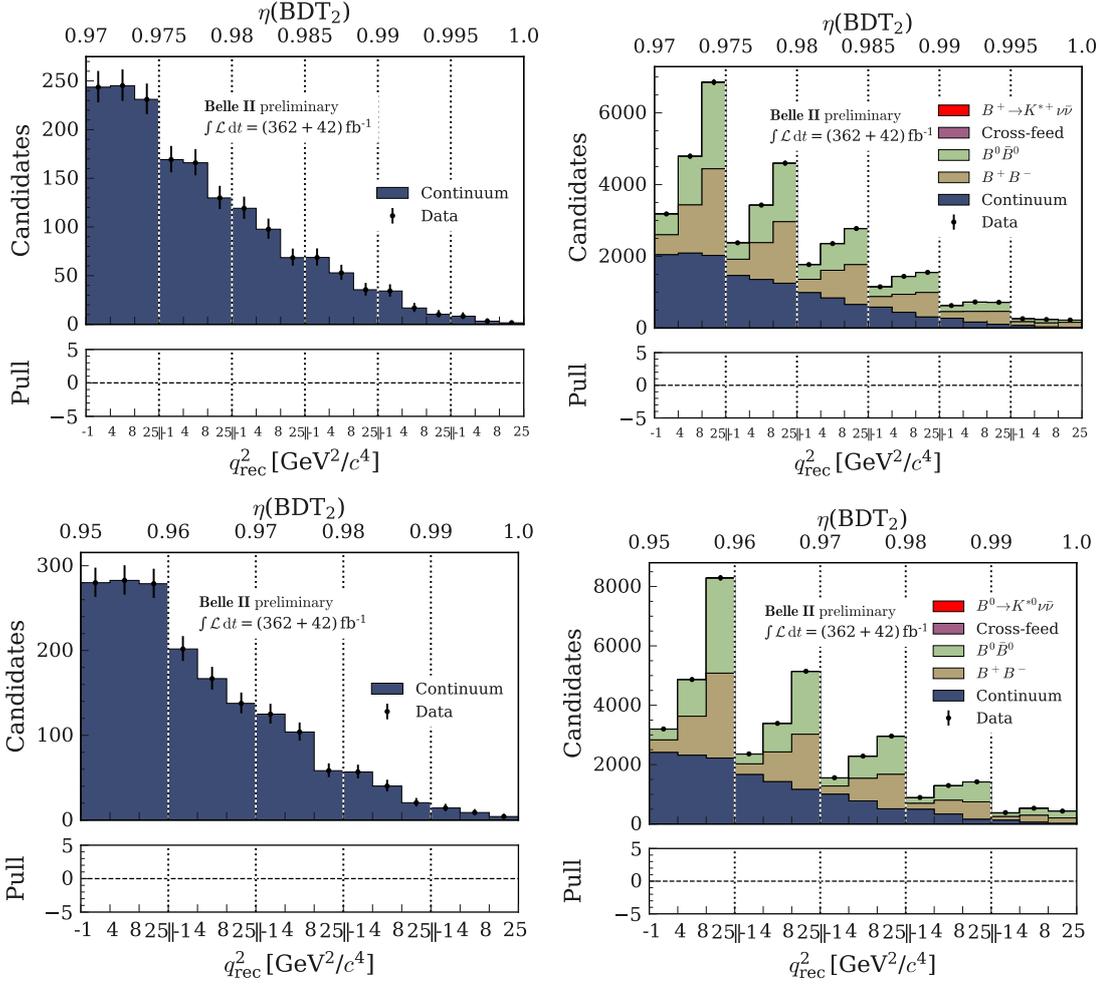


Figure 11.5: Post fit distribution from the Asimov fit of the $B^+ \rightarrow K^{*+} \nu \bar{\nu}$ and $B^0 \rightarrow K^{*0} \nu \bar{\nu}$ channels. The off-resonance region is shown on the *left* and the on-resonance region on the *right*. For each signal channel, the cross-feed indicates the events from the other three signal channels combined.

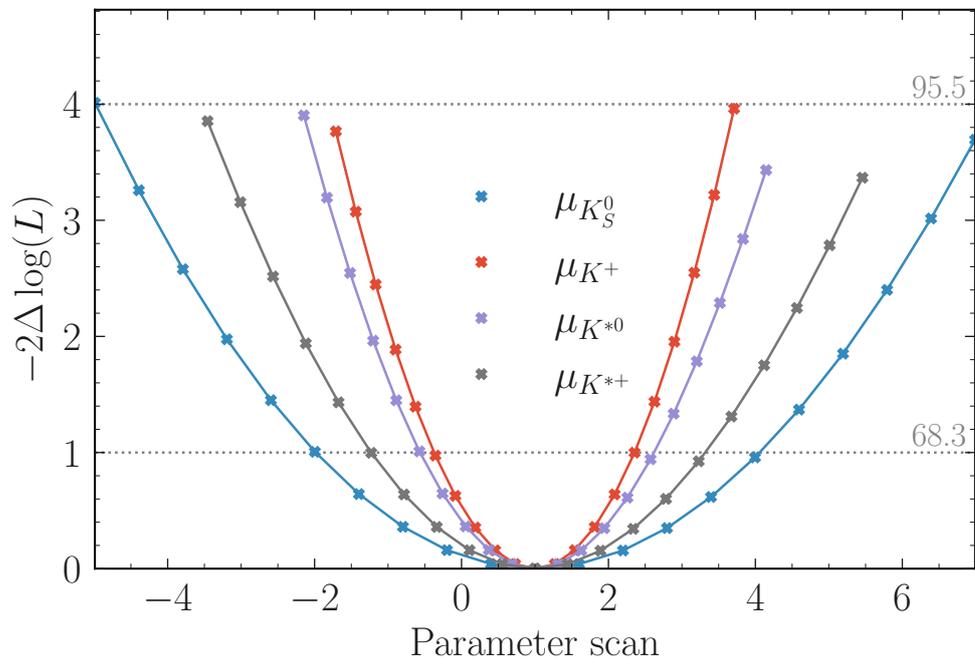


Figure 11.6: Profiled likelihood scan of the Asimov fit for the four signal channels.

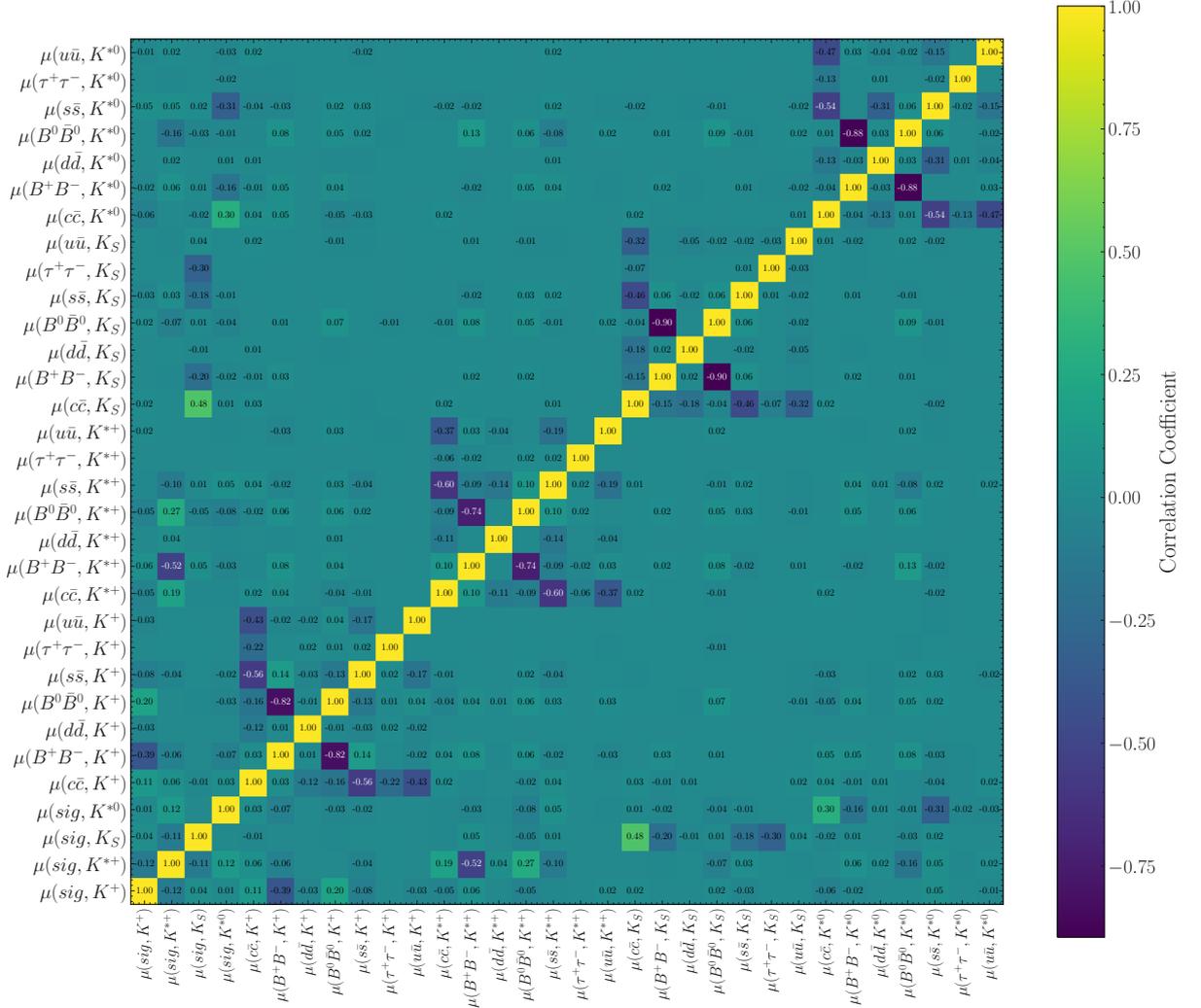


Figure 11.7: Correlation matrix of the uncorrelated model resulting from the Asimov fit. Only correlations of the normalization parameters are reported.

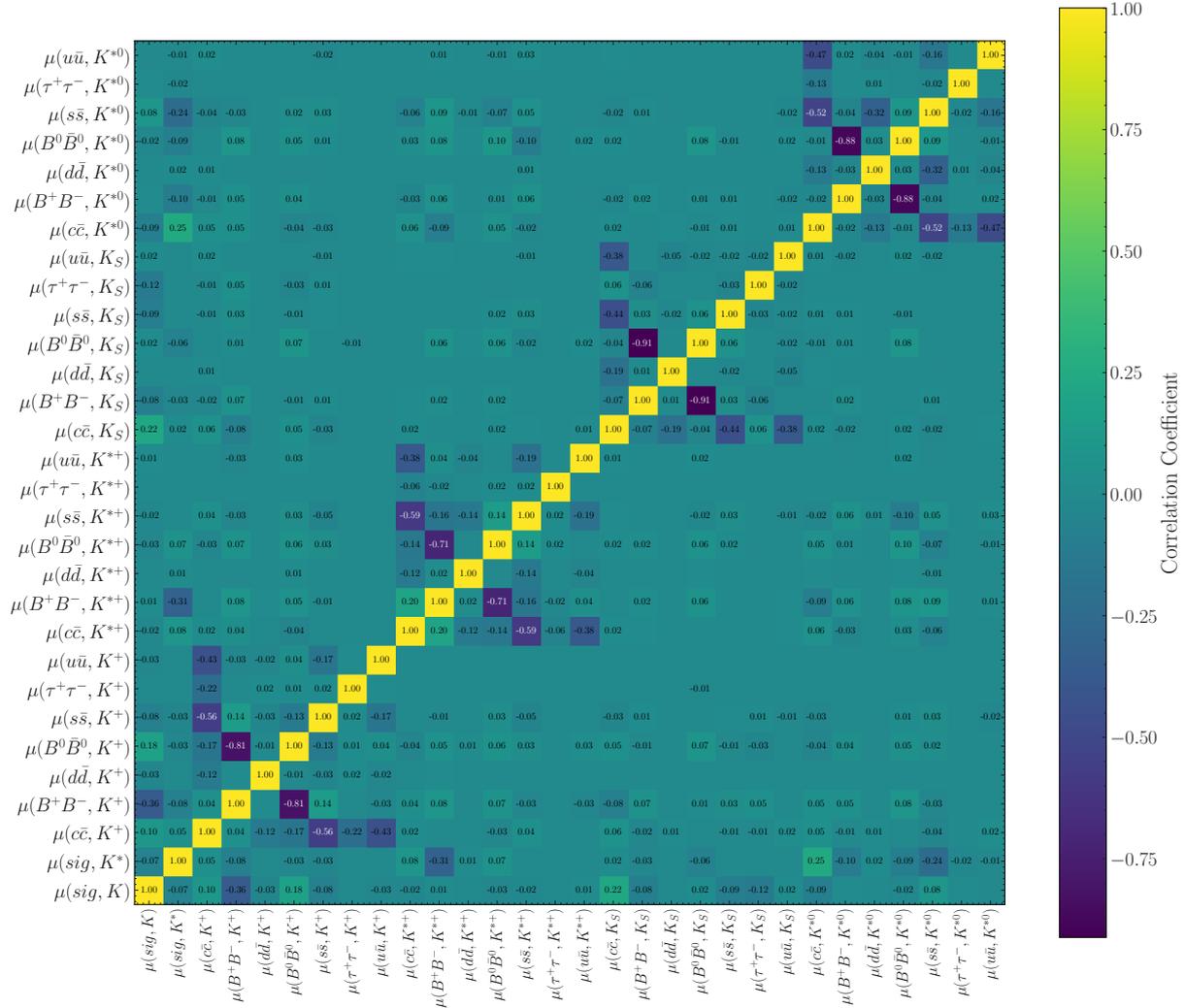


Figure 11.8: Correlation matrix of the isospin correlated model resulting from the Asimov fit. Only correlations of the normalization parameters are reported.

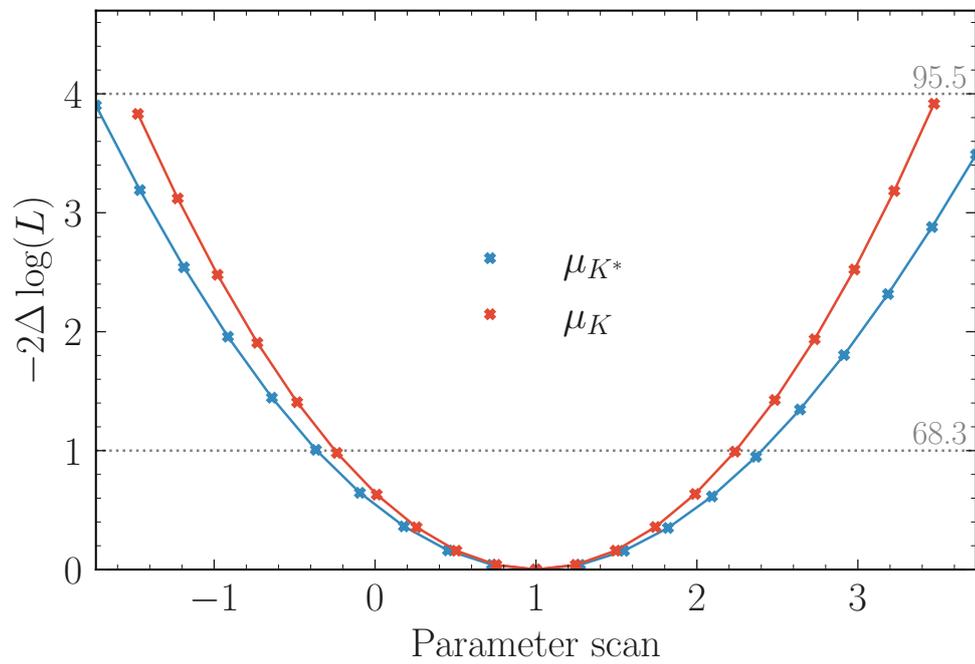


Figure 11.9: Profiled likelihood scans for the pseudoscalar **POI** $\mu_K = \mu(B \rightarrow K\nu\bar{\nu})$ and vector **POI** $\mu_{K^*} = \mu(B \rightarrow K^*\nu\bar{\nu})$ of the isospin averaged model, fit on Asimov data.

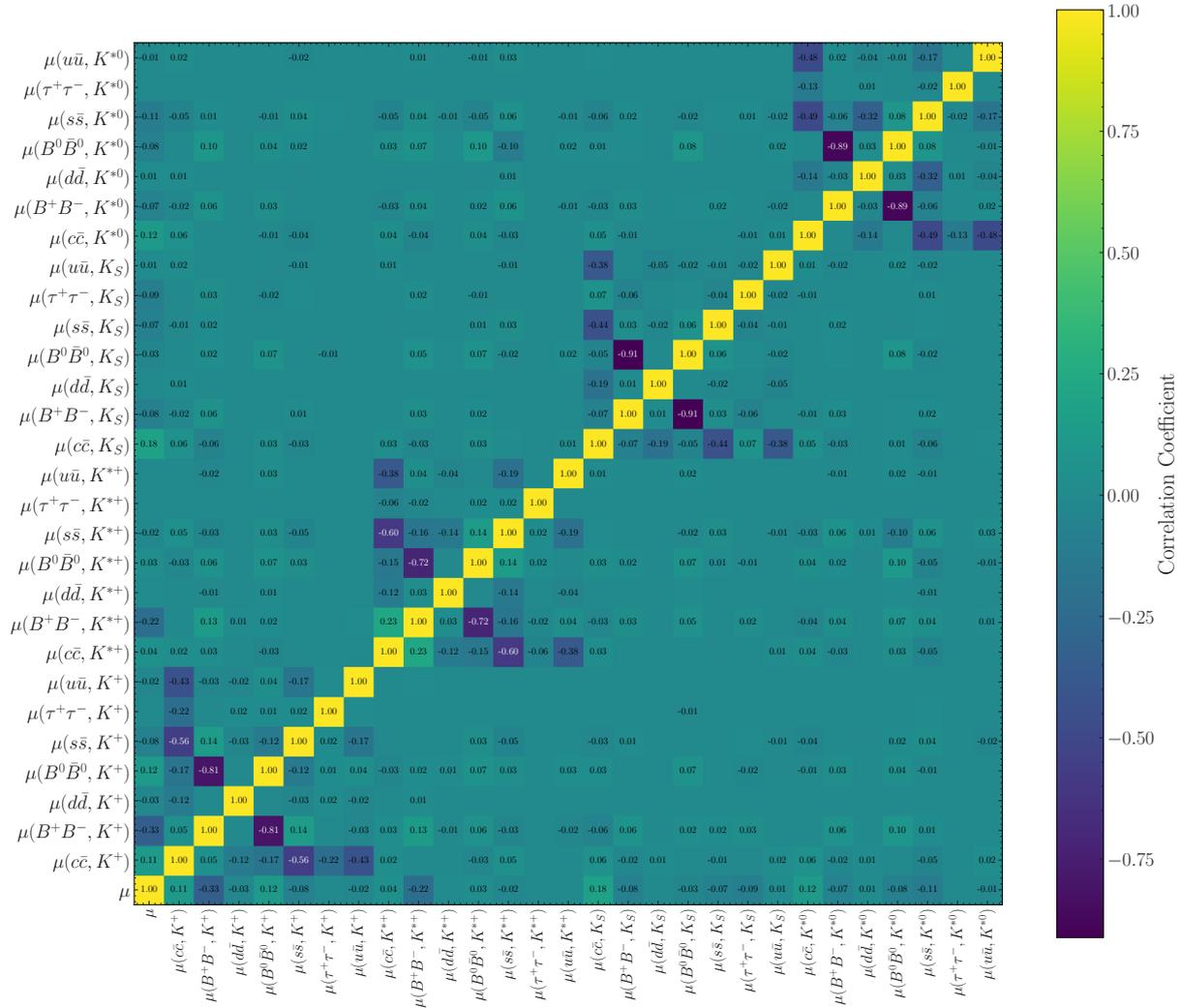


Figure 11.10: Correlation matrix of the fully correlated model resulting from the Asimov fit. Only correlations of the normalization parameters are reported.

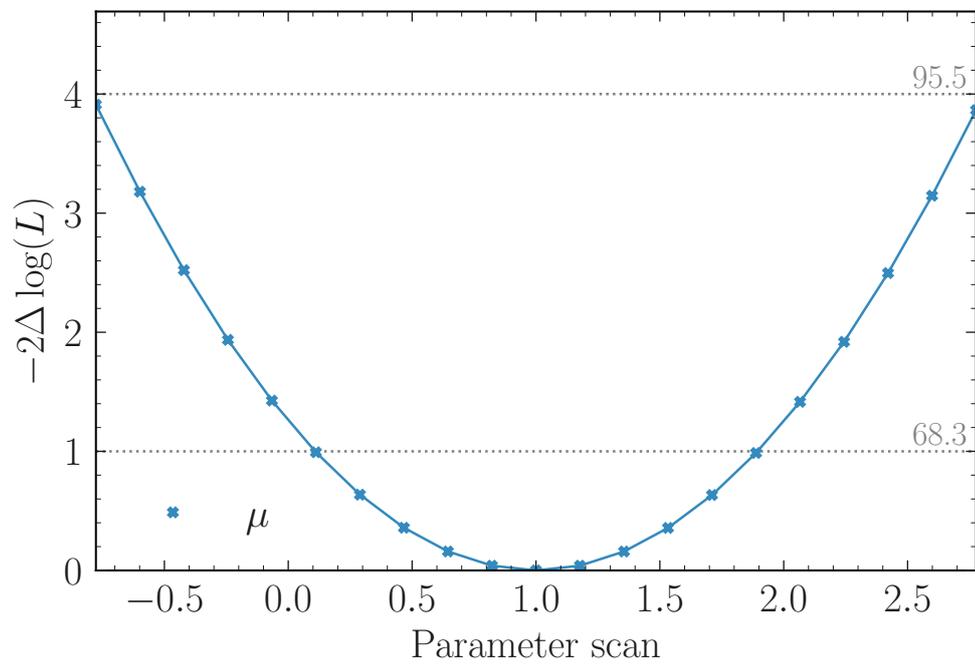


Figure 11.11: Profiled likelihood scan for the fully correlated signal strength across all channels, fit on Asimov data.

Chapter 12

Conclusion

This thesis addresses a fundamental challenge in **HEP**: how to accelerate the discovery of new physics through efficient and accurate reinterpretation of experimental results. The core idea is to reinterpret measurements designed for one theoretical hypothesis in terms of alternative frameworks. I develop, validate, and apply a novel reinterpretation method that preserves statistical rigour while dramatically reducing computational costs compared to a full reanalysis.

The key contribution is a model-agnostic reinterpretation method based on kinematic reweighting. This approach addresses limitations of existing techniques by enabling direct inference on theoretical parameters while correctly accounting for changes in kinematic distributions. Compared to current methods, it offers computational efficiency, statistical accuracy, accessibility, and broad applicability.

The method has been validated through toy examples and then applied to real experimental data. I focus on the Belle II $B^+ \rightarrow K^+ \nu \bar{\nu}$ measurement, which shows a 2.7σ tension with **SM** predictions. I demonstrate reinterpretations in two theoretical contexts:

Weak Effective Theory: I have obtained constraints on the **WET** Wilson coefficients relevant to $b \rightarrow s \nu \bar{\nu}$ processes, providing model-independent bounds on new physics contributions above the electroweak scale. These results represent both a successful application of the reinterpretation method and novel scientific contributions. The presented results are the first direct experimental constraints on the $b \rightarrow s \nu \bar{\nu}$ **WET** Wilson coefficients.

$B^+ \rightarrow K^+ X$ two-body decays: I apply the framework to a two-body light new physics scenario, demonstrating the versatility of the method in constraining specific **BSM** hypotheses with different kinematic signatures. The results demonstrate significantly improved compatibility with the experimental data compared to the **SM**. This illustrates how the method can explore different kinematic signatures and guide future investigations.

This work addresses a bottleneck in particle physics today. The complexity of experimental analyses and the proliferation of theoretical models create a mismatch between

the pace of theoretical development and experimental validation. By enabling rapid, accurate reinterpretation of existing measurements, the method presented here has several important implications:

Enhanced scientific impact: The resources invested in particle physics experiments can yield broader scientific impact when their results can be efficiently applied to test multiple theoretical hypotheses beyond those originally considered. Measurements with open source analysis logic are more likely to be cited and reused in future works, increasing their scientific impact [92].

Democratization of experimental data: By making reinterpretation accessible to theorists without requiring deep experimental expertise or computational resources, the method lowers barriers to scientific collaboration and accelerates the theory-experiment dialogue.

Extension of scientific legacy: Experimental measurements maintain their relevance and utility even as new theoretical frameworks emerge, ensuring that today’s measurements can continue contributing to tomorrow’s discoveries.

The reproducibility and distributability aspects of the method enable powerful combined analyses that can dramatically reduce uncertainties.

The foundation established in this work opens several avenues for future development:

Method extensions: The reweighting framework can be extended to handle more complex scenarios. Alternative approaches to joint number density construction can be explored to improve performance and robustness. The method can be applied within the context of other inference tools and techniques.

Broader applications: While this work has focused on Belle II measurements, the method applies equally to results from other experiments. As of today, colleagues at the LHCb experiment are applying the method and corresponding tool for a number of analyses (see e.g. [135]).

Global fits: The statistical models generated through this approach can be integrated into global fitting frameworks and combined with constraints from other experiments, providing much tighter bounds on theoretical parameters than any single measurement.

The reinterpretation method presented in this thesis contributes to a more sustainable and productive model for scientific progress. The approach advances open science initiatives by making experimental results more accessible and reusable. This directly supports the FAIR data principles that are increasingly recognized as essential for modern scientific practice [136].

The 2.7σ tension observed in the Belle II $B^+ \rightarrow K^+ \nu \bar{\nu}$ measurement illustrates why such advances are important. The ability to rapidly and accurately test multiple theoretical models against the same experimental data is a powerful accelerator of scientific progress.

As the field of particle physics continues to scale towards more data and higher energies,

methods that enhance what we can learn from existing measurements become increasingly valuable. The combination of improved experimental techniques, more sophisticated theoretical frameworks, and advanced statistical methods provides a foundation for such enhanced interpretation. This work contributes one piece to this larger endeavour, helping to ensure that experimental particle physics continues to deliver important scientific discoveries.

List of Publications

Papers

Lorenz Gärtner et al. “Constructing model-agnostic likelihoods, a method for the reinterpretation of particle physics results”. In: *Eur. Phys. J. C* 84.7 (2024), p. 693. DOI: [10.1140/epjc/s10052-024-13038-4](https://doi.org/10.1140/epjc/s10052-024-13038-4). arXiv: [2402.08417](https://arxiv.org/abs/2402.08417) [hep-ph].

Proceedings

Marco Colonna et al. “Redist, a Python Tool for Model-Agnostic Binned-Likelihood Fits in High Energy Physics”. In: *Proceedings of the 24th Python in Science Conference (SciPy 2025)*. SciPy Conference, July 2025. DOI: [10.25080/ppgk2467](https://doi.org/10.25080/ppgk2467). URL: <https://doi.org/10.25080/ppgk2467>.

In publication

Merna Abumusabh et al. “A model-agnostic likelihood for the reinterpretation of the $B^+ \rightarrow K^+ \nu \bar{\nu}$ measurement at Belle II”. In: *Accepted by Phys. Rev. D* (July 2025). arXiv: [2507.12393](https://arxiv.org/abs/2507.12393) [hep-ex].

Appendix A

Theoretical background

A.1 Group generators

The Standard Model gauge group $G_{SM} = SU(3)_C \times SU(2)_W \times U(1)_Y$ is characterized by its generators, which determine the structure of gauge interactions and the number of gauge bosons. Here we follow the conventions of References [23, 25].

A.1.1 U(1) generators

The Abelian group U(1) has a single generator, which can be taken as the identity operator. For U(1)_Y (hypercharge), transformations take the form

$$\psi \rightarrow e^{i\alpha Y} \psi, \quad (\text{A.1})$$

where α is a real parameter and Y is the hypercharge quantum number. This gives rise to a single gauge boson B_μ .

A.1.2 SU(2) generators

The SU(2)_W group has $2^2 - 1 = 3$ generators, conventionally chosen as

$$T_i = \frac{\sigma_i}{2}, \quad i = 1, 2, 3, \quad (\text{A.2})$$

where σ_i are the Pauli matrices:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (\text{A.3})$$

These generators satisfy the SU(2) algebra

$$[T_i, T_j] = i\varepsilon_{ijk} T_k, \quad (\text{A.4})$$

where ε_{ijk} is the totally antisymmetric Levi-Civita symbol. The generators are normalized according to

$$\text{Tr}(T_i T_j) = \frac{1}{2} \delta_{ij}. \quad (\text{A.5})$$

Transformations under $\text{SU}(2)_W$ take the form

$$\psi \rightarrow \exp(i\alpha_i T_i) \psi = \exp\left(i\frac{\alpha_i \sigma_i}{2}\right) \psi, \quad (\text{A.6})$$

where α_i are three real parameters. This generates three gauge bosons W_μ^i .

A.1.3 SU(3) generators

The $\text{SU}(3)_C$ group has $3^2 - 1 = 8$ generators, conventionally chosen as

$$T^a = \frac{\lambda^a}{2}, \quad a = 1, 2, \dots, 8, \quad (\text{A.7})$$

where λ^a are the Gell-Mann matrices:

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (\text{A.8})$$

$$\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \quad \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad (\text{A.9})$$

$$\lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \quad (\text{A.10})$$

The $\text{SU}(3)$ generators satisfy the commutation relations

$$[T^a, T^b] = i f^{abc} T^c, \quad (\text{A.11})$$

where f^{abc} are the totally antisymmetric structure constants of $\text{SU}(3)$. The structure constants are normalized according to

$$f^{ade} f^{bce} = \delta^{ab} N, \quad (\text{A.12})$$

where $N = 3$ for $\text{SU}(3)$. This normalization ensures that the structure constants satisfy the Jacobi identity and are consistent with the trace normalization of the generators.

The generators are normalized according to

$$\text{Tr}(T^a T^b) = \frac{1}{2} \delta^{ab}. \quad (\text{A.13})$$

Transformations under $SU(3)_C$ take the form

$$q \rightarrow \exp(i\alpha_a T^a) q = \exp\left(i\frac{\alpha_a \lambda^a}{2}\right) q, \quad (\text{A.14})$$

where α_a are 8 real parameters and q represents a quark field in the fundamental representation. This generates 8 gauge bosons A_μ^a (gluons).

A.2 Dirac matrices

The Dirac matrices γ^μ are fundamental objects in the theory of relativistic fermions [23, 25]. They satisfy the Clifford algebra

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu} \mathbb{1}, \quad (\text{A.15})$$

where $g^{\mu\nu} = \text{diag}(1, -1, -1, -1)$ is the Minkowski metric tensor and $\mathbb{1}$ is the identity matrix. The anticommutator is defined as $\{A, B\} = AB + BA$.

In the chiral (Weyl) representation, the gamma matrices are given by

$$\gamma^0 = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}, \quad (\text{A.16})$$

where $i = 1, 2, 3$ and σ^i are the Pauli matrices defined in [Appendix A.1](#).

The fifth gamma matrix is defined as

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 = \begin{pmatrix} -\mathbb{1} & 0 \\ 0 & \mathbb{1} \end{pmatrix}. \quad (\text{A.17})$$

This matrix satisfies $(\gamma^5)^2 = \mathbb{1}$ and anticommutes with all other gamma matrices:

$$\{\gamma^5, \gamma^\mu\} = 0. \quad (\text{A.18})$$

The projection operators for left- and right-handed fermions are constructed using γ^5 :

$$P_L = \frac{1 - \gamma^5}{2}, \quad P_R = \frac{1 + \gamma^5}{2}, \quad (\text{A.19})$$

satisfying $P_L + P_R = \mathbb{1}$, $P_L^2 = P_L$, $P_R^2 = P_R$, and $P_L P_R = P_R P_L = 0$.

The charge conjugation matrix C satisfies

$$C\gamma^\mu C^{-1} = -(\gamma^\mu)^T, \quad C\gamma^5 C^{-1} = (\gamma^5)^T, \quad (\text{A.20})$$

where T denotes matrix transposition. In the chiral representation,

$$C = \begin{pmatrix} i\sigma^2 & 0 \\ 0 & -i\sigma^2 \end{pmatrix}. \quad (\text{A.21})$$

The slash notation is commonly used for contractions with four-vectors:

$$\not{p} = \gamma^\mu p_\mu = p_0 \gamma^0 - \vec{p} \cdot \vec{\gamma}. \quad (\text{A.22})$$

A.3 $B \rightarrow K^*$ form factors

For the $B \rightarrow K^*$ transitions, vector operators in Equation (2.54) give rise to four independent form factors [33, 39]:

$$\langle K^*(k, \epsilon) | \bar{s} \gamma^\mu b | B(p) \rangle = \frac{2V(q^2)}{M_B + M_{K^*}} \varepsilon^{\mu\nu\rho\sigma} \epsilon_\nu^* p_\rho k_\sigma, \quad (\text{A.23})$$

$$\begin{aligned} \langle K^*(k, \epsilon) | \bar{s} \gamma^\mu \gamma_5 b | B(p) \rangle &= i \epsilon^{\mu*} (M_B + M_{K^*}) A_1(q^2) \\ &\quad - i (p + k)^\mu (\epsilon^* \cdot q) \frac{A_2(q^2)}{M_B + M_{K^*}} \\ &\quad - i q^\mu (\epsilon^* \cdot q) \frac{2M_{K^*}}{q^2} [A_3(q^2) - A_0(q^2)], \end{aligned} \quad (\text{A.24})$$

where

$$A_3(q^2) = \frac{M_B + M_{K^*}}{2M_{K^*}} A_1(q^2) - \frac{M_B - M_{K^*}}{2M_{K^*}} A_2(q^2), \quad (\text{A.25})$$

$$A_3(0) = A_0(0), \quad (\text{A.26})$$

and $\varepsilon^{\mu\nu\rho\sigma}$ is the totally antisymmetric Levi-Civita symbol in four dimensions, and $V(q^2)$, $A_0(q^2)$, $A_1(q^2)$, and $A_2(q^2)$ are the $B \rightarrow K^*$ form factors.

For the scalar operators, the non-zero matrix element is [33, 39]

$$\langle K^*(k, \epsilon) | \bar{s} \gamma_5 b | B(p) \rangle = i (\epsilon^* \cdot q) \frac{2M_{K^*}}{m_b - m_s} A_0(q^2). \quad (\text{A.27})$$

The tensor operator matrix elements are [33, 39]

$$\langle K^*(k, \epsilon) | \bar{s} \sigma^{\mu\nu} b | B(p) \rangle = 2 \varepsilon^{\mu\nu\rho\sigma} \epsilon_\nu^* p_\sigma k_\rho T_1(q^2) \quad (\text{A.28})$$

$$\begin{aligned} \langle K^*(k, \epsilon) | \bar{s} \sigma^{\mu\nu} \gamma_5 b | B(p) \rangle &= i [\epsilon^{\mu*} (p + k)^\nu - \epsilon^{\nu*} (p + k)^\mu] T_2(q^2) \\ &\quad + i (\epsilon^* \cdot q) [q^\mu (p + k)^\nu - q^\nu (p + k)^\mu] \frac{T_3(q^2)}{M_B^2 - M_{K^*}^2}, \end{aligned} \quad (\text{A.29})$$

where $T_1(q^2)$, $T_2(q^2)$, and $T_3(q^2)$ are the tensor form factors for $B \rightarrow K^*$ transitions, with the constraint $T_1(0) = T_2(0)$.

It is common to replace the form factors $A_2(q^2)$ and $T_3(q^2)$ with the linear combinations [33, 39]

$$A_{12}(q^2) = \frac{(M_B + M_{K^*})^2 (M_B^2 - M_{K^*}^2 - q^2) A_1(q^2) - \lambda(q^2, M_B, M_{K^*}) A_2(q^2)}{16 M_B M_{K^*}^2 (M_B + M_{K^*})}, \quad (\text{A.30})$$

$$T_{23}(q^2) = \frac{(M_B^2 - M_{K^*}^2) (M_B^2 + 3M_{K^*}^2 - q^2) T_2(q^2) - \lambda(q^2, M_B, M_{K^*}) T_3(q^2)}{8 M_B M_{K^*}^2 (M_B - M_{K^*})}, \quad (\text{A.31})$$

where $\lambda(q^2, M_B, M_{K^*})$ is the Källén function.

Scalar currents do not contribute to $B \rightarrow K^*$ transitions, by virtue of the symmetries of QCD. In particular, the invariance of the strong interaction under C , P , and CP transformations implies that these $B \rightarrow K^*$ matrix elements vanish.

A.4 $B \rightarrow K^* \nu \bar{\nu}$ hadronic form factor parameters

The theoretical predictions for $B \rightarrow K^* \nu \bar{\nu}$ decays discussed in [Section 2.4.2](#) rely on precise determinations of the $V(q^2)$, $A_0(q^2)$, $A_1(q^2)$, $A_{12}(q^2)$, $T_1(q^2)$, $T_2(q^2)$ and $T_{23}(q^2)$ hadronic form factors that parametrize the non-perturbative QCD matrix elements. The values for the corresponding 19 hadronic parameters obtained from the Gaussian likelihood provided in Reference [\[49\]](#) and are reported in [Table A.1](#). The values obtained from a joint theoretical prior PDF based on results from References [\[130, 131\]](#) are reported in [Table A.2](#).

Table A.1: The hadronic parameters used to parametrize the $V(q^2)$, $A_0(q^2)$, $A_1(q^2)$, $A_{12}(q^2)$, $T_1(q^2)$, $T_2(q^2)$ and $T_{23}(q^2)$ form factors in the **BSZ** parametrization [48], based on results of Reference [49]. The table shows the mean values (first row) and the covariance matrix (upper triangular part).

| v_0 | v_1 | v_2 | a_0^0 | a_1^0 | a_2^0 | a_0^1 | a_1^1 | a_2^1 | a_0^2 | a_1^2 | a_2^2 | t_0^1 | t_1^1 | t_2^1 | t_0^2 | t_1^2 | t_2^2 | t_0^3 | t_1^3 | t_2^3 | |
|--------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|--------|
| 0.3628 | -1.0893 | 2.8029 | 0.3421 | -1.1474 | 2.3728 | 0.2893 | 0.4599 | 1.2242 | 0.5512 | 0.5781 | 0.3192 | -0.9619 | 2.0931 | 0.5932 | 1.6795 | 0.6212 | 0.9733 | 1.8213 | | | |
| 0.0014 | -0.0008 | -0.0633 | 0.0008 | 0.0039 | -0.0001 | 0.0135 | 0.0506 | 0.0012 | -0.0030 | -0.0630 | 0.0006 | 0.0011 | 0.0130 | 0.0040 | 0.0139 | 0.0011 | 0.0015 | -0.0199 | | | |
| | 0.0660 | 0.3765 | -0.0010 | -0.0010 | -0.0195 | 0.0219 | 0.1498 | -0.0034 | 0.0147 | 0.1876 | -0.0009 | -0.0033 | -0.0733 | -0.0083 | -0.0549 | -0.0025 | -0.0039 | 0.0500 | | | |
| | | 7.8174 | -0.0242 | -0.1684 | -0.0881 | -0.5806 | -1.5977 | -0.0509 | 0.0928 | 4.0359 | -0.0087 | -0.1056 | -1.9914 | -0.1922 | -1.5282 | -0.0424 | 0.0235 | 1.3014 | | | |
| | | | 0.0009 | 0.0036 | -0.0038 | 0.0069 | 0.0273 | 0.0014 | -0.0032 | -0.0635 | 0.0008 | 0.0003 | -0.0059 | 0.0042 | 0.0044 | 0.0011 | 0.0033 | -0.0102 | | | |
| | | | | 0.0286 | 0.0496 | 0.0571 | 0.2767 | 0.0061 | -0.0119 | -0.3353 | 0.0035 | 0.0039 | 0.0102 | 0.0192 | 0.0344 | 0.0052 | 0.0130 | -0.0528 | | | |
| | | | | | 0.5464 | 0.0785 | 0.5791 | -0.0011 | -0.0094 | 0.0201 | -0.0024 | 0.0085 | 0.1791 | -0.0114 | 0.0269 | -0.0014 | -0.0106 | -0.0011 | | | |
| | | | | | | 0.2434 | 1.2378 | 0.0097 | -0.0156 | -0.6045 | 0.0042 | 0.0162 | 0.2251 | 0.0387 | 0.1792 | 0.0093 | 0.0013 | -0.2289 | | | |
| | | | | | | | | 0.0371 | -0.0532 | -2.2307 | 0.0170 | 0.0732 | 0.9817 | 0.1582 | 0.7220 | 0.0370 | -0.0092 | -0.9501 | | | |
| | | | | | | | | | 0.0025 | -0.0069 | -0.1353 | 0.0014 | 0.0007 | -0.0065 | 0.0071 | 0.0098 | 0.0019 | 0.0055 | -0.0183 | | |
| | | | | | | | | | | 0.0464 | 0.4091 | -0.0033 | -0.0008 | 0.0285 | -0.0161 | -0.0127 | -0.0048 | -0.0140 | 0.0571 | | |
| | | | | | | | | | | | 10.012 | -0.0612 | -0.0618 | -0.1433 | -0.3813 | -0.8975 | -0.0937 | -0.2439 | 1.2014 | | |
| | | | | | | | | | | | | 0.0010 | -0.0005 | 0.0189 | 0.0038 | -0.0045 | 0.0012 | 0.0039 | -0.0056 | | |
| | | | | | | | | | | | | | 0.0205 | 0.1149 | 0.0087 | 0.0599 | 0.0002 | 0.0031 | -0.0219 | | |
| | | | | | | | | | | | | | | 2.6624 | -0.0004 | 1.0087 | -0.0053 | -0.1341 | -0.2744 | | |
| | | | | | | | | | | | | | | | 0.0451 | 0.1751 | 0.0057 | 0.0138 | -0.1063 | | |
| | | | | | | | | | | | | | | | | 1.5341 | 0.0075 | -0.0664 | -0.5565 | | |
| | | | | | | | | | | | | | | | | | 0.0039 | -0.0454 | -0.0454 | | |
| | | | | | | | | | | | | | | | | | | 0.0022 | 0.0039 | 0.4790 | |
| | | | | | | | | | | | | | | | | | | | 0.0998 | 0.4790 | 5.9109 |

Table A.2: The hadronic parameters used to parametrize the $V(q^2)$, $A_0(q^2)$, $A_1(q^2)$, $A_{12}(q^2)$, $T_1(q^2)$, $T_2(q^2)$ and $T_{23}(q^2)$ form factors in the **BSZ** parametrization [48], based on results of References [130, 131]. The table shows the mean values (first row) and the covariance matrix (upper triangular part).

| v_0 | v_1 | v_2 | a_0^0 | a_1^0 | a_2^0 | a_0^1 | a_1^1 | a_2^1 | a_{12}^1 | a_{12}^2 | t_0^1 | t_1^1 | t_2^1 | t_1^2 | t_2^2 | t_{23}^0 | t_{23}^1 | t_{23}^2 |
|--------|---------|---------|---------|---------|---------|---------|---------|---------|------------|------------|---------|---------|---------|---------|---------|------------|------------|------------|
| 0.3631 | -1.0903 | 2.7817 | 0.2894 | 0.4597 | 1.2199 | 0.5422 | 0.5465 | 0.3415 | -1.1442 | 2.4375 | 0.3193 | -0.9574 | 2.1215 | 0.5971 | 1.7001 | 0.6216 | 0.9705 | 1.8025 |
| 0.0008 | -0.0013 | -0.0348 | 0.0004 | 0.0018 | -0.0013 | 0.0017 | 0.0067 | 0.0003 | -0.0002 | -0.0012 | 0.0005 | 0.0000 | -0.0085 | 0.0026 | 0.0029 | 0.0006 | 0.0024 | 0.0019 |
| 0.0284 | 0.1068 | -0.0002 | 0.0019 | -0.0058 | 0.0033 | 0.0118 | -0.0004 | 0.0043 | 0.0043 | -0.0242 | -0.0002 | 0.0003 | 0.0030 | -0.0006 | 0.0000 | -0.0004 | -0.0008 | 0.0090 |
| | 4.0579 | -0.0115 | -0.0790 | 0.1297 | -0.0983 | -0.3947 | -0.0043 | -0.0306 | 0.3915 | -0.0139 | 0.0009 | 0.0009 | 0.2689 | -0.0745 | -0.0657 | -0.0168 | -0.0617 | 0.0177 |
| | 0.0004 | 0.0012 | -0.0031 | 0.0018 | 0.0059 | 0.0003 | 0.0000 | 0.0000 | -0.0035 | 0.0003 | 0.0000 | 0.0000 | -0.0062 | 0.0018 | 0.0022 | 0.0004 | 0.0017 | 0.0023 |
| | 0.0174 | 0.0550 | 0.0261 | 0.1422 | 0.0011 | 0.0049 | -0.0235 | 0.0014 | 0.0001 | -0.0247 | 0.0076 | 0.0094 | 0.0018 | 0.0073 | 0.0053 | 0.0018 | 0.0073 | 0.0053 |
| | | 0.5522 | 0.0782 | 0.6026 | -0.0007 | 0.0046 | 0.1593 | -0.0011 | -0.0001 | 0.0382 | -0.0074 | -0.0093 | -0.0004 | -0.0055 | -0.0454 | 0.0020 | 0.0081 | 0.0030 |
| | | 0.1186 | 0.6613 | 0.0047 | 0.0167 | -0.1270 | 0.0014 | -0.0009 | -0.0272 | 0.0064 | 0.0047 | 0.0047 | 0.0020 | 0.0081 | 0.0030 | 0.0064 | 0.0280 | 0.0263 |
| | | 4.3665 | 0.0118 | 0.0759 | 0.0016 | 0.0053 | -0.0024 | -0.0999 | 0.0231 | 0.0110 | 0.0064 | 0.0280 | 0.0263 | 0.0003 | 0.0013 | 0.0004 | 0.0004 | 0.0004 |
| | | 0.0007 | 0.0004 | -0.0195 | 0.0002 | -0.0001 | -0.0041 | 0.0011 | 0.0011 | 0.0011 | 0.0011 | 0.0011 | 0.0011 | 0.0011 | 0.0011 | 0.0003 | 0.0013 | 0.0004 |
| | | 0.0380 | 0.1255 | -0.0001 | 0.0006 | 0.0053 | 0.0003 | 0.0038 | -0.0003 | -0.0004 | 0.0008 | 0.0008 | 0.0008 | 0.0008 | 0.0008 | 0.0003 | 0.0004 | 0.0008 |
| | | | 3.0766 | 0.0014 | 0.0074 | 0.0218 | -0.0026 | 0.0059 | -0.0050 | -0.0131 | 0.0332 | 0.0332 | 0.0332 | 0.0332 | 0.0332 | 0.0050 | -0.0131 | 0.0332 |
| | | | 0.0004 | -0.0004 | -0.0110 | 0.0019 | -0.0004 | 0.0005 | 0.0017 | 0.0023 | 0.0023 | 0.0023 | 0.0023 | 0.0023 | 0.0023 | 0.0005 | 0.0017 | 0.0023 |
| | | | 0.0189 | 0.0704 | 0.0043 | 0.0198 | -0.0003 | 0.0047 | -0.0073 | 0.0737 | 0.0737 | 0.0737 | 0.0737 | 0.0737 | 0.0737 | -0.0063 | -0.0624 | 0.0737 |
| | | | 1.6254 | 0.0331 | 0.1339 | 0.0021 | 0.0143 | 0.0048 | 0.0048 | 0.0048 | 0.0048 | 0.0048 | 0.0048 | 0.0048 | 0.0048 | 0.0021 | 0.0143 | 0.0048 |
| | | | 0.9558 | 0.0012 | 0.0017 | 0.0264 | 0.0012 | 0.0017 | 0.0264 | 0.0012 | 0.0017 | 0.0264 | 0.0012 | 0.0017 | 0.0264 | 0.0012 | 0.0017 | 0.0264 |
| | | | 0.5292 | 0.0957 | 5.7139 | | | | | | | | | | | | | |

APPENDIX A. THEORETICAL BACKGROUND

Appendix B

Statistical inference in particle physics

B.1 Probability density functions

For a measurement outcome depending on a continuous variable x , the probability to find x in an infinitesimal range around a value x' is given by the **PDF** $p(x)$ [65],

$$P(x \in [x', x' + dx]) = p(x')dx. \quad (\text{B.1})$$

The **PDF** is normalized to 1 over the whole range of x , $[a, b]$,

$$\int_a^b dx p(x) = 1. \quad (\text{B.2})$$

PDFs play a fundamental role in statistical analyses. The fundamental goal of any data analysis is to gather information about certain parameters, given large amounts of data. For example, if one attempts to measure the mass of a particle, what is actually obtained is a **PDF** of the mass that indicates how likely certain values are. Ideally, this **PDF** is a narrow peak around the true value of the mass. The width of the **PDF** then defines the uncertainty on the result.

In particle physics, data are often binned in histograms. Consider a set of N measurements, x_1, \dots, x_N . The x -range can be subdivided into m intervals of width Δx_i for $i = 1, \dots, m$, and one can count how many events n_i are in each of those intervals. The value of the **PDF** is then given by normalizing the area under the histogram to 1. For a specific bin i this is given by

$$p_i = \frac{n_i}{\sum_i n_i \Delta x_i}. \quad (\text{B.3})$$

The probability of finding an event in a certain bin i is then

$$P(x \in \Delta x_i) = p_i \Delta x_i. \quad (\text{B.4})$$

In the continuous limit $m \rightarrow \infty$, the system returns to Equation (B.1). From Equation (B.3), it can be easily seen that the PDF is correctly normalized, since

$$\sum_i p_i \Delta x_i = 1. \quad (\text{B.5})$$

Hence, there is a mapping between continuous and binned probability models,

$$\int p(x) dx \leftrightarrow \sum p_i \Delta x_i, \quad (\text{B.6})$$

which can be applied to any of the following definitions in this text.

B.1.1 Cumulative distribution functions

The CDF is the probability of a measurement less than or equal to x [65],

$$F(x) = \int_{-\infty}^x p(x') dx'. \quad (\text{B.7})$$

A *quantile* of α , x_α , is then defined such that the value $F(x_\alpha) = \alpha$, where $0 \leq \alpha \leq 1$. The quantile is then given by

$$x_\alpha = F^{-1}(\alpha). \quad (\text{B.8})$$

B.1.2 Transformation properties

Consider now the transformation properties of PDFs. Suppose there is a random variable x with PDF $p_x(x)$, and a transformation is applied to x such that a new variable y is related to x by the equation

$$y = g(x), \quad (\text{B.9})$$

where g is a differentiable and invertible function. The goal is to find the PDF of y , denoted $p_y(y)$, in terms of the PDF of x .

By the law of total probability, the probability that y lies in the range $[y', y' + dy]$ is the same as the probability that x lies in the range $[x', x' + dx]$, where $x' = g^{-1}(y')$. Therefore, the following holds [65]:

$$P(y \in [y', y' + dy]) = P(x \in [x', x' + dx]). \quad (\text{B.10})$$

Since $P(x \in [x', x' + dx]) = p_x(x') dx$, it follows that:

$$p_y(y') dy = p_x(x') dx. \quad (\text{B.11})$$

Substituting $dx = \left| \frac{dx}{dy} \right| dy$ gives:

$$p_y(y') = p_x(g^{-1}(y')) \left| \frac{dx}{dy} \right|. \quad (\text{B.12})$$

Thus, the PDF of y is related to the PDF of x by the transformation rule:

$$p_y(y) = p_x(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right|. \quad (\text{B.13})$$

B.1.3 Expectation values

The *expectation value* of a random variable x distributed according to the PDF $p(x)$ is given by [65]

$$E[x] = \int_{-\infty}^{\infty} xp(x)dx, \quad (\text{B.14})$$

which is not a function of x , but depends only on the shape of $p(x)$. If $p(x)$ is concentrated in one region, $E[x]$ corresponds to likely observed values of x . If $p(x)$ has separated peaks, then values of x corresponding to $E[x]$ can actually be unlikely values to observe.

The *variance* is defined as [65]

$$V[x] = E[(x - E[x])^2] = E[x^2] - E[x]^2, \quad (\text{B.15})$$

which measures the spread of x around its mean. The *standard deviation* is

$$\sigma[x] = \sqrt{V[x]}. \quad (\text{B.16})$$

In the case of multiple variables, the *covariance* is defined as

$$\text{cov}[x_1, x_2] = E[(x_1 - E[x_1])(x_2 - E[x_2])] = E[x_1x_2] - E[x_1]E[x_2], \quad (\text{B.17})$$

which encodes information on the correlation of the two variables. The covariance matrix is symmetric by construction and the diagonal elements correspond to $\text{cov}[x_i, x_i] = \sigma[x_i]^2$. To have a dimensionless measure of correlation, one can define the *correlation matrix*

$$\rho_{ij} = \frac{\text{cov}[x_i, x_j]}{\sigma[x_i]\sigma[x_j]}. \quad (\text{B.18})$$

The range of any coefficient in the correlation matrix is bounded by $-1 \leq \rho_{ij} \leq 1$.

A way to think about the meaning of the correlation matrix is as follows: a positive correlation coefficient $\rho_{12} > 0$ corresponds to “if I measure $x_1 > E[x_1]$, then it is more likely to obtain $x_2 > E[x_2]$ ”. Conversely, a negative correlation coefficient $\rho_{12} < 0$ corresponds to “if I measure $x_1 > E[x_1]$, then it is more likely to obtain $x_2 < E[x_2]$ ”.

B.2 Conjugate priors

A *conjugate prior* is a prior distribution that, when combined with a specific likelihood function, results in a posterior distribution of the same family as the prior. This property simplifies the process of Bayesian updating, as the posterior can be easily computed and interpreted.

The constraint likelihoods in the HistFactory statistical model are either normal or Poisson PDFs. Consider the two cases:

Normal constraints Most modifiers have normal constraint terms of the form

$$\mathcal{N}(a|\chi, \sigma_\chi) . \quad (\text{B.19})$$

In these cases, a suitable choice of conjugate prior is another normal PDF,

$$\mathcal{N}(\chi|\chi', \sigma'_\chi) . \quad (\text{B.20})$$

Since the product of two normal distributions is another normal distribution [68], the posterior becomes

$$p(\chi) = \mathcal{N}(\chi|\bar{\chi}, \bar{\sigma}) \propto \mathcal{N}(a|\chi, \sigma)\mathcal{N}(\chi|\chi', \sigma'_\chi) , \quad (\text{B.21})$$

with

$$\bar{\chi} = \bar{\sigma}^2 \left(\frac{\chi'}{\sigma'^2_\chi} + \frac{a}{\sigma^2_\chi} \right) , \quad \bar{\sigma}^2 = \frac{\sigma^2_\chi \sigma'^2_\chi}{\sigma^2_\chi + \sigma'^2_\chi} . \quad (\text{B.22})$$

In the limit $\sigma_\chi \ll \sigma'_\chi$,

$$\bar{\chi} \rightarrow a , \quad \bar{\sigma} \rightarrow \sigma_\chi , \quad (\text{B.23})$$

the posterior is (auxiliary) data-driven and not prior-driven.

Poisson constraints For the uncorrelated shape modifier, the individual constraint terms have the form

$$\text{Pois}(a|a\chi) . \quad (\text{B.24})$$

The conjugate prior for the Poisson distribution is the Gamma distribution [68],

$$p_\Gamma(\theta|\alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} e^{-\beta\theta} . \quad (\text{B.25})$$

Combining the above two equations and setting $\theta = a\chi$, the posterior is

$$p(a\chi) = p_\Gamma(a\chi|\alpha + a, \beta + 1) \propto \text{Pois}(a|a\chi)p(a\chi|\alpha, \beta) . \quad (\text{B.26})$$

With the parameter transformation property (see Appendix B.1.2)

$$p(\chi) = \frac{d(a\chi)}{d\chi} p(a\chi) = ap(a\chi) , \quad (\text{B.27})$$

the posterior for χ is simply another Gamma distribution,

$$p(\chi) = p_\Gamma(\chi|\bar{\alpha}, \bar{\beta}) = a p_\Gamma(a\chi|\bar{\alpha}, \bar{\beta}/a) , \quad (\text{B.28})$$

with

$$\bar{\alpha} = \alpha + a , \quad \bar{\beta} = a(\beta + 1) . \quad (\text{B.29})$$

In the limit of $\alpha \rightarrow 0$ and $\beta \rightarrow 0$,

$$p(\chi) \propto \text{Pois}(a|a\chi) , \quad (\text{B.30})$$

such that the posterior is (auxiliary) data-driven and not prior-driven.

B.3 Proof of the Cramér-Rao bound

The *Cramér-Rao bound* is an important result, which sets a lower bound on the variance of any estimator. The variance of an estimator $\hat{\theta}_i \in \hat{\boldsymbol{\theta}}$ with a sampling distribution $p(\hat{\boldsymbol{\theta}}_i|\boldsymbol{\theta}_0)$ is given by,

$$\mathbf{V}_i[\hat{\boldsymbol{\theta}}] = \int d^m \hat{\boldsymbol{\theta}} (\hat{\theta}_i - \mathbf{E}_i[\hat{\boldsymbol{\theta}}])^2 p(\hat{\boldsymbol{\theta}}_i|\boldsymbol{\theta}_0) . \quad (\text{B.31})$$

The Cramér-Rao bound sets a lower limit on the variance in terms of the Fisher information and the bias [67],

$$\mathbf{V}_i[\hat{\boldsymbol{\theta}}] \geq \frac{\partial \mathbf{E}_i[\hat{\boldsymbol{\theta}}]/\partial \theta_i}{\mathbf{I}_{ii}(\hat{\boldsymbol{\theta}})} = \frac{(1 + \partial \mathbf{b}_i/\partial \theta_i)^2}{\mathbf{I}_{ii}(\hat{\boldsymbol{\theta}})} . \quad (\text{B.32})$$

This proof only considers a single estimator, but it is easily extendable to multidimensional estimators, $\hat{\boldsymbol{\theta}}$. Consider a scalar estimator $\hat{\theta}(\mathbf{x})$ with a sampling distribution $p(\hat{\theta}|\theta_0)$. The conditions under which Equation (B.32) holds are

- the range of \mathbf{x} is independent of θ ,
- $\partial/\partial \theta_0$ commutes with $\int d\mathbf{x}$ or equivalently $\int d\hat{\theta}$.

Starting with the normalization condition of any PDF,

$$\int d\hat{\theta} p(\hat{\theta}|\theta_0) = 1 . \quad (\text{B.33})$$

Differentiating with respect to θ_0 gives

$$\int d\hat{\theta} \frac{\partial p(\hat{\theta}|\theta_0)}{\partial \theta_0} = \int d\hat{\theta} \frac{\partial \ln p(\hat{\theta}|\theta_0)}{\partial \theta_0} p(\hat{\theta}|\theta_0) = 0 . \quad (\text{B.34})$$

Furthermore, differentiating $E[\hat{\theta}]$ with respect to θ_0 gives

$$\frac{\partial E[\hat{\theta}]}{\partial \theta_0} = \frac{\partial}{\partial \theta_0} \int d\hat{\theta} \hat{\theta} p(\hat{\theta}|\theta_0) = \int d\hat{\theta} \hat{\theta} \frac{\partial \ln p(\hat{\theta}|\theta_0)}{\partial \theta_0} p(\hat{\theta}|\theta_0) . \quad (\text{B.35})$$

Since $E[\hat{\theta}]$ is a constant, one can multiply Equation (B.34) with it to obtain zero and subtract it from Equation (B.35). This gives

$$\frac{\partial E[\hat{\theta}]}{\partial \theta_0} = \int d\hat{\theta} [\hat{\theta} - E[\hat{\theta}]] \frac{\partial \ln p(\hat{\theta}|\theta_0)}{\partial \theta_0} p(\hat{\theta}|\theta_0) \quad (\text{B.36})$$

An application of Schwarz' inequality yields

$$\int d\hat{\theta} [\hat{\theta} - E[\hat{\theta}]]^2 \int d\hat{\theta} \left[\frac{\partial \ln p(\hat{\theta}|\theta_0)}{\partial \theta_0} \right]^2 p(\hat{\theta}|\theta_0) \geq \left(\frac{\partial E[\hat{\theta}]}{\partial \theta_0} \right)^2 \quad (\text{B.37})$$

The first expression is the definition of the variance $V[\hat{\theta}]$. The second expression is the Fisher information (Equation (B.40))

$$\int d\hat{\theta} \left[\frac{\partial \ln p(\hat{\theta}|\theta_0)}{\partial \theta_0} \right]^2 p(\hat{\theta}|\theta_0) = E \left[\left(\frac{\partial \ln p(\hat{\theta}|\theta_0)}{\partial \theta_0} \right)^2 \right] = I(\hat{\theta}). \quad (\text{B.38})$$

Hence, this gives

$$V[\hat{\theta}] \geq \frac{\partial E[\hat{\theta}]/\partial \theta_0}{I(\hat{\theta})} = \frac{(1 + \partial b/\partial \theta_0)^2}{I(\hat{\theta})}, \quad (\text{B.39})$$

where the last step comes from the definition of the bias. This proves Equation (B.32) [66, 67].

B.4 Fisher information

The Fisher information is a key quantity in statistical inference that quantifies the amount of information a random variable x carries about an unknown parameter θ [67].

For a given PDF $p(x|\theta)$, the Fisher information $I(\theta)$ is defined as the expected value of the squared derivative of the log-likelihood function with respect to the parameter θ ,

$$I(\theta) = E \left[\left(\frac{\partial}{\partial \theta} \log p(x|\theta) \right)^2 \right] = \int dx \left(\frac{\partial}{\partial \theta} \log p(x|\theta) \right)^2 p(x|\theta). \quad (\text{B.40})$$

For multiple parameters, the Fisher information takes the matrix form

$$\mathbf{I}_{ij}(\boldsymbol{\theta}) = E \left[\left(\frac{\partial}{\partial \theta_i} \log p(x|\boldsymbol{\theta}) \right) \left(\frac{\partial}{\partial \theta_j} \log p(x|\boldsymbol{\theta}) \right) \right]. \quad (\text{B.41})$$

The Fisher information provides a measure of the sensitivity of the likelihood function with respect to changes in θ . A high Fisher information value indicates that small changes in θ result in large changes in the likelihood, implying that the data provide more information about θ . Conversely, a low Fisher information value suggests that the likelihood is relatively insensitive to changes in θ , indicating less information about the parameter.

B.5 Proof of consistency of maximum likelihood estimators

To prove the consistency of the MLE, Jensen's inequality is applied to the strictly concave logarithm function [66]:

$$E \left[\ln \left(\frac{p(x|\boldsymbol{\theta})}{p(x|\boldsymbol{\theta}_0)} \right) \right] < \ln \left(E \left[\frac{p(x|\boldsymbol{\theta})}{p(x|\boldsymbol{\theta}_0)} \right] \right) = 0 \quad \forall \boldsymbol{\theta} \neq \boldsymbol{\theta}_0. \quad (\text{B.42})$$

The last step follows from

$$E \left[\frac{p(\mathbf{x}|\boldsymbol{\theta})}{p(\mathbf{x}|\boldsymbol{\theta}_0)} \right] = \int dx p(\mathbf{x}|\boldsymbol{\theta}_0) \frac{p(\mathbf{x}|\boldsymbol{\theta})}{p(\mathbf{x}|\boldsymbol{\theta}_0)} = 1 . \quad (\text{B.43})$$

Hence, from Equation (B.42),

$$E [\ln p(\mathbf{x}|\boldsymbol{\theta}_0)] > E [\ln p(\mathbf{x}|\boldsymbol{\theta})] \quad \forall \boldsymbol{\theta} \neq \boldsymbol{\theta}_0 . \quad (\text{B.44})$$

Consider now two solutions to Equation (5.1), one inconsistent, $\hat{\boldsymbol{\theta}}'$, and one consistent, $\hat{\boldsymbol{\theta}}$. Then from the law of large numbers (Equation (B.67)),

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \ln p(x_i|\hat{\boldsymbol{\theta}}) = E[\ln p(\mathbf{x}|\boldsymbol{\theta}_0)] > E[\ln p(\mathbf{x}|\hat{\boldsymbol{\theta}}')] \quad (\text{B.45})$$

Hence, the consistent solution will asymptotically correspond to the absolute maximum of the likelihood. Note that this proof does not require differentiability of the likelihood. This means that the MLE is consistent, even if not found through Equation (5.2). However, if the likelihood is differentiable and the estimator found through Equation (5.2), it is consistent, but not necessarily unique [66, 67].

B.6 Proof of asymptotic normality of the maximum likelihood estimator

To prove the asymptotic normality of the MLE, the approach from Reference [66] is followed. The condition is enforced that the log-likelihood $\ln p(\mathbf{x}|\theta)$ is twice differentiable in a neighbourhood of θ_0 ¹. By the property of consistency, just discussed above, there exists (at least) one parameter point $\hat{\theta}$ where

$$\partial_{\theta} \ln p(\mathbf{x}|\theta)|_{\hat{\theta}} = \left. \frac{\partial \ln p(\mathbf{x}|\theta)}{\partial \theta} \right|_{\hat{\theta}} = 0 , \quad (\text{B.46})$$

such that $\lim_{N \rightarrow \infty} \hat{\theta} = \theta_0$.

Expanding the negative log-likelihood around a point θ^* which lies between θ_0 and $\hat{\theta}$,

$$-\partial_{\theta} \ln p(\mathbf{x}|\theta)|_{\theta_0} = \partial_{\theta} \ln p(\mathbf{x}|\theta)|_{\hat{\theta}} - \partial_{\theta} \ln p(\mathbf{x}|\theta)|_{\theta_0} = (\hat{\theta} - \theta_0) \partial_{\theta}^2 \ln p(\mathbf{x}|\theta)|_{\theta^*} . \quad (\text{B.47})$$

Hence,

$$\hat{\theta} - \theta_0 = - \frac{\partial_{\theta} \ln p(\mathbf{x}|\theta)|_{\theta_0}}{\partial_{\theta}^2 \ln p(\mathbf{x}|\theta)|_{\theta^*}} , \quad (\text{B.48})$$

which can be written as

$$\sqrt{NI(\theta_0)}(\hat{\theta} - \theta_0) = \left(\frac{\partial_{\theta} \ln p(\mathbf{x}|\theta)|_{\theta_0}}{\sqrt{NI(\theta_0)}} \right) \left(\frac{\partial_{\theta}^2 \ln p(\mathbf{x}|\theta)|_{\theta_0}}{\partial_{\theta}^2 \ln p(\mathbf{x}|\theta)|_{\theta^*}} \right) \left(\frac{\partial_{\theta}^2 \ln p(\mathbf{x}|\theta)|_{\theta_0}}{NI(\theta_0)} \right)^{-1} . \quad (\text{B.49})$$

¹This proof is presented for the case of one parameter θ , but it is easily extendable to multiple parameters.

For the first term on the right-hand side of Equation (B.49) the central limit theorem (Equation (B.68)) can be applied. It is known that

$$E_{\theta_0} [\partial_\theta \ln p(\mathbf{x}|\theta)] = E [\partial_\theta \ln p(\mathbf{x}|\theta)] |_{\theta_0} = 0 , \quad (\text{B.50})$$

and

$$V_{\theta_0} [\partial_\theta \ln p(\mathbf{x}|\theta)] = E_{\theta_0} [(\partial_\theta \ln p(\mathbf{x}|\theta))^2] |_{\theta_0} = I(\theta_0) . \quad (\text{B.51})$$

Hence, it follows that

$$\frac{\partial_\theta \ln p(\mathbf{x}|\theta)|_{\theta_0}}{\sqrt{NI(\theta_0)}} = \frac{\sqrt{N}}{\sqrt{I(\theta_0)}} \sum_i \partial_\theta \ln p(x_i|\theta) \rightarrow^d \mathcal{N}(0, 1) . \quad (\text{B.52})$$

The second term on the right-hand side of Equation (B.49) approaches unity in the large sample limit,

$$\frac{\partial_\theta^2 \ln p(\mathbf{x}|\theta)|_{\theta_0}}{\partial_\theta^2 \ln p(\mathbf{x}|\theta)|_{\theta^*}} \rightarrow 1 , \quad (\text{B.53})$$

by the consistency condition $\hat{\theta} \rightarrow \theta_0$ and the fact that θ^* lies between θ_0 and $\hat{\theta}$.

The third term on the right-hand side of Equation (B.49) also approaches unity in the large sample limit, by the law of large numbers (Equation (B.67)).

Hence, it follows that

$$\sqrt{NI(\theta_0)}(\hat{\theta} - \theta_0) \rightarrow^d \mathcal{N}(0, 1) , \quad (\text{B.54})$$

which is equivalent to Equation (5.3).

B.7 Proof of the Neyman-Pearson lemma

To prove the Neyman-Pearson lemma, the method of Lagrange multipliers can be adapted [66]. The goal is to maximize the power of the test, $1 - \beta(t)$, subject to the constraint $\alpha(t) \leq \alpha$. This constraint ensures that the Type I error probability does not exceed the chosen significance level α .

The Lagrangian to maximize is

$$\begin{aligned} L(t_{cut}, \lambda) &= (1 - \beta(t_{cut})) - \frac{1}{\lambda}(\alpha(t_{cut}) - \alpha) \\ &= \int_{t_{cut}}^{\infty} p(t|H_1)dt - \frac{1}{\lambda} \left(\int_{t_{cut}}^{\infty} p(t|H_0)dt - \alpha \right) , \end{aligned}$$

for a positive-valued λ . This can be written as a function of \mathbf{x} , integrating over a rejection region $\mathbf{x} \in \mathcal{R}$,

$$\begin{aligned} L(\mathbf{x}, \lambda) &= \int_{\mathcal{R}} p(\mathbf{x}|H_1)d^n \mathbf{x} - \frac{1}{\lambda} \left(\int_{\mathcal{R}} p(\mathbf{x}|H_0)d^n \mathbf{x} - \alpha \right) \\ &= \int_{\mathcal{R}} \left(p(\mathbf{x}|H_1) - \frac{1}{\lambda} p(\mathbf{x}|H_0) \right) d^n \mathbf{x} - \frac{1}{\lambda} \alpha . \end{aligned}$$

To maximize this integral, the rejection region must be defined such that

$$\mathcal{R}(\lambda) = \left\{ \mathbf{x} \mid p(\mathbf{x}|H_1) > \frac{1}{\lambda}p(\mathbf{x}|H_0) \right\}, \quad (\text{B.55})$$

or, rearranging, this defines a rejection region where

$$\frac{p(\mathbf{x}|H_0)}{p(\mathbf{x}|H_1)} < \lambda. \quad (\text{B.56})$$

The value of λ is determined by the upper bound of the constraint $\alpha(t) \leq \alpha$,

$$[\alpha(\mathbf{x})]_{\mathcal{R}(\lambda)} = \alpha. \quad (\text{B.57})$$

For the proof of the Neyman-Pearson lemma, consider now an alternative region \mathcal{R}' . Since $p(\mathbf{x}|H_1) > \frac{1}{\lambda}p(\mathbf{x}|H_0)$ in all of \mathcal{R} , any deviation to \mathcal{R}' will contain parts where $p(\mathbf{x}|H_1) < \frac{1}{\lambda}p(\mathbf{x}|H_0)$. Hence, the inequality can be defined as

$$\left[(1 - \beta(\mathbf{x})) - \frac{1}{\lambda}\alpha(\mathbf{x}) \right]_{\mathcal{R}} \geq \left[(1 - \beta(\mathbf{x})) - \frac{1}{\lambda}\alpha(\mathbf{x}) \right]_{\mathcal{R}'}. \quad (\text{B.58})$$

Rearranging this inequality gives

$$[1 - \beta(\mathbf{x})]_{\mathcal{R}} - [1 - \beta(\mathbf{x})]_{\mathcal{R}'} \geq \frac{1}{\lambda} ([\alpha(\mathbf{x})]_{\mathcal{R}} - [\alpha(\mathbf{x})]_{\mathcal{R}'}) \geq 0, \quad (\text{B.59})$$

where the last equality uses the fact that $[\alpha(\mathbf{x})]_{\mathcal{R}} = \alpha$ and $[\alpha(\mathbf{x})]_{\mathcal{R}'} \leq \alpha$. Hence, it follows that

$$[1 - \beta(\mathbf{x})]_{\mathcal{R}} \geq [1 - \beta(\mathbf{x})]_{\mathcal{R}'} \quad (\text{B.60})$$

showing that the rejection region defined by the likelihood ratio results in the most powerful test. This proves the Neyman-Pearson lemma [66].

B.8 Likelihood ratio test statistic for normally distributed data

This section shows the asymptotic distribution of the likelihood ratio test statistic for the case of normally distributed data. This proves to be useful for Wilks' theorem [86] and goodness-of-fit testing.

For example, for a Gaussian likelihood of N data points \mathbf{x} with means $\boldsymbol{\mu}$ and variances $\boldsymbol{\sigma}^2$,

$$p(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\sigma}^2) = \prod_{i=0}^N \mathcal{N}(x_i|\mu_i, \sigma_i^2) = \prod_{i=0}^N \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{(x_i - \mu_i)^2}{2\sigma_i^2}}. \quad (\text{B.61})$$

The likelihood ratio test statistic is then simply

$$t = \sum_{i=0}^N \frac{(x_i - \mu_i)^2}{\sigma_i^2}. \quad (\text{B.62})$$

One might recognize this term as the classic χ^2 variable. So what is the **PDF** of the test statistic in this case? For one data-point

$$p(x_i|\mu_i, \sigma_i^2) = \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{(x_i-\mu_i)^2}{2\sigma_i^2}}. \quad (\text{B.63})$$

The distribution of the test statistic is

$$p(t_i) = p(x_i|\mu_i, \sigma_i^2) \left| \frac{dx_i}{dt_i} \right| = \frac{1}{\sqrt{2\pi t_i}} e^{-t_i/2}, \quad (\text{B.64})$$

where an extra factor of 2 comes from the change in integration limits. This is the χ^2 -distribution for 1 **d.o.f.**. The characteristic function² of this **PDF** is

$$\phi_i(k) = \int_0^\infty dt_i p(t_i) e^{ikt_i} = (1 - 2ik)^{-1/2}. \quad (\text{B.65})$$

Under the assumption of independent data points, the characteristic function of the sum $t = \sum_{i=0}^N t_i$ is the product of the individual characteristic functions

$$\phi(k) = \prod_{i=0}^N \phi_i(k) = (1 - 2ik)^{-N/2}, \quad (\text{B.66})$$

which is the characteristic function of a χ^2 -distribution with N **d.o.f.**. If the system parametrizes $\mu(x|\boldsymbol{\theta})$ with m parameters and infers these parameters from the data, this results in losing m independent pieces of information and the resulting test statistic **PDF** will be a χ^2 -distribution with $N - m$ **d.o.f.**. This is an important example, since from the central limit theorem in [Equation \(B.68\)](#) it is known that in the asymptotic limit, all parameters become normally distributed [65].

The above calculation proves useful in two cases:

- **Wilks' theorem:** For one **POI**, in the large sample limit where the Wald approximation of [Equation \(5.21\)](#) holds and $\hat{\mu}$ is normally distributed, the test statistic t is distributed as a χ^2 -distribution with 1 **d.o.f.** [86].
- **Goodness-of-fit test:** In this case, the asymptotic distribution of the test statistic is a χ^2 -distribution with $N - m$ **d.o.f.** [65].

B.9 Convergence of random variables

In the context of parameter estimation and statistical inference, two foundational results in probability theory are particularly important: the *Law of large numbers* and the *Central limit theorem*. These theorems provide insight into the behaviour of sample-based estimators in the asymptotic limit of large sample sizes. They are key for proving fundamental properties of *maximum likelihood estimators* (see [Section 5.1.1](#)).

²The Fourier transform of the **PDF**.

Law of large numbers Consider a sample $\mathbf{x} = (x_1, \dots, x_N)$ drawn from a PDF $p(x)$. The law of large numbers states that for any function $a(x)$ with finite variance $V[a(x)]$, the sample average converges to the expectation value in the limit of large N [67],

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N a(x_i) = E[a(x)] = \int dx a(x)p(x). \quad (\text{B.67})$$

Central limit theorem This theorem states that for a sample $\mathbf{x} = (x_1, \dots, x_N)$ of independent and identically distributed random variables with finite mean $\mu = E[x_i]$ and variance $\sigma^2 = V[x_i]$, $\frac{\sqrt{N}}{\sigma}(\bar{x} - \mu)$ converges in distribution to the standard normal distribution $\mathcal{N}(0, 1)$ as N becomes large,

$$\frac{\sqrt{N}}{\sigma}(\bar{x} - \mu) \rightarrow^d \mathcal{N}(0, 1) \quad (\text{B.68})$$

where $\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$ is the sample mean [65, 66]. This theorem forms the basis for the use of normal approximations in hypothesis testing and confidence interval construction.

The proof of the central limit theorem can be shown through the use of characteristic functions. The characteristic function $\phi_x(k)$ of a PDF $p(x)$ is the expectation value of e^{ikx} , which is the Fourier transform of the PDF [65, 67],

$$\phi_x(k) = E_x [e^{ikx}] = \int dx e^{ikx} p(x). \quad (\text{B.69})$$

Consider the definition of the variable

$$Z_N = \frac{\bar{x} - \mu}{\sigma/\sqrt{N}} = \sum_{i=1}^N \frac{x_i - \mu}{\sqrt{N}\sigma} = \sum_{i=1}^N \frac{1}{\sqrt{N}} Y_i, \quad (\text{B.70})$$

with $Y_i = (x_i - \mu)/\sigma$. The expectation value of Y_i is $E_x[Y_i] = 0$ by the law of large numbers (Equation (B.67)), and $E_x[Y_i^2] = 1$. The proof proceeds by showing that the characteristic function of Z_N is equivalent to the one of a standard normal, with mean 0 and variance 1, in the limit of $N \rightarrow \infty$.

The characteristic function of Z_N is

$$\phi_{Z_N}(k) = \prod_{i=1}^N \phi_{Y_i/\sqrt{N}}(k) = \prod_{i=1}^N \phi_{Y_i} \left(\frac{k}{\sqrt{N}} \right), \quad (\text{B.71})$$

using the property $\phi_{ax}(k) = \phi_x(ak)$ for some constant a . Since all x_i are equally distributed,

$$\phi_{Z_N}(k) = \left(\phi_{Y_i} \left(\frac{k}{\sqrt{N}} \right) \right)^N. \quad (\text{B.72})$$

The characteristic function $\phi_{Y_i}(k/\sqrt{N})$ can be expanded for large N as

$$\phi_{Y_i}\left(\frac{k}{\sqrt{N}}\right) = \sum_{m=0}^{\infty} \frac{k^m}{m!} \left. \frac{\partial^m \phi_{Y_i}(k/\sqrt{N})}{\partial (k/\sqrt{N})^m} \right|_{k/\sqrt{N}=0} = 1 - \frac{k^2}{2N} + o\left(\frac{k^2}{N}\right), \quad (\text{B.73})$$

where $o(k^2/N)$ refers to all terms which approach zero faster than k^2/N as $N \rightarrow \infty$.

Neglecting all higher order terms, this gives

$$\lim_{N \rightarrow \infty} \phi_{Z_N}(k) = \lim_{N \rightarrow \infty} \left(1 - \frac{k^2}{2N}\right) = e^{-\frac{k^2}{2}}, \quad (\text{B.74})$$

which is the characteristic function of a standard normal distribution [65, 67].

B.10 Posterior validation

The reliability of Bayesian posterior inference critically depends on the convergence and mixing properties of the **MCMC** sampling. The quality of **MCMC** chains is validated following established best practices [68, 137, 138].

B.10.1 Potential scale reduction factor

The potential scale reduction factor \hat{R} quantifies convergence by comparing between-chain and within-chain variance [68]. The improved split- \hat{R} diagnostic [137] is used, which splits each chain in half to increase sensitivity to non-stationarity. For parameter θ , split- \hat{R} is defined as:

$$\hat{R} = \sqrt{\frac{\widehat{\text{Var}}^+(\theta|\mathbf{x})}{W}} \quad (\text{B.75})$$

where $\widehat{\text{Var}}^+(\theta|\mathbf{x}) = \frac{n-1}{n}W + \frac{1}{n}B$ combines within-chain variance W and between-chain variance B and n is the number of draws per chain. Values $\hat{R} < 1.01$ indicate excellent convergence, while $\hat{R} \geq 1.1$ suggests insufficient sampling [137].

B.10.2 Effective sample size

The effective sample size (ESS) accounts for autocorrelation within chains, quantifying how many independent samples the correlated **MCMC** output is equivalent to. Both bulk ESS (for central quantiles) and tail ESS (for extreme quantiles) are computed [137]. Following recommendations, $\text{ESS} > 400$ is required for reliable inference, with $\text{ESS} > 1000$ preferred for high precision [137].

B.10.3 Visual diagnostics

Trace plots display parameter evolution across iterations, providing assessment of chain behaviour. Well-mixed chains exhibit rapid fluctuation around a stable mean without systematic trends or chain separation [138].

Rank plots offer sensitivity for detecting mixing problems. For each parameter, all samples across chains are ranked, then histograms of ranks per chain are plotted. Uniform rank distributions indicate proper mixing, while systematic patterns reveal convergence issues [137].

Appendix C

Model-agnostic likelihoods for the reinterpretation of particle physics results

C.1 Individual contributions

The work presented in [Chapter 7](#) is a joint effort of Lorenz Gärtner (L. G.), Nikolai Krug (N. K.; former Nikolai Hartmann), Lukas Heinrich (L. H.), Malin Horstmann (M. H.), Thomas Kuhr (T. K.), M eril Reboud (M. R.), Slavomira Stefkova (S. S.), and Danny van Dyk (D. v. D.).

The initial idea for the method was developed by L. G. as an extension of the “brief idea” proposed by L. H. in Reference [\[139\]](#).

The idea for the toy study was developed by L. G. and S. S., based on the plan to apply it to the $B^+ \rightarrow K^+ \nu \bar{\nu}$ analysis at Belle II [\[2\]](#).

The implementation of the method was done by L. G. with the guidance of N. K., while L. H. provided support with the custom modifier implementation. The code is available at this [repository](#) [\[83\]](#), and contributions can be traced back through the commit history. In the `examples` folder, one can find the study presented in this work.

Theoretical input is provided by M. R. and D. v. D., who also developed the EOS package [\[47, 115\]](#) used to obtain the theoretical predictions. The implementation of [Equation \(2.62\)](#) in EOS was done by L. G. with the help of M. R. and D. v. D.

The statistical inference is performed using `pyhf` [\[71, 72\]](#), which was developed by L. H. The Bayesian `pyhf` package [\[74\]](#) used for the Bayesian inference in this work was developed by Matthew Feickert, L. H., and M. H.

T. K. provided general guidance and support throughout the project.

C.2 Optimizing the kinematic binning

To obtain suggestions on the number of bins to use for the kinematic variable(s), one can follow the following procedure:

For a large set of models covering your parameter space as thoroughly as possible, compute the expected yields for a finer and finer binning in the kinematic **d.o.f.** (a new joint number density and new weights need to be computed every time). At each step, compute the difference to the results of the previous step and stop when reaching a predefined convergence condition. The maximum number of bins over all the looped models should give a good estimate of the number of kinematic bins to use.

C.3 A recipe for application of this reinterpretation method

To assist with easy application of this reinterpretation method to any analysis, I provide a simple 4-step guide on what needs to be done to reinterpret a result from **HEP**. This guide focuses on the discrete approach of [Section 7.1.1](#).

1. *Samples.* Gather your post-reconstruction samples and ensure that they contain information on all kinematic **d.o.f.** as well as the reconstruction variable.
2. *Null joint number density.* From these samples, build the null joint number density by binning samples in bins of the reconstruction variable times the kinematic **d.o.f.** (see [Appendix C.2](#) on how to optimize the kinematic binning).
3. *Weights.* Identify your null prediction used for producing the original **MC** samples. Choose your alternative theoretical prediction(s) and ensure that the support of the null distribution covers the full range of the alternative distribution (this can also be done by setting an upper bound on the weights). Compute the weights as the ratio of the bin-integrated alternative to the bin-integrated null distribution (as in [Equation \(7.6\)](#)).
4. *Inference.* Either by making use of the code in [\[83\]](#) or by implementing [Equation \(7.7\)](#), compute the expected yields given the alternative prediction, making use of the joint number density and the computed weights. Using either **pyhf** [\[71, 72\]](#) or alternative tools, statistical inference can be used to compute results for the alternative theory.

Appendix D

A model-agnostic likelihood for the Belle II $B^+ \rightarrow K^+ \nu \bar{\nu}$ measurement

D.1 Individual contributions

The work presented in [Chapters 8](#) and [9](#) is a joint effort by the Belle II collaboration. The main proponents of this work are Lorenz Gärtner (L. G.), Nikolai Krug (N. K.; former Nikolai Hartmann), Thomas Kuhr (T. K.), and Slavomira Stefkova (S. S.).

The initial idea for the project was developed by T. K. and S. S. The details of the analysis were discussed and refined by L. G., with guidance from N. K., T. K. and S. S.

The method was implemented by L. G., with guidance from N. K., T. K. and S. S.

Valuable input on the analysis and the corresponding publication [[116](#)] was provided by all members of the Belle II collaboration. Particular gratitude goes to Swagato Banerjee, Gaetano de Marino, Sasha Glazov, Eldar Ganiev, Enrico Graziani, Romulus Godang, Bob Kowalewski, James Libby, Elisa Manoni, Mikihiro Nakao, Leo Piilonen, Soeren Prell, Yoshi Sakai, Alan Schwartz, Wei Shan, Diego Tonelli, Phillip Urquijo, Roberta Volpe, and Bruce Yabsley, who provided detailed comments and were particularly active during the review process of this work.

D.2 Final-state radiation

Final-state radiation potentially changes the q^2 spectrum, depending on how q^2 is defined. If no final-state radiation is present

$$q^2 = (p_\nu + p_{\bar{\nu}})^2 = (p_B - p_K)^2. \quad (\text{D.1})$$

Otherwise, the two definitions differ, $(p_\nu + p_{\bar{\nu}})^2 \neq (p_B - p_K)^2$, and consequently their distributions will change slightly. The comparison is shown in [Figure D.1](#), where final-state radiation was included using PHOTOS [[140](#)]. Differences in the distributions are at

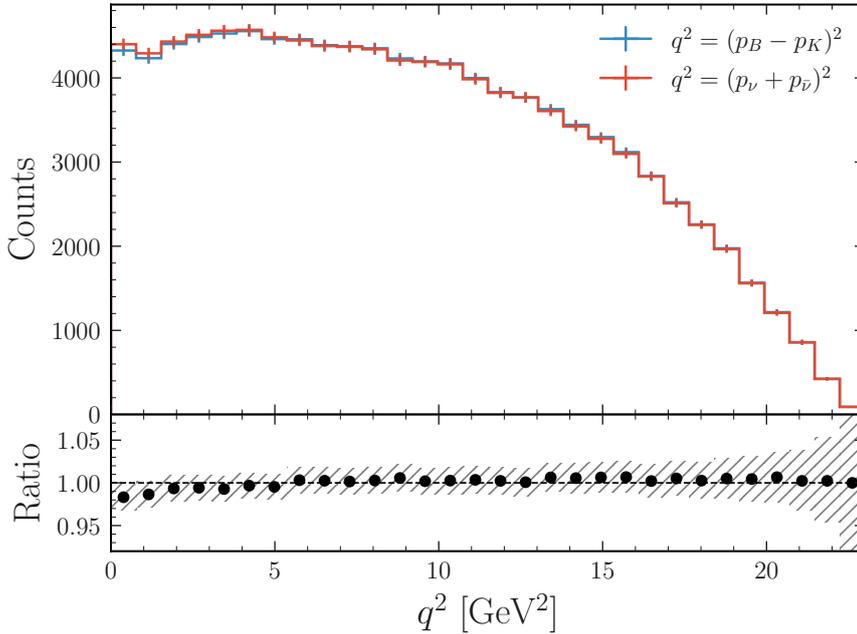


Figure D.1: The generator level q^2 distributions for a set of 100 000 simulated MC samples, including final-state radiation.

most around 2% in the lowest q^2 bins. The higher q^2 bins are less affected by final-state radiation.

As a comparison, the form factor uncertainties constitute an uncertainty of around 6% in the low q^2 region. Given the small differences in the distributions, depending on the two definitions, I neglect this effect in the reinterpretation study.

D.3 Kinematic q^2 binning

When creating the number densities, the binning of the kinematic variables is an important choice. To enable reinterpretation studies with narrow resonances, I consider relatively fine binnings and test options of 50, 100, and 250 bins in the kinematic range $q^2 \in [0, 22.9]$ GeV², corresponding to bin widths of 0.45, 0.23, and 0.09 GeV², respectively. The resulting best-fit values of the POIs are summarised in Table D.1. For all three binning choices, the best-fit values agree within uncertainties. The uncertainty in $C_{\text{SL}} + C_{\text{SR}}$ is difficult to estimate due to limited sensitivity and is therefore expected to fluctuate. Since potential use cases may require rather fine binning, I adopt 100 bins for this reinterpretation study.

Table D.1: Best-fit values for the **WET** and $B^+ \rightarrow K^+ X$ reinterpretation, under different q^2 binning scenarios.

| Model | Parameters | 50 q^2 bins | 100 q^2 bins | 250 q^2 bins |
|------------------------------|---------------------------------|----------------|----------------|----------------|
| WET | $C_{\text{VL}} + C_{\text{VR}}$ | 11.4 ± 3.2 | 11.4 ± 3.2 | 11.4 ± 3.2 |
| | $C_{\text{SL}} + C_{\text{SR}}$ | 0.1 ± 15.8 | 0.1 ± 13.7 | 0.2 ± 10.0 |
| | C_{TL} | 8.4 ± 3.0 | 8.4 ± 3.0 | 8.4 ± 3.0 |
| $B^+ \rightarrow K^+ X$ | μ_X | 9.1 ± 2.6 | 9.0 ± 2.6 | 9.0 ± 2.6 |
| $\Gamma_X = 0.1 \text{ GeV}$ | $m_X \text{ [GeV]}$ | 2.1 ± 0.1 | 2.1 ± 0.1 | 2.1 ± 0.1 |

D.4 Impact of removing unphysical q^2 values in the hadronic tagging analysis

In [Section 8.2](#) I discussed that 0.16% of events in the **HTA** analysis have an unphysical generator-level q^2 assigned. In [Table D.2](#) I summarise the best-fit values of the **POIs** with and without these unphysical samples. Since the impact on the best-fit point is negligible, I decide to keep the unphysical samples in the analysis without reweighting them in the reinterpretation process.

Table D.2: Best-fit values for a model with all samples and one with the unphysical q^2 events removed.

| Model | Parameters | All samples | Unphys. q^2 removed |
|------------------------------|---------------------------------|----------------|-----------------------|
| SM | μ | 4.6 ± 1.3 | 4.6 ± 1.3 |
| WET | $C_{\text{VL}} + C_{\text{VR}}$ | 11.4 ± 3.2 | 11.4 ± 3.2 |
| | $C_{\text{SL}} + C_{\text{SR}}$ | 0.1 ± 13.7 | 0.1 ± 16.9 |
| | C_{TL} | 8.4 ± 3.0 | 8.4 ± 3.0 |
| $B^+ \rightarrow K^+ X$ | μ_X | 9.0 ± 2.6 | 9.0 ± 2.6 |
| $\Gamma_X = 0.1 \text{ GeV}$ | $m_X \text{ [GeV]}$ | 2.1 ± 0.1 | 2.1 ± 0.1 |

Appendix E

A Weak Effective Theory reinterpretation of the Belle II $B^+ \rightarrow K^+ \nu \bar{\nu}$ measurement

E.1 Belle II HEPData reinterpretation inventory

To enable reinterpretation under any new physics model with the model-agnostic likelihood [34], the necessary information from Belle II will be published on [HEPData](#) [122, 123]. The release will include the following components:

- The **SM** $B^+ \rightarrow K^+ \nu \bar{\nu}$ differential branching ratio as a function of q^2 ;
- Signal selection efficiency as a function of q^2 ;
- Binned joint number densities:
 - **ITA**: x-axis: q^2 , y-axis: $q_{\text{rec}}^2 \times \eta(\text{BDT}_2)$ (flattened), z-axis: events (weighted);
 - **HTA**: x-axis: q^2 , y-axis: $\eta(\text{BDTh})$ (flattened), z-axis: events (weighted);
- **pyhf** combined likelihood in JSON format:
 - Containing templates for signal and background after all selections, binned in $q_{\text{rec}}^2 \times \eta(\text{BDT}_2)$ (**ITA**) and $\eta(\text{BDTh})$ (**HTA**);
- The code to reproduce the **WET** reinterpretation results obtained in this analysis.

E.2 Weak Effective Theory posterior validation

To validate the **MCMC** sampling, I apply the diagnostics described in [Appendix B.10](#) to the **WET** posterior samples discussed in [Section 9.2](#) ([Figure 9.1](#)). I find $\hat{R} = 1.000$ for all parameters, indicating complete convergence across the chains, with bulk ESS > 2400 and tail ESS > 3100 for all parameters, ensuring high-precision inference.

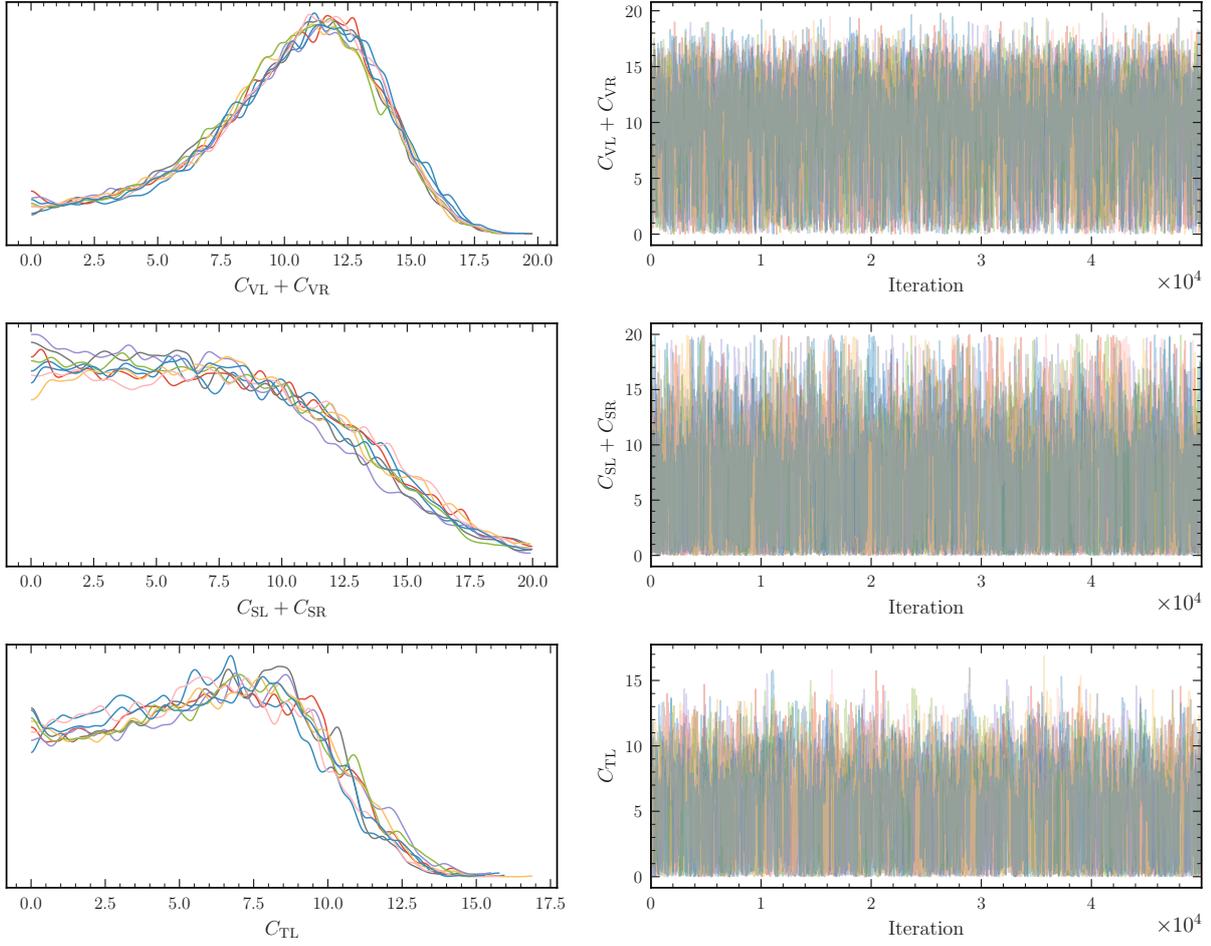


Figure E.1: Chain-wise marginal posterior distributions (*left*) and trace plots (*right*) for the WET Wilson coefficients in Equation (2.62), from the posterior discussed in Section 9.2 (Figure 9.1). Each colour represents a different MCMC chain. No trends or chain separation are visible.

Trace plots (see Figure E.1) show ideal “hairy caterpillar” behaviour with rapid mixing and no visible trends. Rank plots (see Figure E.2) display uniform distributions across all chains, confirming good mixing.

These results provide confidence that the posterior constraints presented in Table 9.1 are statistically robust and suitable for physics interpretation. The diagnostic validation ensures that reported credible intervals accurately reflect posterior uncertainty and that no additional sampling is required.

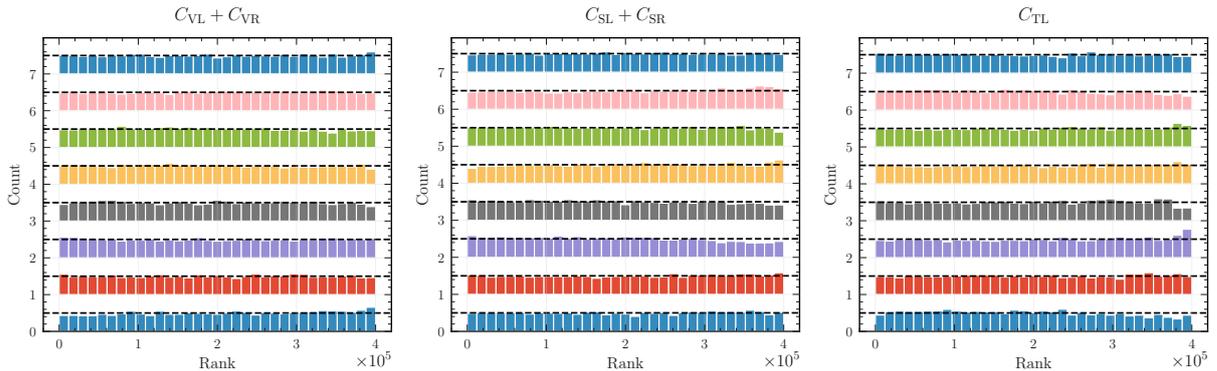


Figure E.2: Rank plots for the **WET** Wilson coefficients in Equation (2.62), from the posterior discussed in Section 9.2 (Figure 9.1). Each colour represents a different **MCMC** chain. The uniform distributions indicate good mixing.

E.3 Weak Effective Theory posteriors for the inclusive and hadronic tagging analyses

In Chapter 9, I presented the marginal posterior of the **WET** Wilson coefficients based on the combined **ITA** and **HTA** likelihood. Here, I present the marginal posteriors for the **ITA** and **HTA** analyses separately.

I adopt the same approach as in Section 9.2 but now only use the **ITA** or **HTA** likelihood, respectively. Priors are chosen to be uniform for all Wilson coefficients with support $[0, 20]$. I also use the same symmetrization technique as described in Section 9.2.

E.3.1 Marginal posterior for the inclusive tagging analysis

The 1- and 2-dimensional marginalizations of the posterior based on the **ITA** likelihood and the resulting credible intervals at 68% and 95% probability are shown in Figure E.3. Posterior mode and credible intervals for the absolute values of the Wilson coefficients are shown in Table E.1.

Table E.1: The mode of the posterior, and **HDI** at 68% and 95% for the (sums of the) **WET** Wilson coefficients in Equation (2.62), derived from the posterior in Figure E.3.

| Parameters | Mode | 68% HDI | 95% HDI |
|---------------------|------|-------------|-------------|
| $ C_{VL} + C_{VR} $ | 10.9 | [7.0, 14.7] | [1.0, 16.3] |
| $ C_{SL} + C_{SR} $ | 0.0 | [1.9, 14.3] | [0.0, 18.1] |
| $ C_{TL} $ | 11.8 | [4.9, 13.6] | [0.0, 14.5] |

There is a clear deviation from the **SM** prediction in the vector sector, as expected from

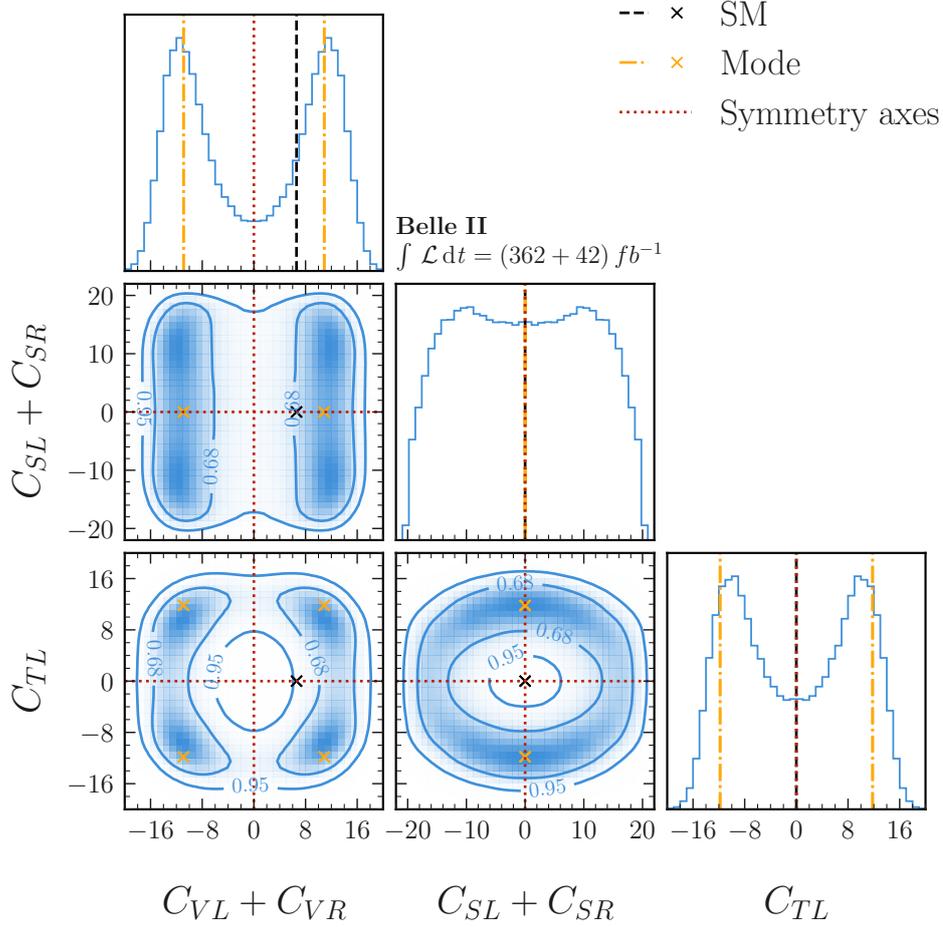


Figure E.3: The marginalized posterior for the Wilson coefficients in Equation (2.62), based on the ITA likelihood. This analysis adopts the convention that $C_{VL} + C_{VR}$, $C_{SL} + C_{SR}$, and C_{TL} are real-valued. Diagonal and off-diagonal panels show the 1-dimensional and 2-dimensional sample density PDFs on a linear scale, respectively. The overall scale is omitted, as all relevant information is contained in the shape of the distribution. The contours indicate 68% and 95% credible intervals. The dashed black lines and cross mark the SM point; the dash-dotted yellow lines and cross indicate the posterior mode; dotted red lines mark the symmetry axes used for sample symmetrization.

the result of the original **ITA** analysis, which reported a signal strength of $\mu = 5.4 \pm 1.5$ [2]. Since the Wilson coefficients enter quadratically into the differential branching ratio (see [Equation \(2.62\)](#)), one expects the posterior of the dominant vector Wilson coefficient to peak approximately at

$$|C_{\text{VL}} + C_{\text{VR}}| \simeq \sqrt{5.4} C_{\text{VL}}^{\text{SM}} \simeq 15.3, \quad (\text{E.1})$$

assuming negligible contributions from the other Wilson coefficients. The posterior indeed peaks below this value, with an additional non-zero contribution from the tensor sector.

Furthermore, the analysis finds that the posterior distribution peaks around a non-zero value for the tensor contribution. This indicates that a pure **SM** signal template does not provide the best description of the data, and the combination with a different kinematic distribution is preferred.

E.3.2 Marginal posterior for hadronic tagging analysis

The 1- and 2-dimensional marginalizations of the posterior based on the **HTA** likelihood and the resulting credible intervals at 68% and 95% probability are shown in [Figure E.4](#). Credible intervals for the absolute values of the Wilson coefficients are shown in [Table E.2](#).

Table E.2: The mode of the posterior, and **HDI** at 68% and 95% for the (sums of the) **WET** Wilson coefficients in [Equation \(2.62\)](#), derived from the posterior in [Figure E.4](#).

| Parameters | 68% HDI | 95% HDI |
|-----------------------------------|------------|-------------|
| $ C_{\text{VL}} + C_{\text{VR}} $ | [0.0, 8.3] | [0.0, 14.1] |
| $ C_{\text{SL}} + C_{\text{SR}} $ | [0.0, 8.6] | [0.0, 14.3] |
| $ C_{\text{TL}} $ | [0.0, 5.7] | [0.0, 9.7] |

In this case, the posterior peaks at zero for all (sums of) Wilson coefficients. This can be attributed to the choice of the binning variable, $\eta(\text{BDTh})$, which shows only a weak correlation with the kinematic variable q^2 (see [Section 8.2](#)). Therefore, the posterior mode is not a sensible measure. Consequently, the **HTA** analysis provides less constraining power on the individual Wilson coefficients than the **ITA** analysis, where the binning variable $\eta(\text{BDT}_2)$ exhibits a stronger correlation with q^2 (see [Section 8.2](#) and [Figure 8.5](#)).

Nonetheless, the **HTA** likelihood provides complementary information to the **ITA** likelihood, as it is sensitive to the overall normalization of the signal.

E.4 Goodness-of-fit

Here, I report the goodness-of-fit tests for the **SM** and background-only hypotheses. For details on the goodness-of-fit test, see [Section 5.3](#) and the paragraph around [Equation \(9.9\)](#). Here, more details on the results reported in [Table 9.4](#) are provided.

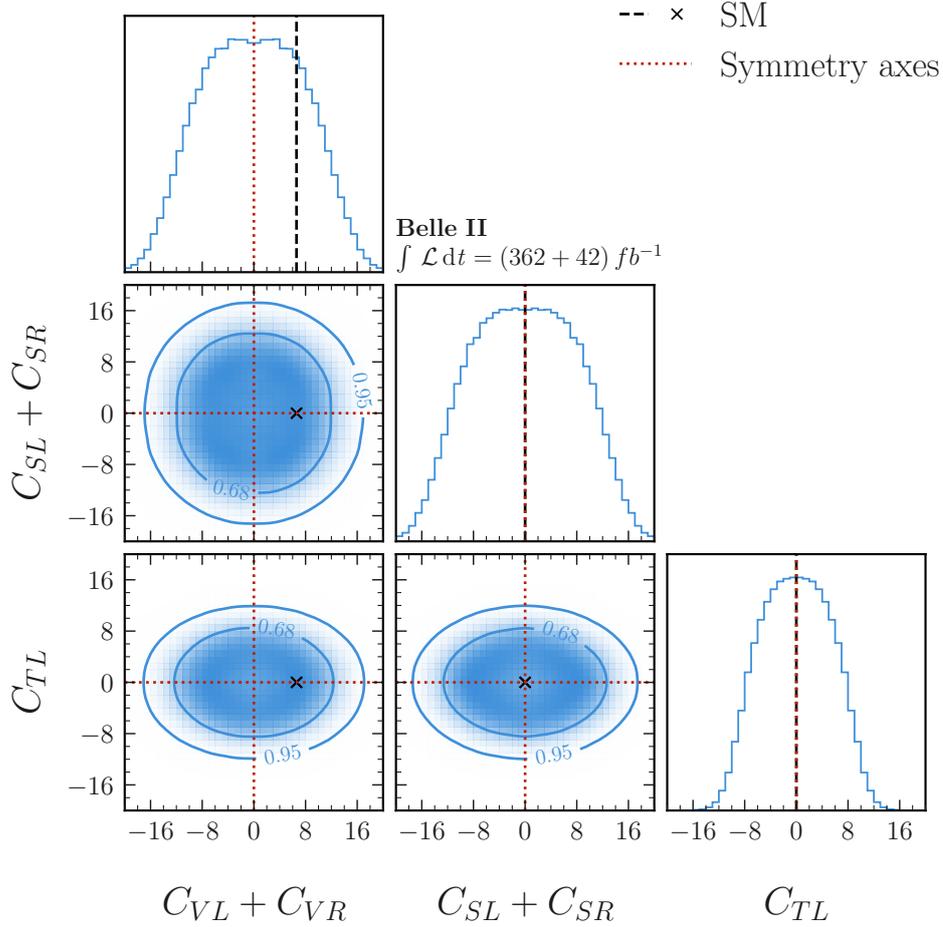


Figure E.4: The marginalized posterior for the Wilson coefficients in Equation (2.62), based on the HTA likelihood. This analysis adopts the convention that $C_{VL} + C_{VR}$, $C_{SL} + C_{SR}$, and C_{TL} are real-valued. Diagonal and off-diagonal panels show the 1-dimensional and 2-dimensional sample density PDFs on a linear scale, respectively. The overall scale is omitted, as all relevant information is contained in the shape of the distribution. The contours indicate 68% and 95% credible intervals. The dashed black lines and cross mark the SM point; dotted red lines mark the symmetry axes used for sample symmetrization.

The test statistic distributions are shown in [Figure E.5](#). The expected N_{dof} is obtained when fitting to the toy test statistic values. For the **SM** case, the expected $N_{\text{dof}} = 29$, as there is only the signal strength as a free parameter and 30 analysis bins. For the background-only hypothesis, there are no free parameters, and hence the expected $N_{\text{dof}} = 30$. The P -value for the **SM** hypothesis, $P = 0.58$, indicates a very good fit to the data. The P -value for the background-only hypothesis, $P = 0.13$, indicates an acceptable fit to the data.

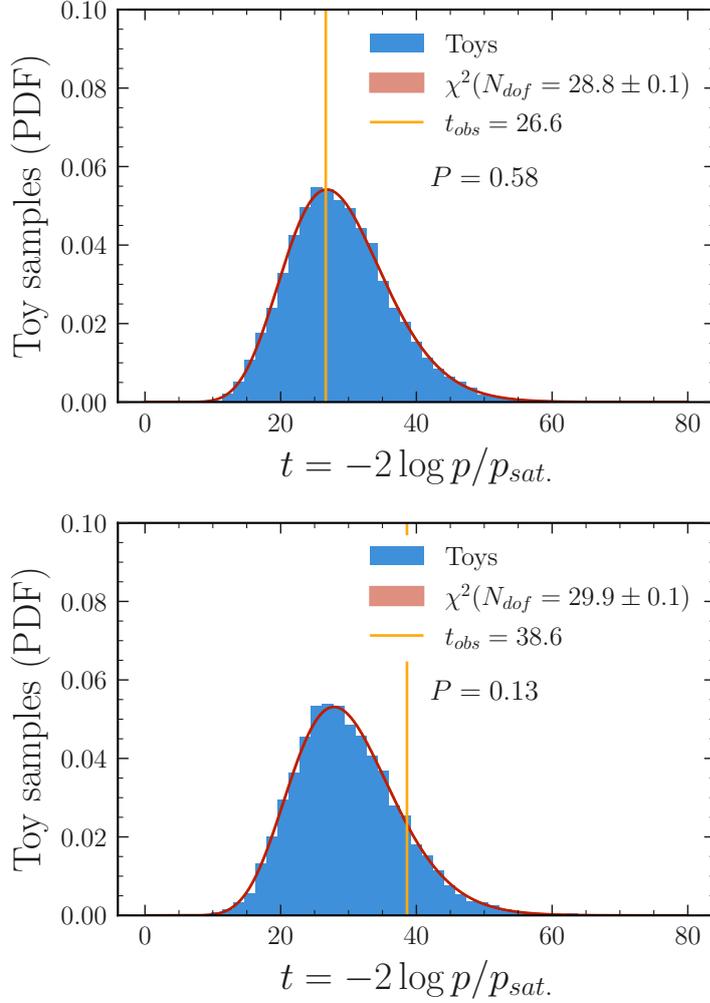


Figure E.5: The distribution of the goodness-of-fit test statistic for the SM (*top*) and background-only (*bottom*) hypotheses, each obtained from 10 000 toys. The values obtained from a fit to the data are shown as vertical yellow lines. The asymptotic χ^2 -distributions (red band) correspond to the best-fit N_{dof} , which are shown with the corresponding uncertainty band.

Appendix F

A light new physics reinterpretation of the Belle II $B^+ \rightarrow K^+ \nu \bar{\nu}$ measurement

F.1 $B^+ \rightarrow K^+ X$ posterior validation

To validate the **MCMC** sampling, I apply the diagnostics described in **Appendix B.10** to all $B^+ \rightarrow K^+ X$ posterior samples discussed in **Section 10.2** (**Figure 10.3**). For all three marginal posteriors, I find $\hat{R} = 1.000$ for all parameters, indicating complete convergence across the chains, with bulk ESS > 5800 and tail ESS > 4200 for all parameters, ensuring high-precision inference.

Trace plots (see **Figures F.1 to F.3**) show ideal “hairy caterpillar” behaviour with rapid mixing and no visible trends. Rank plots (see **Figure F.4**) display uniform distributions across all chains, confirming good mixing.

These results provide confidence that the posterior constraints presented in **Table 10.2** are statistically robust and suitable for physics interpretation. The diagnostic validation ensures that reported credible intervals accurately reflect posterior uncertainty and that no additional sampling is required.

F.2 Signal injection study

To quantify the ability to correctly identify a specific parameter point in the data, I perform an injection study in which the model is fitted to toy data drawn from the **PDF** at a given point in m_X and μ_X . The model considered in this study corresponds to the case $\Gamma_X = 0.1$ GeV. In **Figure F.5**, I show the pulls

$$\frac{\hat{\eta} - \eta_0}{\hat{\sigma}_\eta}, \tag{F.1}$$

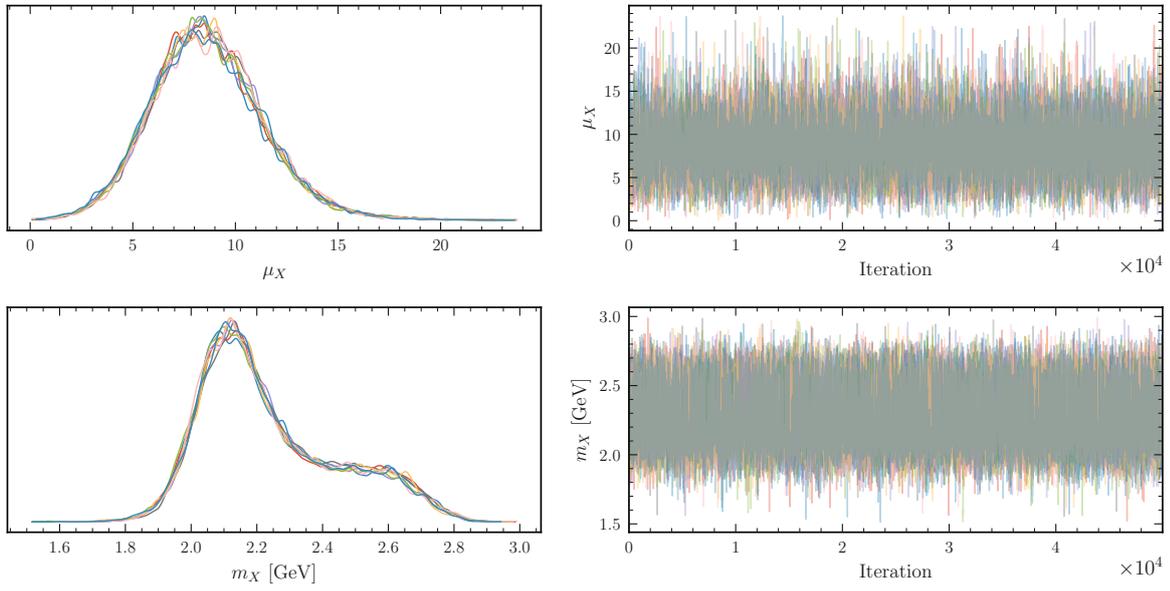


Figure F.1: Chain-wise marginal posterior distributions (*left*) and trace plots (*right*) for the $B^+ \rightarrow K^+ X$ signal strength μ_X and mass m_X in Equation (10.2), from the $\Gamma_X = 0.1$ GeV posterior discussed in Section 10.2 (Figure 10.3). Each colour represents a different MCMC chain. No trends or chain separation are visible.

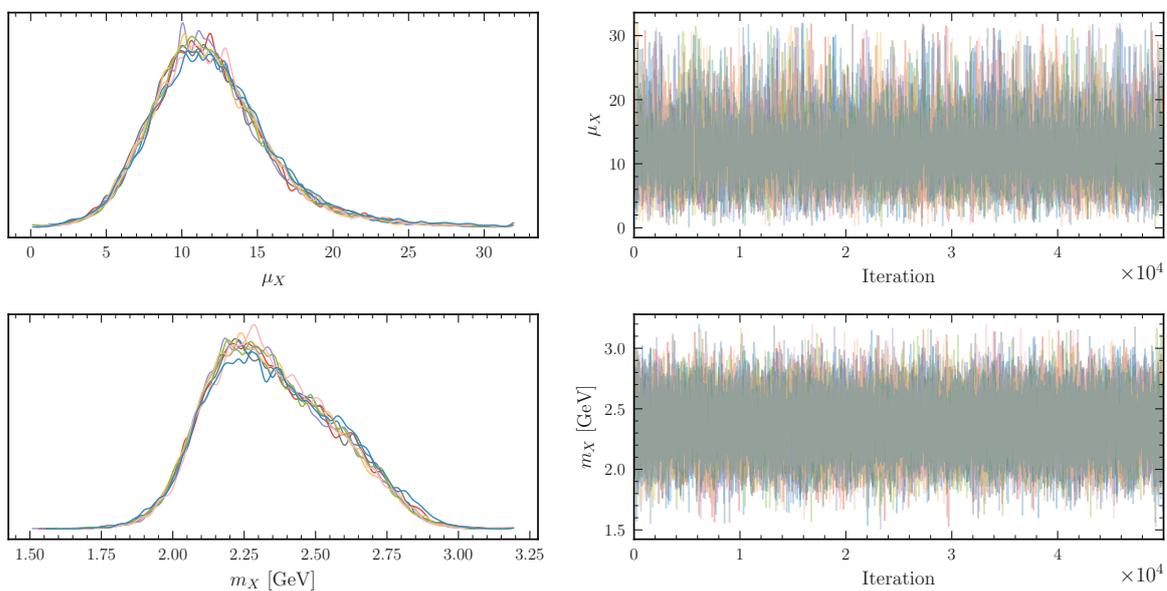


Figure F.2: Chain-wise marginal posterior distributions (*left*) and trace plots (*right*) for the $B^+ \rightarrow K^+ X$ signal strength μ_X and mass m_X in Equation (10.2), from the $\Gamma_X = 0.5$ GeV posterior discussed in Section 10.2 (Figure 10.3). Each colour represents a different MCMC chain. No trends or chain separation are visible.

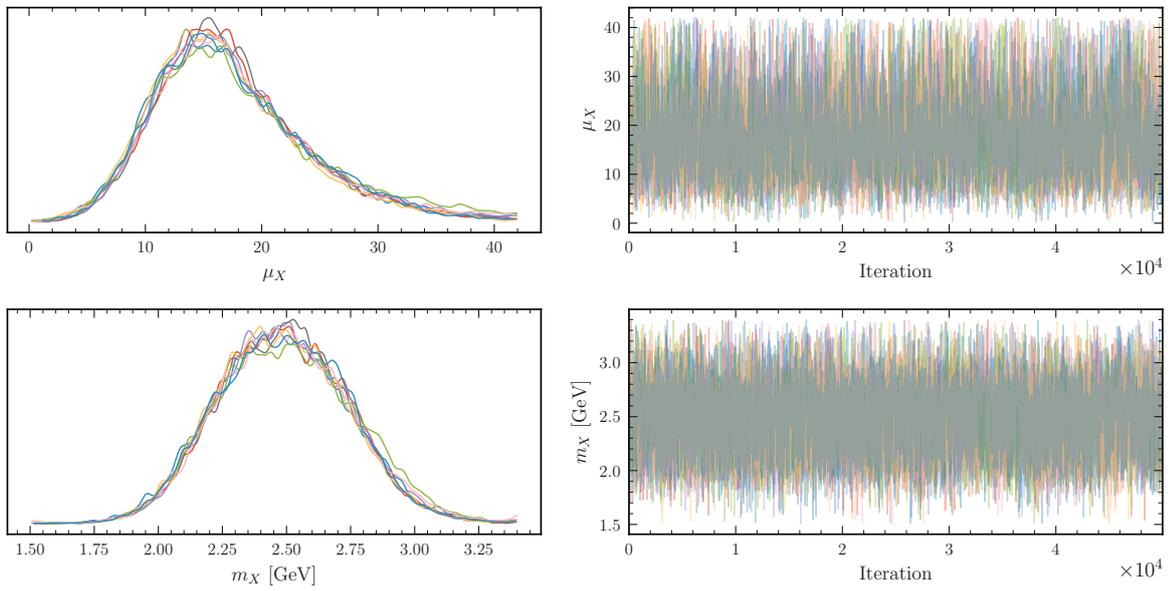


Figure F.3: Chain-wise marginal posterior distributions (*left*) and trace plots (*right*) for the $B^+ \rightarrow K^+ X$ signal strength μ_X and mass m_X in Equation (10.2), from the $\Gamma_X = 1.0$ GeV posterior discussed in Section 10.2 (Figure 10.3). Each colour represents a different MCMC chain. No trends or chain separation are visible.

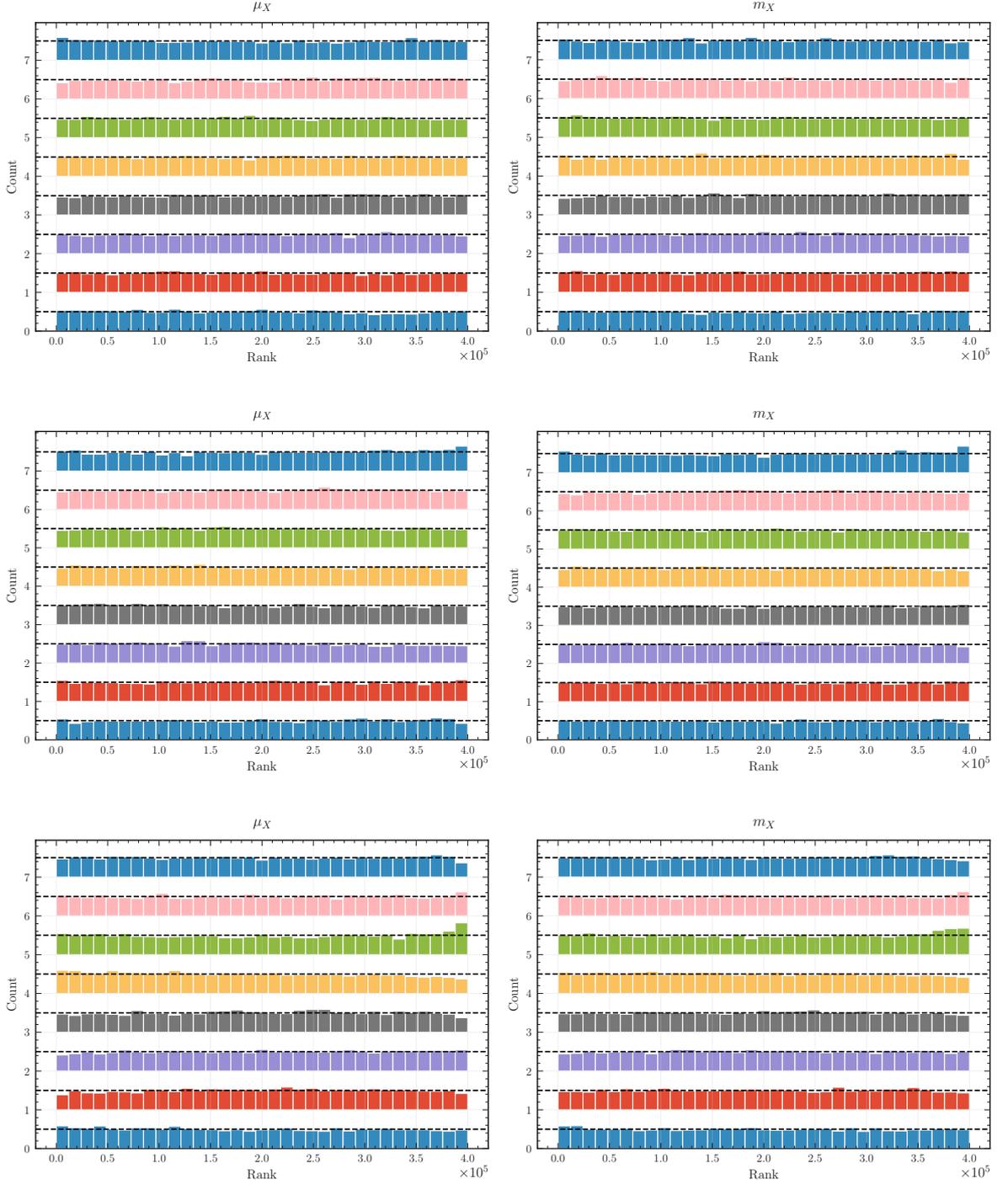


Figure F.4: Rank plots for the $B^+ \rightarrow K^+X$ signal strength μ_X and mass m_X in Equation (10.2), from the posteriors discussed in Section 10.2 (Figure 10.3). The *top/middle/bottom* two panels represent the $\Gamma_X = 0.1/0.5/1.0$ GeV cases, respectively. Each colour represents a different MCMC chain. The uniform distributions indicate good mixing.

in the 2-dimensional parameter space of m_X and μ_X .

Ideally, the pulls are expected to follow an uncorrelated multivariate normal distribution, as observed for the point $m_X = 2$ GeV and $\mu_X = 10$. However, due to the complex nature of the model and the low efficiency for $m_X > 3$ GeV, distinct features arise depending on the injected point. For instance, if $\mu_X = 0$ exactly, m_X becomes completely unconstrained, leading to the observed gap in the pulls for μ_X at the corresponding injected points. This strong parameter correlation also contributes to deviations from normality in the pull distributions. Additionally, the mass is bounded as $0 \leq m_X \leq 4.8$ GeV, resulting in a sharp cutoff of the pull distribution at the lower bound. Finally, it should be noted that q_{rec}^2 is binned into only three intervals, $[-1, 4, 8, 25]$ GeV², which further limits the precision with which the mass can be determined.

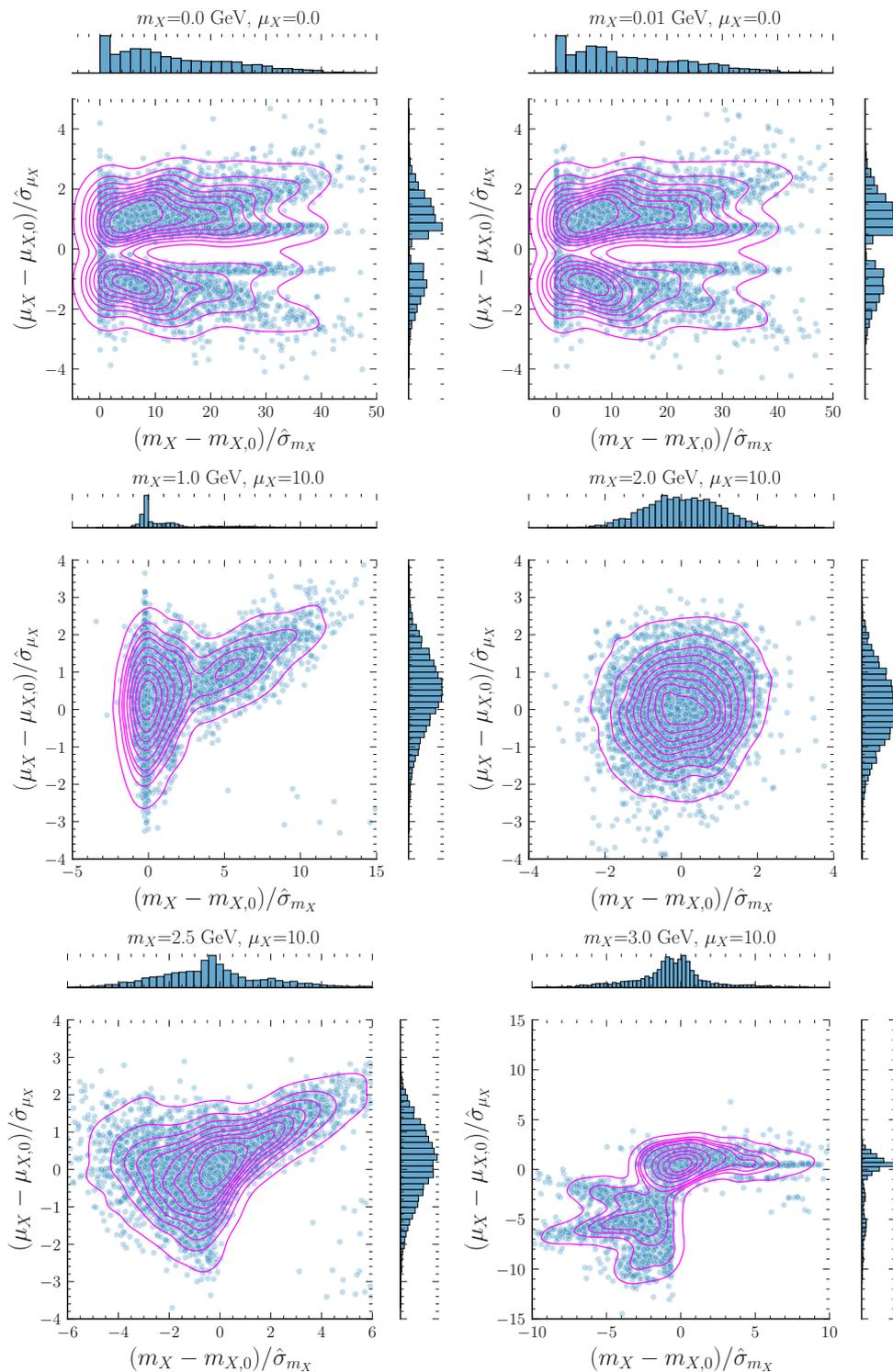


Figure F.5: The pulls in m_X and μ_X for each injection study, based on 5000 toy fits. The injected parameter point is displayed at the top of each plot.

Appendix G

A combined analysis of $B \rightarrow K^{(*)}\nu\bar{\nu}$ decays at Belle II

G.1 Individual contributions

The work presented in [Chapters 8](#) and [9](#) is a joint effort by the Belle II collaboration. The main proponents of this work are Valerio Bertacchi (V. B.), Yulan Fan, Lorenz Gärtner (L. G.), Eldar Ganiev, Alexander Glazov (A. G.), Yubo Han, Danylo Kulakov, Arsenii Kucher, Meihong Liu, Yuriy Onishchuk, Sebastiano Raiz, Niharika Rout, Caspar Schmitt and Slavomira Stefkova.

The theory predictions and decorrelated variations due to the form factor uncertainties presented in [Section 11.1](#) were computed by L. G.

The statistical inference presented in [Section 11.5](#) was performed by L. G. in collaboration with V. B. and with guidance from A. G. In detail, the construction of the individual `pyhf` models for all decay channels was refined by L. G. from the previous analysis [\[2\]](#). The combination of the `pyhf` models and the initial fitting strategy were designed by L. G. The profiled likelihood scans for the uncorrelated, isospin average and 4-channel correlated models were also performed by L. G. The post-fit correlations were computed by L. G. and plotted by V. B. The impact of the systematic uncertainties on the [POIs](#) in [Table 11.3](#) were computed by V. B.

List of acronyms

| | |
|---------------|---|
| ARICH | Aerogel Ring-Imaging Cherenkov detector |
| BCL | Bourelly-Caprini-Lellouch |
| BDT | Boosted Decision Tree |
| BSZ | Bharucha-Straub-Zwicky |
| BSM | Beyond the Standard Model |
| CDC | Central Drift Chamber |
| CDF | Cumulative Distribution Function |
| CKM | Cabibbo-Kobayashi-Maskawa |
| <i>CP</i> | Charge-Parity |
| d.o.f. | degrees of freedom |
| ECL | Electromagnetic Calorimeter |
| EFT | Effective Field Theory |
| FAIR | Findable, Accessible, Interoperable, Reusable |
| FCNC | Flavour Changing Neutral Current |
| GIM | Glashow-Iliopoulos-Maiani |
| HDI | Highest Density Interval |
| HEP | High-Energy Physics |
| HSR | Highest-Sensitivity Region |
| HS3 | HEP Statistics Serialization Standard |
| HTA | Hadronic Tagging Analysis |
| ITA | Inclusive Tagging Analysis |
| KLM | K_L^0 and muon detector |
| LHC | Large Hadron Collider |

| | |
|-------------|---------------------------------|
| MAP | Maximum A Posteriori estimator |
| MC | Monte Carlo |
| MCMC | Markov Chain Monte Carlo |
| mDST | mini Data Summary Table |
| MLE | Maximum Likelihood Estimator |
| OPE | Operator Product Expansion |
| PDF | Probability Density Function |
| PID | Particle Identification |
| PMNS | Pontecorvo-Maki-Nakagawa-Sakata |
| POI | Parameter Of Interest |
| PXD | Pixel Vertex Detector |
| QCD | Quantum Chromodynamics |
| QED | Quantum Electrodynamics |
| ROE | Rest-Of-Event |
| ROI | Region Of Interest |
| SM | Standard Model |
| SR | Signal Region |
| SVD | Silicon Vertex Detector |
| TOP | Time-Of-Propagation counter |
| UV | ultraviolet |
| vev | vacuum expectation value |
| VXD | Vertex Detector |
| WET | Weak Effective Theory |

Bibliography

- [1] Oxford English Dictionary. *scientific method*, *n.* Oxford English Dictionary. July 2023. URL: <https://doi.org/10.1093/OED/8843968527>.
- [2] I. Adachi et al. “Evidence for $B^+ \rightarrow K^+ \nu \bar{\nu}$ decays”. In: *Phys. Rev. D* 109.11 (2024), p. 112006. DOI: [10.1103/PhysRevD.109.112006](https://doi.org/10.1103/PhysRevD.109.112006). arXiv: [2311.14647](https://arxiv.org/abs/2311.14647) [[hep-ex](#)].
- [3] Jason Aebischer et al. “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](https://doi.org/10.1007/JHEP09(2017)158). arXiv: [1704.06639](https://arxiv.org/abs/1704.06639) [[hep-ph](#)].
- [4] Elizabeth E. Jenkins, Aneesh V. Manohar, and Peter Stoffer. “Low-Energy Effective Field Theory below the Electroweak Scale: Operators and Matching”. In: *JHEP* 03 (2018). [Erratum: *JHEP* 12, 043 (2023)], p. 016. DOI: [10.1007/JHEP03\(2018\)016](https://doi.org/10.1007/JHEP03(2018)016). arXiv: [1709.04486](https://arxiv.org/abs/1709.04486) [[hep-ph](#)].
- [5] Elizabeth E. Jenkins, Aneesh V. Manohar, and Peter Stoffer. “Low-Energy Effective Field Theory below the Electroweak Scale: Anomalous Dimensions”. In: *JHEP* 01 (2018). [Erratum: *JHEP* 12, 042 (2023)], p. 084. DOI: [10.1007/JHEP01\(2018\)084](https://doi.org/10.1007/JHEP01(2018)084). arXiv: [1711.05270](https://arxiv.org/abs/1711.05270) [[hep-ph](#)].
- [6] Damir Bečirević et al. “Scalar leptoquarks from grand unified theories to accommodate the B -physics anomalies”. In: *Phys. Rev. D* 98.5 (2018), p. 055003. DOI: [10.1103/PhysRevD.98.055003](https://doi.org/10.1103/PhysRevD.98.055003). arXiv: [1806.05689](https://arxiv.org/abs/1806.05689) [[hep-ph](#)].
- [7] I. Doršner et al. “Physics of leptoquarks in precision experiments and at particle colliders”. In: *Phys. Rept.* 641 (2016), pp. 1–68. DOI: [10.1016/j.physrep.2016.06.001](https://doi.org/10.1016/j.physrep.2016.06.001). arXiv: [1603.04993](https://arxiv.org/abs/1603.04993) [[hep-ph](#)].
- [8] Andrei Angelescu et al. “Single leptoquark solutions to the B -physics anomalies”. In: *Phys. Rev. D* 104.5 (2021), p. 055017. DOI: [10.1103/PhysRevD.104.055017](https://doi.org/10.1103/PhysRevD.104.055017). arXiv: [2103.12504](https://arxiv.org/abs/2103.12504) [[hep-ph](#)].
- [9] Thomas E. Browder et al. “Impact of $B \rightarrow K \nu \bar{\nu}$ measurements on beyond the Standard Model theories”. In: *Phys. Rev. D* 104.5 (2021), p. 053007. DOI: [10.1103/PhysRevD.104.053007](https://doi.org/10.1103/PhysRevD.104.053007). arXiv: [2107.01080](https://arxiv.org/abs/2107.01080) [[hep-ph](#)].
- [10] S. Navas et al. “Review of particle physics”. In: *Phys. Rev. D* 110.3 (2024), p. 030001. DOI: [10.1103/PhysRevD.110.030001](https://doi.org/10.1103/PhysRevD.110.030001).
- [11] A. Leike. “The Phenomenology of extra neutral gauge bosons”. In: *Phys. Rept.* 317 (1999), pp. 143–250. DOI: [10.1016/S0370-1573\(98\)00133-1](https://doi.org/10.1016/S0370-1573(98)00133-1). arXiv: [hep-ph/9805494](https://arxiv.org/abs/hep-ph/9805494).

- [12] Andrzej J. Buras, Fulvia De Fazio, and Jennifer Girrbach. “The Anatomy of Z' and Z with Flavour Changing Neutral Currents in the Flavour Precision Era”. In: *JHEP* 02 (2013), p. 116. DOI: [10.1007/JHEP02\(2013\)116](https://doi.org/10.1007/JHEP02(2013)116). arXiv: [1211.1896](https://arxiv.org/abs/1211.1896) [[hep-ph](#)].
- [13] Jorge Martin Camalich et al. “Quark Flavor Phenomenology of the QCD Axion”. In: *Phys. Rev. D* 102.1 (2020), p. 015023. DOI: [10.1103/PhysRevD.102.015023](https://doi.org/10.1103/PhysRevD.102.015023). arXiv: [2002.04623](https://arxiv.org/abs/2002.04623) [[hep-ph](#)].
- [14] R. D. Peccei and Helen R. Quinn. “CP Conservation in the Presence of Instantons”. In: *Phys. Rev. Lett.* 38 (1977), pp. 1440–1443. DOI: [10.1103/PhysRevLett.38.1440](https://doi.org/10.1103/PhysRevLett.38.1440).
- [15] R. D. Peccei and Helen R. Quinn. “Constraints Imposed by CP Conservation in the Presence of Instantons”. In: *Phys. Rev. D* 16 (1977), pp. 1791–1797. DOI: [10.1103/PhysRevD.16.1791](https://doi.org/10.1103/PhysRevD.16.1791).
- [16] Jorge Martin Camalich et al. “Quark Flavor Phenomenology of the QCD Axion”. In: *Phys. Rev. D* 102.1 (2020), p. 015023. DOI: [10.1103/PhysRevD.102.015023](https://doi.org/10.1103/PhysRevD.102.015023). arXiv: [2002.04623](https://arxiv.org/abs/2002.04623) [[hep-ph](#)].
- [17] Alfredo Walter Mario Guerrero and Stefano Rigolin. “ALP Production in Weak Mesonic Decays”. In: *Fortsch. Phys.* 71.2-3 (2023), p. 2200192. DOI: [10.1002/prop.202200192](https://doi.org/10.1002/prop.202200192). arXiv: [2211.08343](https://arxiv.org/abs/2211.08343) [[hep-ph](#)].
- [18] Sebastian Bruggisser, Lara Grabitz, and Susanne Westhoff. “Global analysis of the ALP effective theory”. In: *JHEP* 01 (2024), p. 092. DOI: [10.1007/JHEP01\(2024\)092](https://doi.org/10.1007/JHEP01(2024)092). arXiv: [2308.11703](https://arxiv.org/abs/2308.11703) [[hep-ph](#)].
- [19] Anastasiia Filimonova, Ruth Schäfer, and Susanne Westhoff. “Probing dark sectors with long-lived particles at BELLE II”. In: *Phys. Rev. D* 101.9 (2020), p. 095006. DOI: [10.1103/PhysRevD.101.095006](https://doi.org/10.1103/PhysRevD.101.095006). arXiv: [1911.03490](https://arxiv.org/abs/1911.03490) [[hep-ph](#)].
- [20] Alakabha Datta et al. “Dark photon and dark Z mediated B meson decays”. In: *JHEP* 03 (2023), p. 108. DOI: [10.1007/JHEP03\(2023\)108](https://doi.org/10.1007/JHEP03(2023)108). arXiv: [2210.15662](https://arxiv.org/abs/2210.15662) [[hep-ph](#)].
- [21] Murat Abdughani and Yakefu Reyimuaji. “Constraining light dark matter and mediator with $B^+ \rightarrow K^+ \nu \bar{\nu}$ data”. In: *Phys. Rev. D* 110.5 (2024), p. 055013. DOI: [10.1103/PhysRevD.110.055013](https://doi.org/10.1103/PhysRevD.110.055013). arXiv: [2309.03706](https://arxiv.org/abs/2309.03706) [[hep-ph](#)].
- [22] Alexander Berezhnoy and Dmitri Melikhov. “ $B \rightarrow K^* M_X$ vs $B \rightarrow K M_X$ as a probe of a scalar-mediator dark matter scenario”. In: *EPL* 145.1 (2024), p. 14001. DOI: [10.1209/0295-5075/ad1d03](https://doi.org/10.1209/0295-5075/ad1d03). arXiv: [2309.17191](https://arxiv.org/abs/2309.17191) [[hep-ph](#)].
- [23] Matthew D. Schwartz. *Quantum Field Theory and the Standard Model*. Cambridge University Press, Mar. 2014. ISBN: 978-1-107-03473-0, 978-1-107-03473-0.
- [24] Mark Thomson. *Modern particle physics*. New York: Cambridge University Press, Oct. 2013. ISBN: 978-1-107-03426-6, 978-1-139-52536-7. DOI: [10.1017/CB09781139525367](https://doi.org/10.1017/CB09781139525367).
- [25] Michael E. Peskin and Daniel V. Schroeder. *An Introduction to quantum field theory*. Reading, USA: Addison-Wesley, 1995. ISBN: 978-0-201-50397-5, 978-0-429-50355-9, 978-0-429-49417-8. DOI: [10.1201/9780429503559](https://doi.org/10.1201/9780429503559).

- [26] Steven Weinberg. *The quantum theory of fields. Vol. 3: Supersymmetry*. Cambridge University Press, June 2013. ISBN: 978-0-521-67055-5, 978-1-139-63263-8, 978-0-521-67055-5.
- [27] W. Buchmuller and C. Ludeling. “Field Theory and Standard Model”. In: *2005 European School of High-Energy Physics*. Sept. 2006. arXiv: [hep-ph/0609174](#).
- [28] Yuval Grossman and Philip Tanedo. “Just a taste: lectures on flavor physics.” In: *Theoretical Advanced Study Institute in Elementary Particle Physics: Anticipating the Next Discoveries in Particle Physics*. 2018, pp. 109–295. DOI: [10.1142/9789813233348_0004](#). arXiv: [1711.03624 \[hep-ph\]](#).
- [29] S. L. Glashow, J. Iliopoulos, and L. Maiani. “Weak Interactions with Lepton-Hadron Symmetry”. In: *Phys. Rev. D* 2 (1970), pp. 1285–1292. DOI: [10.1103/PhysRevD.2.1285](#).
- [30] Gerhard Buchalla, Andrzej J. Buras, and Markus E. Lautenbacher. “Weak Decays beyond Leading Logarithms”. In: *Rev. Mod. Phys.* 68 (1996), pp. 1125–1144. DOI: [10.1103/RevModPhys.68.1125](#). arXiv: [hep-ph/9512380](#).
- [31] Andrzej J. Buras. “Weak Hamiltonian, CP violation and rare decays”. In: *Les Houches Summer School in Theoretical Physics, Session 68: Probing the Standard Model of Particle Interactions*. June 1998, pp. 281–539. arXiv: [hep-ph/9806471](#).
- [32] Carlo Giunti and Alexander Studenikin. “Neutrino electromagnetic interactions: a window to new physics”. In: *Rev. Mod. Phys.* 87 (2015), p. 531. DOI: [10.1103/RevModPhys.87.531](#). arXiv: [1403.6344 \[hep-ph\]](#).
- [33] Tobias Felkl, Sze Lok Li, and Michael A. Schmidt. “A tale of invisibility: constraints on new physics in $b \rightarrow s\nu\nu$ ”. In: *JHEP* 12 (2021), p. 118. DOI: [10.1007/JHEP12\(2021\)118](#). arXiv: [2111.04327 \[hep-ph\]](#).
- [34] Lorenz Gärtner et al. “Constructing model-agnostic likelihoods, a method for the reinterpretation of particle physics results”. In: *Eur. Phys. J. C* 84.7 (2024), p. 693. DOI: [10.1140/epjc/s10052-024-13038-4](#). arXiv: [2402.08417 \[hep-ph\]](#).
- [35] Jason Aebischer et al. “WCxf: an exchange format for Wilson coefficients beyond the Standard Model”. In: *Comput. Phys. Commun.* 232 (2018), pp. 71–83. DOI: [10.1016/j.cpc.2018.05.022](#). arXiv: [1712.05298 \[hep-ph\]](#).
- [36] W. G. Parrott, C. Bouchard, and C. T. H. Davies. “Standard Model predictions for $B \rightarrow Kl^+l^-$, $B \rightarrow Kl_1^+l_2^-$, $B \rightarrow K\nu\bar{\nu}$ using form factors from $N_f = 2 + 1 + 1$ lattice QCD”. In: *Phys. Rev. D* 107.1 (2023). [Erratum: *Phys.Rev.D* 107, 119903 (2023)], p. 014511. DOI: [10.1103/PhysRevD.107.014511](#). arXiv: [2207.13371 \[hep-ph\]](#).
- [37] Danny van Dyk et al. *eos/eos: EOS Version 1.10.16*. Version v1.0.16. see also [the GitHub webpage](#). DOI: [10.5281/zenodo.15783309](#). URL: <https://doi.org/10.5281/zenodo.15783309>.
- [38] Laurent Lellouch. “Flavor physics and lattice quantum chromodynamics”. In: *Les Houches Summer School: Session 93: Modern perspectives in lattice QCD: Quantum field theory and high performance computing*. Apr. 2011, pp. 629–698. arXiv: [1104.5484 \[hep-lat\]](#).
- [39] Nico Gubernari, Ahmet Kokulu, and Danny van Dyk. “ $B \rightarrow P$ and $B \rightarrow V$ Form Factors from B -Meson Light-Cone Sum Rules beyond Leading Twist”. In: *JHEP* 01 (2019), p. 150. DOI: [10.1007/JHEP01\(2019\)150](#). arXiv: [1811.00983 \[hep-ph\]](#).

- [40] Claude Bourrely, Irinel Caprini, and Laurent Lellouch. “Model-independent description of $B \rightarrow \pi l \nu$ decays and a determination of $|V_{ub}|$ ”. In: *Phys. Rev. D* 79 (2009). [Erratum: *Phys.Rev.D* 82, 099902 (2010)], p. 013008. DOI: [10.1103/PhysRevD.82.099902](https://doi.org/10.1103/PhysRevD.82.099902). arXiv: [0807.2722](https://arxiv.org/abs/0807.2722) [hep-ph].
- [41] Aoife Bharucha, David M. Straub, and Roman Zwicky. “ $B \rightarrow V \ell^+ \ell^-$ in the Standard Model from light-cone sum rules”. In: *JHEP* 08 (2016), p. 098. DOI: [10.1007/JHEP08\(2016\)098](https://doi.org/10.1007/JHEP08(2016)098). arXiv: [1503.05534](https://arxiv.org/abs/1503.05534) [hep-ph].
- [42] Andrzej J. Buras et al. “ $B \rightarrow K^{(*)} \nu \bar{\nu}$ decays in the Standard Model and beyond”. In: *JHEP* 02 (2015), p. 184. DOI: [10.1007/JHEP02\(2015\)184](https://doi.org/10.1007/JHEP02(2015)184). arXiv: [1409.4557](https://arxiv.org/abs/1409.4557) [hep-ph].
- [43] Damir Bečirević, Gioacchino Piazza, and Olcyr Sumensari. “Revisiting $B \rightarrow K^{(*)} \nu \bar{\nu}$ decays in the Standard Model and beyond”. In: *Eur. Phys. J. C* 83.3 (2023), p. 252. DOI: [10.1140/epjc/s10052-023-11388-z](https://doi.org/10.1140/epjc/s10052-023-11388-z). arXiv: [2301.06990](https://arxiv.org/abs/2301.06990) [hep-ph].
- [44] James Gratx, Markus Hopfer, and Roman Zwicky. “Generalised helicity formalism, higher moments and the $B \rightarrow K_{JK}(\rightarrow K\pi)\bar{\ell}_1\ell_2$ angular distributions”. In: *Phys. Rev. D* 93.5 (2016), p. 054008. DOI: [10.1103/PhysRevD.93.054008](https://doi.org/10.1103/PhysRevD.93.054008). arXiv: [1506.03970](https://arxiv.org/abs/1506.03970) [hep-ph].
- [45] Y. Aoki et al. “FLAG Review 2021”. In: *Eur. Phys. J. C* 82.10 (2022), p. 869. DOI: [10.1140/epjc/s10052-022-10536-1](https://doi.org/10.1140/epjc/s10052-022-10536-1). arXiv: [2111.09849](https://arxiv.org/abs/2111.09849) [hep-lat].
- [46] W. G. Parrott, C. Bouchard, and C. T. H. Davies. “ $B \rightarrow K$ and $D \rightarrow K$ form factors from fully relativistic lattice QCD”. In: *Phys. Rev. D* 107.1 (2023), p. 014510. DOI: [10.1103/PhysRevD.107.014510](https://doi.org/10.1103/PhysRevD.107.014510). arXiv: [2207.12468](https://arxiv.org/abs/2207.12468) [hep-lat].
- [47] Danny van Dyk et al. “EOS: a software for flavor physics phenomenology”. In: *Eur. Phys. J. C* 82.6 (2022), p. 569. DOI: [10.1140/epjc/s10052-022-10177-4](https://doi.org/10.1140/epjc/s10052-022-10177-4). arXiv: [2111.15428](https://arxiv.org/abs/2111.15428) [hep-ph].
- [48] Aoife Bharucha, David M. Straub, and Roman Zwicky. “ $B \rightarrow V \ell^+ \ell^-$ in the Standard Model from light-cone sum rules”. In: *JHEP* 08 (2016), p. 098. DOI: [10.1007/JHEP08\(2016\)098](https://doi.org/10.1007/JHEP08(2016)098). arXiv: [1503.05534](https://arxiv.org/abs/1503.05534) [hep-ph].
- [49] Nico Gubernari et al. “Dispersive analysis of $B \rightarrow K^{(*)}$ and $B_s \rightarrow \phi$ form factors”. In: *JHEP* 12 (2023). [Erratum: *JHEP* 01, 125 (2025)], p. 153. DOI: [10.1007/JHEP12\(2023\)153](https://doi.org/10.1007/JHEP12(2023)153). arXiv: [2305.06301](https://arxiv.org/abs/2305.06301) [hep-ph].
- [50] Bernard Aubert et al. “Observation of CP violation in the B^0 meson system”. In: *Phys. Rev. Lett.* 87 (2001), p. 091801. DOI: [10.1103/PhysRevLett.87.091801](https://doi.org/10.1103/PhysRevLett.87.091801). arXiv: [hep-ex/0107013](https://arxiv.org/abs/hep-ex/0107013).
- [51] Kazuo Abe et al. “Observation of large CP violation in the neutral B meson system”. In: *Phys. Rev. Lett.* 87 (2001), p. 091802. DOI: [10.1103/PhysRevLett.87.091802](https://doi.org/10.1103/PhysRevLett.87.091802). arXiv: [hep-ex/0107061](https://arxiv.org/abs/hep-ex/0107061).
- [52] W. Altmannshofer et al. “The Belle II Physics Book”. In: *PTEP* 2019.12 (2019). Ed. by E. Kou and P. Urquijo. [Erratum: *PTEP* 2020, 029201 (2020)], p. 123C01. DOI: [10.1093/ptep/ptz106](https://doi.org/10.1093/ptep/ptz106). arXiv: [1808.10567](https://arxiv.org/abs/1808.10567) [hep-ex].
- [53] Kazunori Akai, Kazuro Furukawa, and Haruyo Koiso. “SuperKEKB Collider”. In: *Nucl. Instrum. Meth. A* 907 (2018), pp. 188–199. DOI: [10.1016/j.nima.2018.08.017](https://doi.org/10.1016/j.nima.2018.08.017). arXiv: [1809.01958](https://arxiv.org/abs/1809.01958) [physics.acc-ph].

- [54] Yukiyoishi Ohnishi et al. “Accelerator design at SuperKEKB”. In: *PTEP* 2013 (2013), 03A011. DOI: [10.1093/ptep/pts083](https://doi.org/10.1093/ptep/pts083).
- [55] I. Adachi et al. “Measurement of the integrated luminosity of data samples collected during 2019-2022 by the Belle II experiment*”. In: *Chin. Phys. C* 49.1 (2025), p. 013001. DOI: [10.1088/1674-1137/ad806c](https://doi.org/10.1088/1674-1137/ad806c). arXiv: [2407.00965](https://arxiv.org/abs/2407.00965) [[hep-ex](#)].
- [56] I. Adachi et al. “Detectors for extreme luminosity: Belle II”. In: *Nucl. Instrum. Meth. A* 907 (2018), pp. 46–59. DOI: [10.1016/j.nima.2018.03.068](https://doi.org/10.1016/j.nima.2018.03.068).
- [57] T. Abe et al. “Belle II Technical Design Report”. In: (Nov. 2010). arXiv: [1011.0352](https://arxiv.org/abs/1011.0352) [[physics.ins-det](#)].
- [58] B. Shwartz and B. Shwartz. “Electromagnetic calorimeter of the Belle II detector”. In: *J. Phys. Conf. Ser.* 928.1 (2017), p. 012021. DOI: [10.1088/1742-6596/928/1/012021](https://doi.org/10.1088/1742-6596/928/1/012021).
- [59] S. Bähr et al. “The neural network first-level hardware track trigger of the Belle II experiment”. In: *Nucl. Instrum. Meth. A* 1073 (2025), p. 170279. DOI: [10.1016/j.nima.2025.170279](https://doi.org/10.1016/j.nima.2025.170279). arXiv: [2402.14962](https://arxiv.org/abs/2402.14962) [[hep-ex](#)].
- [60] T. Kuhr et al. “The Belle II Core Software”. In: *Comput. Softw. Big Sci.* 3.1 (2019), p. 1. DOI: [10.1007/s41781-018-0017-9](https://doi.org/10.1007/s41781-018-0017-9). arXiv: [1809.04299](https://arxiv.org/abs/1809.04299) [[physics.comp-ph](#)].
- [61] Valerio Bertacchi et al. “Track finding at Belle II”. In: *Comput. Phys. Commun.* 259 (2021), p. 107610. DOI: [10.1016/j.cpc.2020.107610](https://doi.org/10.1016/j.cpc.2020.107610). arXiv: [2003.12466](https://arxiv.org/abs/2003.12466) [[physics.ins-det](#)].
- [62] D. J. Lange. “The EvtGen particle decay simulation package”. In: *Nucl. Instrum. Meth. A* 462 (2001). Ed. by S. Erhan, P. Schlein, and Y. Rozen, pp. 152–155. DOI: [10.1016/S0168-9002\(01\)00089-4](https://doi.org/10.1016/S0168-9002(01)00089-4).
- [63] Torbjörn Sjöstrand et al. “An introduction to PYTHIA 8.2”. In: *Comput. Phys. Commun.* 191 (2015), pp. 159–177. DOI: [10.1016/j.cpc.2015.01.024](https://doi.org/10.1016/j.cpc.2015.01.024). arXiv: [1410.3012](https://arxiv.org/abs/1410.3012) [[hep-ph](#)].
- [64] S. Agostinelli et al. “GEANT4 - A Simulation Toolkit”. In: *Nucl. Instrum. Meth. A* 506 (2003), pp. 250–303. DOI: [10.1016/S0168-9002\(03\)01368-8](https://doi.org/10.1016/S0168-9002(03)01368-8).
- [65] Glen Cowan. *Statistical Data Analysis*. Clarendon Press, 1998.
- [66] G. A. Young and R. L. Smith. *Essentials of Statistical Inference*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2005.
- [67] Frederick James. *Statistical methods in experimental physics*. 2006. ISBN: 978-981-256-795-6, 978-981-270-527-3.
- [68] Andrew Gelman et al. *Bayesian Data Analysis*. Third. Chapman & Hall/CRC Texts in Statistical Science Series. CRC, 2013. ISBN: 9781439840955 1439840954. URL: <https://stat.columbia.edu/~gelman/book/>.
- [69] Kyle Cranmer et al. “Publishing statistical models: Getting the most out of particle physics experiments”. In: *SciPost Phys.* 12.1 (2022), p. 037. DOI: [10.21468/SciPostPhys.12.1.037](https://doi.org/10.21468/SciPostPhys.12.1.037). arXiv: [2109.04981](https://arxiv.org/abs/2109.04981) [[hep-ph](#)].
- [70] Kyle Cranmer et al. “HistFactory: A tool for creating statistical models for use with RooFit and RooStats”. In: (June 2012).
- [71] Lukas Heinrich et al. “pyhf: pure-Python implementation of HistFactory statistical models”. In: *J. Open Source Softw.* 6.58 (2021), p. 2823. DOI: [10.21105/joss.02823](https://doi.org/10.21105/joss.02823).

- [72] Lukas Heinrich, Matthew Feickert, and Giordon Stark. *pyhf: v0.7.6*. Version 0.7.6. DOI: [10.5281/zenodo.1169739](https://doi.org/10.5281/zenodo.1169739). URL: <https://doi.org/10.5281/zenodo.1169739>.
- [73] Roger J. Barlow and Christine Beeston. “Fitting using finite Monte Carlo samples”. In: *Comput. Phys. Commun.* 77 (1993), pp. 219–228. DOI: [10.1016/0010-4655\(93\)90005-W](https://doi.org/10.1016/0010-4655(93)90005-W).
- [74] Matthew Feickert, Lukas Heinrich, and Malin Horstmann. “Bayesian Methodologies with pyhf”. In: *EPJ Web Conf.* 295 (2024), p. 06004. DOI: [10.1051/epjconf/202429506004](https://doi.org/10.1051/epjconf/202429506004). arXiv: [2309.17005](https://arxiv.org/abs/2309.17005) [stat.CO].
- [75] Rene Brun et al. *root-project/root: v6.18/02*. Version v6-18-02. DOI: [10.5281/zenodo.3895860](https://doi.org/10.5281/zenodo.3895860). URL: <https://doi.org/10.5281/zenodo.3895860>.
- [76] Pauli Virtanen et al. “SciPy 1.0—Fundamental Algorithms for Scientific Computing in Python”. In: *Nature Meth.* 17 (2020), p. 261. DOI: [10.1038/s41592-019-0686-2](https://doi.org/10.1038/s41592-019-0686-2). arXiv: [1907.10121](https://arxiv.org/abs/1907.10121) [cs.MS].
- [77] F. James. “MINUIT Function Minimization and Error Analysis: Reference Manual Version 94.1”. In: (1994).
- [78] James Bradbury et al. *JAX: composable transformations of Python+NumPy programs*. Version 0.3.13. 2018. URL: <http://github.com/google/jax>.
- [79] Glen Cowan et al. “Asymptotic formulae for likelihood-based tests of new physics”. In: *Eur. Phys. J. C* 71 (2011). [Erratum: *Eur.Phys.J.C* 73, 2501 (2013)], p. 1554. DOI: [10.1140/epjc/s10052-011-1554-0](https://doi.org/10.1140/epjc/s10052-011-1554-0). arXiv: [1007.1727](https://arxiv.org/abs/1007.1727) [physics.data-an].
- [80] Oriol Abril-Pla et al. “PyMC: a modern, and comprehensive probabilistic programming framework in Python”. In: *PeerJ Computer Science* 9 (2023). URL: <https://api.semanticscholar.org/CorpusID:261469458>.
- [81] Lorenz Gärtner. *lorenzennio/pyhfcorr: v1.0.0*. Version v1.0.0. DOI: [10.5281/zenodo.16794057](https://doi.org/10.5281/zenodo.16794057). URL: <https://doi.org/10.5281/zenodo.16794057>.
- [82] Robert McLeod et al. *pydata/numexpr: NumExpr v2.6.9*. Version v2.6.9. DOI: [10.5281/zenodo.2483274](https://doi.org/10.5281/zenodo.2483274). URL: <https://doi.org/10.5281/zenodo.2483274>.
- [83] Lorenz Gärtner. *lorenzennio/redis: redis version 1.0.4*. Version v1.0.4. DOI: [10.5281/zenodo.15968344](https://doi.org/10.5281/zenodo.15968344). URL: <https://doi.org/10.5281/zenodo.15968344>.
- [84] Alexander L. Read. “Presentation of search results: The CL_s technique”. In: *J. Phys. G* 28 (2002). Ed. by M. R. Whalley and L. Lyons, pp. 2693–2704. DOI: [10.1088/0954-3899/28/10/313](https://doi.org/10.1088/0954-3899/28/10/313).
- [85] Abraham Wald. “Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations is Large”. In: *Transactions of the American Mathematical Society* 54.3 (1943), pp. 426–482. ISSN: 00029947, 10886850. URL: <http://www.jstor.org/stable/1990256> (visited on 08/12/2025).
- [86] S. S. Wilks. “The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses”. In: *Annals Math. Statist.* 9.1 (1938), pp. 60–62. DOI: [10.1214/aoms/1177732360](https://doi.org/10.1214/aoms/1177732360).
- [87] Harold Jeffreys. *The Theory of Probability*. Oxford Classic Texts in the Physical Sciences. 1939. ISBN: 978-0-19-850368-2, 978-0-19-853193-7.
- [88] Robert E. Kass and Adrian E. Raftery. “Bayes Factors”. In: *J. Am. Statist. Assoc.* 90.430 (1995), pp. 773–795. DOI: [10.1080/01621459.1995.10476572](https://doi.org/10.1080/01621459.1995.10476572).

- [89] Minas Karamanis. “Bayesian Computation in Astronomy: Novel methods for parallel and gradient-free inference”. Other thesis. Mar. 2023. DOI: [10.7488/era/2932](https://doi.org/10.7488/era/2932). arXiv: [2303.16134](https://arxiv.org/abs/2303.16134) [[astro-ph.IM](#)].
- [90] Robert D. Cousins. “Generalization of Chisquare Goodness-of-Fit Test for Binned Data Using Saturated Models, with Application to Histograms”. In: 2013. URL: <https://api.semanticscholar.org/CorpusID:5936965>.
- [91] John Kruschke. *Doing Bayesian Data Analysis*. 2nd ed. Boston: Academic Press.
- [92] Jonathan Butterworth et al. “Reinterpretation and preservation of data and analyses in HEP”. In: (Mar. 2025). arXiv: [2504.00256](https://arxiv.org/abs/2504.00256) [[hep-ph](#)].
- [93] Stephen Bailey et al. “Data and Analysis Preservation, Recasting, and Reinterpretation”. In: (Mar. 2022). arXiv: [2203.10057](https://arxiv.org/abs/2203.10057) [[hep-ph](#)].
- [94] Waleed Abdallah et al. “Reinterpretation of LHC Results for New Physics: Status and Recommendations after Run 2”. In: *SciPost Phys.* 9.2 (2020), p. 022. DOI: [10.21468/SciPostPhys.9.2.022](https://doi.org/10.21468/SciPostPhys.9.2.022). arXiv: [2003.07868](https://arxiv.org/abs/2003.07868) [[hep-ph](#)].
- [95] CERN. *CERN Open Data Portal*. <https://opendata.cern.ch/>. Accessed: 2025-07-11. 2024.
- [96] Christian Bierlich et al. “Robust Independent Validation of Experiment and Theory: Rivet version 3”. In: *SciPost Phys.* 8 (2020), p. 026. DOI: [10.21468/SciPostPhys.8.2.026](https://doi.org/10.21468/SciPostPhys.8.2.026). arXiv: [1912.05451](https://arxiv.org/abs/1912.05451) [[hep-ph](#)].
- [97] HS3 Collaboration. *Hep Statistics Serialization Standard*. <https://github.com/hep-statistics-serialization-standard/hep-statistics-serialization-standard>. Accessed: 2025-07-07.
- [98] Jonas Eschle et al. “zfit: scalable pythonic fitting”. In: (Oct. 2019). DOI: [10.1016/j.softx.2020.100508](https://doi.org/10.1016/j.softx.2020.100508). arXiv: [1910.13429](https://arxiv.org/abs/1910.13429) [[physics.data-an](#)].
- [99] Kyle Cranmer and Itay Yavin. “RECAST: Extending the Impact of Existing Analyses”. In: *JHEP* 04 (2011), p. 038. DOI: [10.1007/JHEP04\(2011\)038](https://doi.org/10.1007/JHEP04(2011)038). arXiv: [1010.2506](https://arxiv.org/abs/1010.2506) [[hep-ex](#)].
- [100] Manuel Drees et al. “CheckMATE: Confronting your Favourite New Physics Model with LHC Data”. In: *Comput. Phys. Commun.* 187 (2015), pp. 227–265. DOI: [10.1016/j.cpc.2014.10.018](https://doi.org/10.1016/j.cpc.2014.10.018). arXiv: [1312.2591](https://arxiv.org/abs/1312.2591) [[hep-ph](#)].
- [101] Daniel Dercks et al. “CheckMATE 2: From the model to the limit”. In: *Comput. Phys. Commun.* 221 (2017), pp. 383–418. DOI: [10.1016/j.cpc.2017.08.021](https://doi.org/10.1016/j.cpc.2017.08.021). arXiv: [1611.09856](https://arxiv.org/abs/1611.09856) [[hep-ph](#)].
- [102] Eric Conte, Benjamin Fuks, and Guillaume Serret. “MadAnalysis 5, A User-Friendly Framework for Collider Phenomenology”. In: *Comput. Phys. Commun.* 184 (2013), pp. 222–256. DOI: [10.1016/j.cpc.2012.09.009](https://doi.org/10.1016/j.cpc.2012.09.009). arXiv: [1206.1599](https://arxiv.org/abs/1206.1599) [[hep-ph](#)].
- [103] Eric Conte and Benjamin Fuks. “Confronting new physics theories to LHC data with MADANALYSIS 5”. In: *Int. J. Mod. Phys. A* 33.28 (2018), p. 1830027. DOI: [10.1142/S0217751X18300272](https://doi.org/10.1142/S0217751X18300272). arXiv: [1808.00480](https://arxiv.org/abs/1808.00480) [[hep-ph](#)].
- [104] Anders Kvellestad, Pat Scott, and Martin White. “GAMBIT and its application in the search for physics Beyond the Standard Model”. In: *Prog. Part. Nucl. Phys.* 113 (2020), p. 103769. DOI: [10.1016/j.pnpnp.2020.103769](https://doi.org/10.1016/j.pnpnp.2020.103769). arXiv: [1912.04079](https://arxiv.org/abs/1912.04079) [[hep-ph](#)].

- [105] Jonathan M. Butterworth et al. “Constraining new physics with collider measurements of Standard Model signatures”. In: *JHEP* 03 (2017), p. 078. DOI: [10.1007/JHEP03\(2017\)078](https://doi.org/10.1007/JHEP03(2017)078). arXiv: [1606.05296](https://arxiv.org/abs/1606.05296) [hep-ph].
- [106] Giordon Stark, Camila Aristimuno Ots, and Mike Hance. “Reduce, Reuse, Reinterpret: an end-to-end pipeline for recycling particle physics results”. In: (June 2023). DOI: [10.21468/SciPostPhysCodeb.27](https://doi.org/10.21468/SciPostPhysCodeb.27). arXiv: [2306.11055](https://arxiv.org/abs/2306.11055) [hep-ex].
- [107] Giordon Stark, Camila Aristimuno Ots, and Mike Hance. “Codebase release 0.5 for mapyde”. In: *SciPost Phys. Codebases* (2024), 27–r0.5. DOI: [10.21468/SciPostPhysCodeb.27-r0.5](https://doi.org/10.21468/SciPostPhysCodeb.27-r0.5). URL: <https://scipost.org/10.21468/SciPostPhysCodeb.27-r0.5>.
- [108] Serguei Chatrchyan et al. “Interpretation of Searches for Supersymmetry with Simplified Models”. In: *Phys. Rev. D* 88.5 (2013), p. 052017. DOI: [10.1103/PhysRevD.88.052017](https://doi.org/10.1103/PhysRevD.88.052017). arXiv: [1301.2175](https://arxiv.org/abs/1301.2175) [hep-ex].
- [109] Sabine Kraml et al. “SModelS: a tool for interpreting simplified-model results from the LHC and its application to supersymmetry”. In: *Eur. Phys. J. C* 74 (2014), p. 2868. DOI: [10.1140/epjc/s10052-014-2868-5](https://doi.org/10.1140/epjc/s10052-014-2868-5). arXiv: [1312.4175](https://arxiv.org/abs/1312.4175) [hep-ph].
- [110] Federico Ambrogio et al. “SModelS v1.2: long-lived particles, combination of signal regions, and other novelties”. In: *Comput. Phys. Commun.* 251 (2020), p. 106848. DOI: [10.1016/j.cpc.2019.07.013](https://doi.org/10.1016/j.cpc.2019.07.013). arXiv: [1811.10624](https://arxiv.org/abs/1811.10624) [hep-ph].
- [111] F. Abudinén et al. “Search for $B^+ \rightarrow K^+ \nu \bar{\nu}$ Decays Using an Inclusive Tagging Method at Belle II”. In: *Phys. Rev. Lett.* 127.18 (2021), p. 181802. DOI: [10.1103/PhysRevLett.127.181802](https://doi.org/10.1103/PhysRevLett.127.181802). arXiv: [2104.12624](https://arxiv.org/abs/2104.12624) [hep-ex].
- [112] Kåre Fridell et al. “Decoding the $B \rightarrow K \nu \bar{\nu}$ excess at Belle II: Kinematics, operators, and masses”. In: *Phys. Rev. D* 109.11 (2024), p. 115006. DOI: [10.1103/PhysRevD.109.115006](https://doi.org/10.1103/PhysRevD.109.115006). arXiv: [2312.12507](https://arxiv.org/abs/2312.12507) [hep-ph].
- [113] Wolfgang Altmannshofer et al. “Light new physics in $B \rightarrow K^{(*)} \nu \bar{\nu}$?” In: *Phys. Rev. D* 109.7 (2024), p. 075008. DOI: [10.1103/PhysRevD.109.075008](https://doi.org/10.1103/PhysRevD.109.075008). arXiv: [2311.14629](https://arxiv.org/abs/2311.14629) [hep-ph].
- [114] Emidio Gabrielli et al. “Explaining the $B^+ \rightarrow K^+ \nu \bar{\nu}$ excess via a massless dark photon”. In: *Eur. Phys. J. C* 84.5 (2024), p. 460. DOI: [10.1140/epjc/s10052-024-12818-2](https://doi.org/10.1140/epjc/s10052-024-12818-2). arXiv: [2402.05901](https://arxiv.org/abs/2402.05901) [hep-ph].
- [115] Danny van Dyk et al. *EOS Version 1.0.11*. Version v1.0.11. DOI: [10.5281/zenodo.10600399](https://doi.org/10.5281/zenodo.10600399). URL: <https://doi.org/10.5281/zenodo.10600399>.
- [116] Merna Abumusabh et al. “A model-agnostic likelihood for the reinterpretation of the $B^+ \rightarrow K^+ \nu \bar{\nu}$ measurement at Belle II”. In: *Accepted by Phys. Rev. D* (July 2025). arXiv: [2507.12393](https://arxiv.org/abs/2507.12393) [hep-ex].
- [117] T. E. Browder et al. “A Search for $B \rightarrow \tau \nu$ ”. In: *Phys. Rev. Lett.* 86 (2001). Ed. by K. K. Gan and R. Kass, pp. 2950–2954. DOI: [10.1103/PhysRevLett.86.2950](https://doi.org/10.1103/PhysRevLett.86.2950). arXiv: [hep-ex/0007057](https://arxiv.org/abs/hep-ex/0007057).
- [118] O. Lutz et al. “Search for $B \rightarrow h^{(*)} \nu \bar{\nu}$ with the full Belle $\Upsilon(4S)$ data sample”. In: *Phys. Rev. D* 87.11 (2013), p. 111103. DOI: [10.1103/PhysRevD.87.111103](https://doi.org/10.1103/PhysRevD.87.111103). arXiv: [1303.3719](https://arxiv.org/abs/1303.3719) [hep-ex].

- [119] J. P. Lees et al. “Search for $B \rightarrow K^{(*)}\nu\bar{\nu}$ and invisible quarkonium decays”. In: *Phys. Rev. D* 87.11 (2013), p. 112005. DOI: [10.1103/PhysRevD.87.112005](https://doi.org/10.1103/PhysRevD.87.112005). arXiv: [1303.7465](https://arxiv.org/abs/1303.7465) [hep-ex].
- [120] P. del Amo Sanchez et al. “Search for the Rare Decay $B \rightarrow K\nu\bar{\nu}$ ”. In: *Phys. Rev. D* 82 (2010), p. 112002. DOI: [10.1103/PhysRevD.82.112002](https://doi.org/10.1103/PhysRevD.82.112002). arXiv: [1009.1529](https://arxiv.org/abs/1009.1529) [hep-ex].
- [121] J. Grygier et al. “Search for $B \rightarrow h\nu\bar{\nu}$ decays with semileptonic tagging at Belle”. In: *Phys. Rev. D* 96.9 (2017). [Addendum: *Phys. Rev. D* 97, 099902 (2018)], p. 091101. DOI: [10.1103/PhysRevD.96.091101](https://doi.org/10.1103/PhysRevD.96.091101). arXiv: [1702.03224](https://arxiv.org/abs/1702.03224) [hep-ex].
- [122] Belle-II Collaboration. *A model-agnostic likelihood for the reinterpretation of the $B^+ \rightarrow K^+\nu\bar{\nu}$ measurement at Belle II*. HEPData (collection). <https://doi.org/10.17182/hepdata.166082>. 2025.
- [123] Belle-II Collaboration. *Evidence for $B^+ \rightarrow K^+\nu\bar{\nu}$ Decays*. HEPData (collection). <https://doi.org/10.17182/hepdata.146803>. 2024.
- [124] F. Abudinén et al. “Measurement of the integrated luminosity of the Phase 2 data of the Belle II experiment”. In: *Chin. Phys. C* 44.2 (2020), p. 021001. DOI: [10.1088/1674-1137/44/2/021001](https://doi.org/10.1088/1674-1137/44/2/021001). arXiv: [1910.05365](https://arxiv.org/abs/1910.05365) [hep-ex].
- [125] Bernard Aubert et al. “A measurement of the total width, the electronic width, and the mass of the $\Upsilon(10580)$ resonance”. In: *Phys. Rev. D* 72 (2005), p. 032005. DOI: [10.1103/PhysRevD.72.032005](https://doi.org/10.1103/PhysRevD.72.032005). arXiv: [hep-ex/0405025](https://arxiv.org/abs/hep-ex/0405025).
- [126] Kyle Cranmer. “Practical Statistics for the LHC”. In: *2011 European School of High-Energy Physics*. 2014, pp. 267–308. DOI: [10.5170/CERN-2014-003.267](https://doi.org/10.5170/CERN-2014-003.267). arXiv: [1503.07622](https://arxiv.org/abs/1503.07622) [physics.data-an].
- [127] Florian Urs Bernlochner et al. “Cover your bases: asymptotic distributions of the profile likelihood ratio when constraining effective field theories in high-energy physics”. In: *SciPost Phys. Core* 6 (2023), p. 013. DOI: [10.21468/SciPostPhysCore.6.1.013](https://doi.org/10.21468/SciPostPhysCore.6.1.013). arXiv: [2207.01350](https://arxiv.org/abs/2207.01350) [physics.data-an].
- [128] S. E. Minson, M. Simons, and J. L. Beck. “Bayesian inversion for finite fault earthquake source models I—theory and algorithm”. In: *Geophysical Journal International* 194.3 (June 2013), pp. 1701–1726. ISSN: 0956-540X. DOI: [10.1093/gji/ggt180](https://doi.org/10.1093/gji/ggt180). eprint: https://academic.oup.com/gji/article-pdf/194/3/1701/39595731/gji_194_3_1701.pdf. URL: <https://doi.org/10.1093/gji/ggt180>.
- [129] Eilam Gross and Ofer Vitells. “Trial factors for the look elsewhere effect in high energy physics”. In: *Eur. Phys. J. C* 70 (2010), pp. 525–530. DOI: [10.1140/epjc/s10052-010-1470-8](https://doi.org/10.1140/epjc/s10052-010-1470-8). arXiv: [1005.1891](https://arxiv.org/abs/1005.1891) [physics.data-an].
- [130] Nico Gubernari, Danny van Dyk, and Javier Virto. “Non-local matrix elements in $B_{(s)} \rightarrow \{K^{(*)}, \phi\}\ell^+\ell^-$ ”. In: *JHEP* 02 (2021), p. 088. DOI: [10.1007/JHEP02\(2021\)088](https://doi.org/10.1007/JHEP02(2021)088). arXiv: [2011.09813](https://arxiv.org/abs/2011.09813) [hep-ph].
- [131] R. R. Horgan et al. “Rare B decays using lattice QCD form factors”. In: *PoS LATTICE2014* (2015), p. 372. DOI: [10.22323/1.214.0372](https://doi.org/10.22323/1.214.0372). arXiv: [1501.00367](https://arxiv.org/abs/1501.00367) [hep-lat].

BIBLIOGRAPHY

- [132] S. H. Lee et al. “Evidence for $B^0 \rightarrow \pi^0\pi^0$ ”. In: *Phys. Rev. Lett.* 91 (2003). Ed. by H. W. K. Cheung and T. S. Pratt, p. 261801. DOI: [10.1103/PhysRevLett.91.261801](https://doi.org/10.1103/PhysRevLett.91.261801). arXiv: [hep-ex/0308040](https://arxiv.org/abs/hep-ex/0308040).
- [133] J. Grygier et al. “Search for $B \rightarrow h\nu\bar{\nu}$ decays with semileptonic tagging at Belle”. In: *Phys. Rev. D* 96.9 (2017). [Addendum: *Phys.Rev.D* 97, 099902 (2018)], p. 091101. DOI: [10.1103/PhysRevD.96.091101](https://doi.org/10.1103/PhysRevD.96.091101). arXiv: [1702.03224](https://arxiv.org/abs/1702.03224) [[hep-ex](#)].
- [134] Johannes Grygier et al. *Analysis of the Rare Decay $B \rightarrow h\nu\bar{\nu}$ with Semileptonic Tag*. Tech. rep. Belle Note 1413. Belle Note #1413, Version 4b. Belle Collaboration, Jan. 2017. URL: https://belle.kek.jp/secured/belle_note/gn1413/belle_note_hnuu_v4b.pdf.
- [135] Marco Colonna et al. “Redist, a Python Tool for Model-Agnostic Binned-Likelihood Fits in High Energy Physics”. In: *Proceedings of the 24th Python in Science Conference (SciPy 2025)*. SciPy Conference, July 2025. DOI: [10.25080/ppgk2467](https://doi.org/10.25080/ppgk2467). URL: <https://doi.org/10.25080/ppgk2467>.
- [136] Mark D. Wilkinson et al. “The FAIR Guiding Principles for scientific data management and stewardship”. In: *Sci. Data* 3.1 (2016), p. 160018. DOI: [10.1038/sdata.2016.18](https://doi.org/10.1038/sdata.2016.18).
- [137] Aki Vehtari et al. “Rank-Normalization, Folding, and Localization: An Improved \hat{R} for Assessing Convergence of MCMC (with Discussion)”. In: *Bayesian Analysis* 16.2 (2021). ISSN: 1936-0975. DOI: [10.1214/20-ba1221](https://doi.org/10.1214/20-ba1221). URL: <http://dx.doi.org/10.1214/20-BA1221>.
- [138] Vivekananda Roy. *Convergence diagnostics for Markov chain Monte Carlo*. 2019. arXiv: [1909.11827](https://arxiv.org/abs/1909.11827) [[stat.CO](#)].
- [139] Kyle Cranmer and Lukas Heinrich. *Recasting through reweighting*. DOI: [10.5281/zenodo.1013926](https://doi.org/10.5281/zenodo.1013926). URL: <https://doi.org/10.5281/zenodo.1013926>.
- [140] Elisabetta Barberio and Zbigniew Was. “PHOTOS: A Universal Monte Carlo for QED radiative corrections. Version 2.0”. In: *Comput. Phys. Commun.* 79 (1994), pp. 291–308. DOI: [10.1016/0010-4655\(94\)90074-4](https://doi.org/10.1016/0010-4655(94)90074-4).

Disclaimer on the use of language tools

In the preparation of this thesis, I made limited use of language tools and large language models (LLMs), for non-substantive editorial purposes. These tools were used to correct typographical and grammatical errors, improve sentence clarity and fluency, and perform light paraphrasing to enhance readability. At no point were these tools employed to generate original content, conduct research, develop arguments, or contribute to the intellectual substance of this work. All ideas, analyses, and conclusions are entirely my own.

Acknowledgements

My deep gratitude goes to my advisors, collaborators, friends, family, and everyone else who has made the past years such a wonderful experience. The work presented in this thesis would not have been possible without the close guidance, idea-sparking discussions, and heart-warming support you gave me.

First and foremost I want to thank Prof. Dr. Thomas Kuhr for being a great supporter of my work, but also beyond. Our discussions always helped me to think clearly and find solutions. You encouraged me to pursue new ideas and created many opportunities for me, which made the last years extremely diverse and incredibly interesting.

Dr. Nikolai Krug, I want to thank for the many great discussions we had over the years. Watching you solve problems and deep-diving into ideas, while always remaining curious and playful, was truly inspiring. I appreciate the love we share for statistics and tough programming puzzles, and everything you taught me.

Prof. Dr. Slavomira Stefkova, you have been a guiding light since the early days of my PhD. I am amazed by your ability to always see things in a positive light, and find amazing compromises. I will never forget the fun we had in all our projects and trips together.

Prof. Dr. Danny van Dyk, your deep expertise in so many topics has always amazed me. It was always a pleasure to work with you, and of course I will never forget our summers in La Clusaz together.

Dr. M eril Reboud, thank you for your guidance on so many occasions. You were always happy to discuss with me and support me in all our projects together. I loved our time in La Clusaz together and was super happy to have a fellow hiking companion.

Prof. Dr. Lukas Heinrich, thank you for being an inspiration in so many ways. I am grateful for our projects together and what I could learn from you, navigating and bringing together so many fields and topics.

I also want to thank Dr. Matthew Feikert and Dr. Giordon Stark for being inspiring colleagues over the years. You have taught me a lot about scientific software.

I want to thank Prof. Dr. Thomas Kuhr, Prof. Dr. Otmar Biebel, Prof. Dr. Danny van Dyk, Prof. Dr. Joseph Mohr and Dr. Sascha Mehlhase for agreeing to be on my defence committee.

I want to thank everyone who has read and reviewed parts of this thesis, and has given me advice along the way. Special thanks go to Dr. Thomas Lueck, Dr. Nikolai Krug, Prof. Dr. Slavomira Stefkova, Dr. Markus Prim, and Caspar Schmitt.

I want to thank all my close colleagues who have made the last years an unforgettable journey. You have guided me through my initial days, until the very end of my thesis work. I felt the deepest support along every step of the way.

Lastly, I want to thank my family and friends, without whom I would definitely not be where I am today. I owe you so much, and will always be thankful for the strength you provide me with. Your support means the universe to me.