THE UNIVERSITY
of ADELAIDE

Measurement of Semileptonic B Decays and Global Fits of Leptoquark Models

Philip Grace

September 2023

Supervisors:
Paul Jackson
Martin White

A thesis submitted toward the degree of
Doctor of Philosphy
The School of Physics, Chemistry and Earth Sciences
The University of Adelaide

Abstract

Flavour physics offers some of the most promising hints of physics beyond the Standard Model (BSM). Precision measurements in the last decade have shown significant deviations from SM predictions, which have been found to favour particular BSM interpretations. This work contributes to two aspects of the ongoing quest for deeper understanding of flavour physics: precision measurements of rare processes at the Belle II experiment, and global fits with flavour constraints in order to make conclusions about specific BSM models.

The Full Event Interpretation (FEI) is a novel tagging algorithm developed by the Belle II collaboration. Here we demonstrate the use of the semileptonic FEI via a normalisation mode approach, which circumvents the ongoing difficulties in calibrating the semileptonic FEI. We use this approach to extract $|V_{ub}|$, obtaining a result that is consistent with world averages and has an uncertainty of 7% that is statistically dominated. There are avenues for improving the semileptonic FEI tagging in the future, and the statistical uncertainty may be greatly reduced with larger datasets.

Anomalous measurements may be placed in a broader context by performing global fits. This allows us to map the high-dimensional parameter spaces of BSM models and determine which regions are favoured by all available results. This thesis presents preliminary results of a global fit of leptoquark models with a selection of flavour constraints, and investigates the viability of observing the non-excluded models at the LHC. We focus in particular on identifying regions of the parameter space which are still viable in light of the recent $R(K^{(*)})$ measurement by LHCb, which removes a previous long-standing anomaly.

This thesis also details software work done regarding Belle II's skimming procedure. This work helps to prepare Belle II for future data collection, when an optimised data processing pipeline will be critically necessary.

Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint award of this degree.

I give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

Philip Grace

1 June 2023

Acknowledgements

I would firstly like to thank my supervisors, Paul Jackson and Martin White, for their mentorship throughout my PhD. Thank you for helping me to maintain the correct perspective when I was deep in the weeds of analysis work. During the many times I thought I was not making enough progress, you brought patience and graciousness, always giving encouragement with a healthy dose of humour.

I would also like to thank the Belle II and GAMBIT communities, both of which provided crucial feedback and suggestions in all aspects of my PhD. In particular, thank you to the Sydney Belle II group for our invaluable weekly meetings.

Two others deserve individual mention for their impact on my Belle II work. Red, thank you for first encouraging me to get involved with Belle II, and for your mentorship and encouragement in the early stages of my PhD. Shanette, thank you for being the best Belle II buddy anyone could ask for, and for putting up with me as an office mate.

A PhD is a difficult undertaking in the best of times—even more so when a global pandemic hits during the second year, shutting down conferences, travel, and all in-person interactions. I owe an enormous debt of gratitude to friends and family for getting me through. Thank you to all the other HDR students in the Physics building; our Zoom calls during lockdown kept me connected, and our in-person chats over coffee were highlights of my PhD. And thank you to my family, both immediate and in-laws, for your continual support.

Finally, thank you Sarah. Thank you for all the support you have provided for me, even while you were working on your own PhD. Neither of us would have gotten through our PhDs in the midst of COVID without each other. As I come to the end of my PhD, there is one key difference from how I started: I'm now married to a wonderful and brilliant woman.

Contents

Part I.

Background

1. Introduction
2. The Standard Model
3. Belle II

1. Introduction

The Standard Model of particle physics (SM) is the framework for modern particle physics, describing all known fundamental particles and their interactions. The SM was developed over the course of the 20^{th} century, culminating in its current formulation around 1970. The story of its development is an interplay between theory and experiment, with unexpected experimental results requiring new theoretical understanding, and theory guiding experiments to investigate unstudied phenomena.

The SM is remarkably successful at predicting the behaviour of fundamental particles. However, there are several areas in which the SM falls short. The SM contains no description of gravitation, and there are serious challenges to making it compatible with general relativity. None of the particles in the SM are appropriate dark matter candidates. Certain SM processes have differing rates for matter and antimatter, due to violation of charge-parity (*CP*) symmetry; however, the *CP* violation in the SM is several orders of magnitude too small to allow for baryogenesis. Finally, in the last decade, experimentalists have measured statistically significant deviations from SM predictions in the realm of flavour physics. These shortcomings continue to drive experimentalists to search for hints of new phenomena, and theorists to build extensions to the SM.

The Belle II experiment is an asymmetric detector built on the site of the SuperKEKB collider, an e^+e^- accelerator operating at the $\Upsilon(4S)$ resonance. A major goal of the Belle II collaboration is to produce large numbers of B mesons, on which to perform precision measurements of the flavour sector of the SM. SuperKEKB has achieved a world record instantaneous luminosity, and the collaboration plans to collect $50\,\text{ab}^{-1}$ of data by 2030. This large dataset will allow for measurements with very low statistical uncertainty that will either confirm or refute the previously measured flavour anomalies.

The flavour anomalies may be hints of physics beyond the SM. In particular, leptoquark models have recently gained popularity due to their ability to explain the flavour anomalies. We can arrive at a complete picture of the constraints placed on leptoquark models by current measurements through a *global fit*. In a global fit, we combine flavour measurements and high energy searches to map the parameter space and determine which parameter combinations are favoured by the available results.

The analysis work in this thesis has two main parts: an analysis of $B \to \pi\ell\nu$ decays in Belle II data, and a global fit to leptoquark models using a selection of flavour constraints. The outline of this thesis is as follows.

In Chapter 2, I introduce the SM by way of its particle content and its formulation as a quantum field theory. I describe how flavour physics emerges from electroweak symmetry breaking, and discuss aspects of B meson physics which are relevant to the analysis in Chapter 5. I then include a brief overview of the current status of the flavour anomalies. Finally, I introduce the leptoquark models which are the focus of the global fits in Chapter 6.

Chapter 3 introduces SuperKEKB and the Belle II detector. I describe the setup of the accelerator and detector, and the background processes inherent in the experiment. The collaboration has built a large software framework to analyse the collected data, and this Chapter includes an overview of the framework design and software stack. I then introduce analysis techniques in B physics which are directly relevant to Chapter 5.

Chapter 4 describes software I developed for the skimming aspect of Belle II's data processing pipeline. This work reduced the amount of boilerplate code in the Belle II collaboration software framework, and reduced the amount of human intervention required during data production campaigns.

Chapter 5 details my analysis of $B \to \pi \ell \nu$ decays using the Full Event Interpretation tagging algorithm (FEI). Previous analyses have predominantly used the hadronic FEI, and this is one of the first analyses to use the semileptonic FEI. The semileptonic FEI is currently uncalibrated, but this Chapter demonstrates a normalisation mode approach which enables its use in the absence of calibration. This Chapter covers all aspects of the analysis, including particle reconstruction, systematic corrections, template fitting, and the final fit to the q^2 spectrum to extract $|V_{ub}|$.

Finally, Chapter 6 shows the results of a global fit to two leptoquark models with flavour constraints. I introduce the idea of global fitting and provide an overview of the software framework GAMBIT which was used to perform the global fit. I then explain the models analysed and the setup of the global fit, and present the results. These results are preliminary, and the analysis has clear directions for future work.

2. The Standard Model

In this Chapter, we will introduce the particle content of the SM, and its mathematical basis. We will discuss how the Higgs field gives masses to the gauge bosons and fermions, and how it also gives rise to the flavour structure of the SM quark sector. We will give an overview of interesting processes involving B mesons, including several recent anomalous measurements which may be indications of Physics Beyond the Standard Model (BSM). We will conclude the Chapter with a discussion of models which may explain these anomalies.

2.1. Particle content

In the SM, matter is composed of fundamental point-like particles, interacting via the strong, weak, and electromagnetic forces [1]. The particles of the SM can be categorised into bosons and fermions, according to whether they have integer spin or half-integer spin, respectively. This particle content is tabulated in Figure 2.1. The fermions comprise the "matter" sector of the particles, while the bosons are the mediators of the fundamental forces. The Higgs boson is unique, as it is the only known spin-0 fundamental particle, and plays a pivotal role in the theory of the electromagnetic and weak forces.

[1]: Thomson (2013), *Modern Particle Physics*

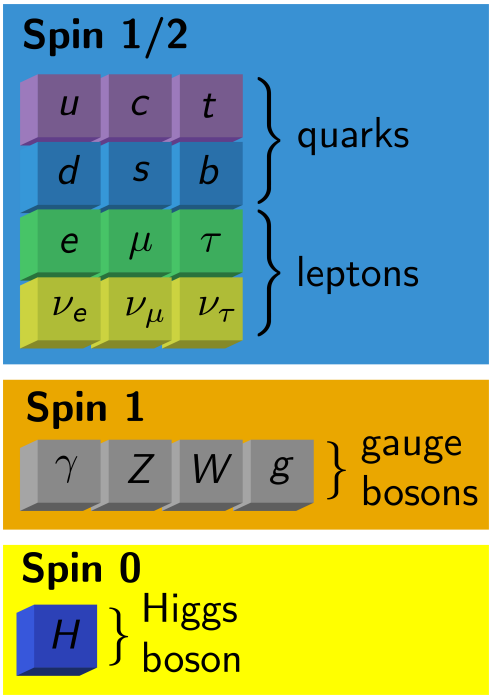


Figure 2.1: Particle content of the Standard Model.

The matter sector of the SM is composed of quarks and leptons. There are six species of quark, which are arranged in three generations. In each generation, there is an "up-type" and a "down-type" quark, with fractional charges of $2/3$ and $-1/3$ times the electron charge, respectively. The lepton sector is also arranged in three generations, with each generation containing a charged lepton and an associated neutrino. The hadrons which comprise atomic nuclei are composed of the lightest of quarks species—u and d as "valence quarks," and u, d, and s as "sea quarks." Studying the heavier quarks requires *high energy physics* (HEP) experiments. The second heaviest species, the b quark, is a primary focus of this thesis.

The spin-1 bosons are force-carrier particles, which mediate interactions between particles which carry the appropriate charge. The photon mediates electromagnetism, which acts on particles with electric charge (*i.e.* quarks, charged leptons, and $W^{\pm}$ bosons). The $W^{\pm}$ and Z^0 bosons mediate the weak force, which acts on particles with *weak isospin*—this includes the left-chiral components of all SM fermions, as well as the $W^{\pm}$ and Z^0 bosons themselves. The gluon mediates the strong force, which acts on particles with *colour charge*; namely, quarks and gluons.

Having described the particles and interactions in the SM, we will now discuss the mathematics of the SM.

2.2. Quantum field theory

[2]: Peskin et al. (1995), *An Introduction to Quantum Field Theory*
[3]: Zee (2003), *Quantum Field Theory in a Nutshell*
[4]: Yang et al. (1954), "Conservation of Isotopic Spin and Isotopic Gauge Invariance"
[5]: Salam (1968), "Weak and Electromagnetic Interactions"
[6]: Weinberg (1967), "A Model of Leptons"
[7]: Glashow (1961), "Partial-Symmetries of Weak Interactions"

The mathematical basis of the SM is *quantum field theory* (QFT), which describes all particles in terms of fields [2–7]. The central object in a QFT is the *Lagrangian density* $\mathscr{L}$ (also referred to as simply the *Lagrangian*). We define the *action* as the integral of the Lagrangian over all spacetime, $S = \int \mathrm{d}^4x\, \mathscr{L}$. The Lagrangian is an object constructed from the fields in the theory, and it fully dictates the interactions of the particles in the theory.

Symmetries play a key role in the formulation of QFTs, as they determine which terms appear in the Lagrangian. To construct a QFT, we first start with some collection of particles, each with an associated field. We then postulate that a chosen set of transformations of the fields should leave the action unchanged—these are the symmetries of the theory. Finally, we write down a Lagrangian containing all combinations of the fields which obey the symmetries. The resulting Lagrangian terms describe all possible interactions in the theory. The symmetries of the SM may be categorised as external symmetries and internal (or gauge) symmetries.

The external symmetries of the SM come from special relativity. Namely, the action is required to be invariant under translations, rotations, and Lorentz boosts. These symmetries form the Poincaré group, and this is the full set of external symmetries of the SM. The previously-mentioned property of spin is a direct consequence of imposing this symmetry.

The internal symmetries of the SM are *gauge symmetries*. These symmetries are with respect to transformations of internal degrees of freedom of the fields. QFTs are often referred to by the group of their gauge symmetries. Gauge transformations are spacetime-dependent transformations applied to a field $\psi(x)$ in the following way,

$$\psi(x) \to \psi'(x) = e^{i\alpha(x)}\psi(x)\,. \tag{2.1}$$

The gauge group of the SM is $\mathrm{SU}(3)_C \otimes \mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$. $\mathrm{SU}(3)_C$ describes the strong force [8, 9], and $\mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$ describes the electroweak force[1]. The interactions of a field are determined by the manner in which they transform under each of these groups. For example, quarks transform as colour triplets under $\mathrm{SU}(3)_C$, whereas leptons are singlets under this group and so do not interact strongly.

[8]: Politzer (1973), "Reliable Perturbative Results for Strong Interactions?"
[9]: Wilson (1974), "Confinement of Quarks"
1: In Section 2.3, we will discuss how the electroweak group is broken to produce the weak and electromagnetic forces.

The gauge symmetries of a QFT determine which *gauge bosons* are present in the theory. A QFT with a given group will contain a set of gauge bosons equal in number to the generators of that group. For example, quantum electrodynamics is a U(1) gauge theory, in which the action is invariant under a global phase rotation of the electron fields,

$$\psi(x) \to \psi'(x) = e^{i\theta}\psi(x)\,. \tag{2.2}$$

In order to preserve gauge invariance under this transformation, we must introduce an additional term to the Lagrangian. This term describes the interaction of the electron field with a massless gauge boson, which we can readily identify as the photon. This procedure does not introduce mass terms, and the self-interaction terms of the photon cancel out because the elements of U(1) commute with each other.

The SM gauge group also contains the non-abelian groups SU(3) and SU(2), and this property has important consequences for the resultant gauge bosons. Here we apply the same procedure of adding terms to the Lagrangian to preserve gauge invariance. However, unlike the abelian case, the self-interaction terms do not cancel, so the gauge bosons carry the charge of the force they mediate. As a result, gluons and $W^{\pm}/Z^0$ bosons self-interact. The strong force has a large coupling constant, so the gluon self-interaction is a unique and defining feature of the study of the strong force, quantum chromodynamics (QCD).

We will briefly mention one final aspect of constructing QFTs: renormalisability [10]. Divergent integrals are unavoidable in QFTs, but renormalisation is a well-defined procedure for absorbing these infinities via a redefinition of the masses and couplings. The effective masses and couplings then become functions of the energy scale of our experiment. It can be shown that in four spacetime dimensions, a Lagrangian is only renormalisable if none of its terms have a mass dimension greater than four[2]. Terms beyond this introduce divergences which cannot be absorbed into the masses and couplings. One can propose an unlimited number of combinations of the fields which obey a chosen gauge symmetry, but the renormalisability requirement restricts the allowed terms to a finite number.

[10]: Wilson et al. (1974), "The renormalization group and the ϵ expansion"

2: Working in natural units, where both $\hbar$ and c are set to 1.

Renormalisation can also be considered from another viewpoint. We only expect the SM to be valid up to some energy scale, beyond which some hypothetical heavier physics comes into effect. Renormalisation is a systematic way of encoding this lack of knowledge. Renormalised theories are only valid up to a certain energy scale—or, equivalently, down to some minimum distance. We will discuss in Section 2.7 how non-renormalisable terms provide methods of probing physics beyond the SM.

2.3. Electroweak symmetry breaking

The Higgs mechanism is a crucial piece of the SM, fulfilling several roles in the model [11–13]. Firstly, it is responsible for breaking the electroweak symmetry group, giving rise to the weak and electromagnetic forces. Secondly, it is through the symmetry breaking that the $W^{\pm}$ and Z^0 bosons acquire mass. Finally, the associated particle, the Higgs boson, also gives mass to the SM fermions through a separate mechanism. The Higgs mechanism is the simplest extension to the SM which achieves these aims. It was proposed in the 1960's, with the Higgs boson put forward as an experimental prediction. In 2012, almost fifty years later, the Higgs boson was discovered by ATLAS and CMS [14, 15]. This discovery provided strong evidence for the Higgs mechanism.

[11]: Higgs (1964), "Broken Symmetries and the Masses of Gauge Bosons"
[12]: Guralnik et al. (1964), "Global Conservation Laws and Massless Particles"
[13]: Englert et al. (1964), "Broken Symmetry and the Mass of Gauge Vector Mesons"

[14]: ATLAS Collaboration (2012), "Observation of a New Particle in the Search for the Standard Model Higgs Boson with the ATLAS Detector at the LHC"
[15]: CMS Collaboration (2012), "Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC"

The Higgs mechanism is formulated in the following way. We start with the assumption that the Higgs field is a complex scalar doublet, $\phi = (\phi^+, \phi^0)^{\mathsf{T}}$, and add a potential term to the SM Lagrangian,

$$V(\phi) = \mu^2 \phi^\dagger \phi + \frac{\lambda^2}{2} (\phi^\dagger \phi)^2 \,. \tag{2.3}$$

For $\mu^2 < 0$, this potential can be visualised to look somewhat like the base of a wine bottle, with an unstable maximum at $\phi = 0$, and a surface of degenerate stable minima. In the early universe, the vacuum expectation value (VEV) of the Higgs field was zero. As the universe cooled, the Higgs

field settled into the stable minimum, acquiring a non-zero VEV, $v \approx 246\,\text{GeV}$.

This non-zero VEV is the driver of the *spontaneous symmetry breaking* of the electroweak group into the weak and electromagnetic forces. The gauge bosons of the electroweak group are labelled W^1, W^2, W^3 and B. All of these bosons are massless, as there is no gauge-invariant way to include mass terms "by hand." However, when we expand the kinetic energy terms of the Higgs field about the non-zero VEV, we obtain mass terms for mixtures of the electroweak bosons. Thus, the mass states of the gauge bosons in the broken symmetry are linear combinations of the massless gauge bosons in the higher-energy theory. The $W^\pm$ bosons are mixtures of W^1 and W^2, and the Z^0 and photon are mixtures of W^3 and B. The $W^\pm$ and Z^0 masses are proportional to the Higgs VEV, and the photon remains massless.

We have now covered two of the roles of the Higgs in breaking the electroweak symmetry and giving the weak gauge bosons mass. We now turn to the third role: giving rise to fermion masses and consequent flavour-changing processes in the SM.

2.4. The physics of flavour

Flavour physics refers to phenomena which can change different species of quarks into each other, and different species of leptons into each other. The flavour structure of the SM is another consequence of the Higgs field. As we will see in this Section, the Higgs provides fermions with mass, and also allows for quarks to transition between flavours.

All SM fermions are massless prior to electroweak symmetry breaking, but they acquire mass in the following way. The SM Lagrangian contains Yukawa terms, which govern the interactions of fermions with the Higgs field. Once the Higgs field acquires its non-zero VEV, we can expand these Yukawa terms around the VEV. In doing so, we obtain gauge invariant mass terms for the fermions. The fermion masses are directly related to the Yukawa couplings, which are free parameters of the SM.

A crucial property of the Yukawa couplings is that they are not diagonal, but couple between quark generations. As a result, the quark flavour eigenstates are not mass eigenstates. We can obtain the mass eigenstates by diagonalising the Yukawa matrices via unitary transformations. From this, we find that the mixing of flavour eigenstates to form the observable states is governed by a 3×3 matrix known as the Cabibbo-

Maskawa-Kobayashi (CKM) matrix,

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{\text{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}, \tag{2.4}$$

where q refers to the quark mass eigenstate, and q' to the quark flavour eigenstate.

The Lagrangian term governing interactions of quarks and $W^{\pm}$ bosons then takes the following form.

$$-\frac{g}{\sqrt{2}} (\bar{u}_L,\ \bar{c}_L,\ \bar{t}_L) \gamma^{\mu} W^{+}_{\mu} V_{\text{CKM}} \begin{pmatrix} \bar{d}_L \\ \bar{s}_L \\ \bar{b}_L \end{pmatrix} + \text{h.c.}\,, \tag{2.5}$$

where q_L is the left-chiral component of the quark field, g is the weak gauge coupling, and γ^{μ} are the Dirac gamma matrices. This interaction term allows for quarks of one generation to transition to another. The amplitudes of the processes $u_i \to d_j$ and $d_j \to u_i$ are proportional to V_{ij} (where i and j are generation indices).

Our understanding of the CKM matrix was developed through both theoretical and experimental advances. In the 1960's, physicists had only discovered the lightest three quark species. Cabibbo proposed a unitary 2×2 matrix parametrised by a single mixing angle to explain the properties of decays involving strange quarks [16]. This led to the prediction of the charm quark. However, this model was not sufficient to explain the violation of charge-parity (CP) symmetry observed in neutral kaon flavour oscillations. A necessary condition of CP violation is that the mixing matrix must have an irreducible complex phase, which is not possible in a 2×2 matrix. In 1973, Kobayashi and Maskawa proposed a 3×3 matrix parametrised by three mixing angles and a complex phase [17]. This model predicted a third generation of quarks[3]. The bottom and top quarks were observed in 1977 and 1995, respectively. Kobayashi and Maskawa received the 2008 Nobel Prize in Physics for their work that predicted CP violation in the neutral B meson system.

[16]: Cabibbo (1963), "Unitary Symmetry and Leptonic Decays"

[17]: Kobayashi et al. (1973), "CP-Violation in the Renormalizable Theory of Weak Interaction"

3: Interestingly, the third generation of quarks was proposed before the discovery of the charm quark in 1974.

The CKM matrix can be parametrised using three mixing angles and a phase factor [17, 18]. However, the more commonly used parametrisation is the *Wolfenstein parametrisation* [19]. This parametrisation makes use of the assumed unitarity of the matrix and the experimentally observed hierarchy of matrix elements. Four $\mathcal{O}(1)$ parameters are introduced—λ, A, ρ and η—and an expansion is performed in powers of λ.

[17]: Kobayashi et al. (1973), "CP-Violation in the Renormalizable Theory of Weak Interaction"

[18]: Chau et al. (1984), "Comments on the Parametrization of the Kobayashi-Maskawa Matrix"

[19]: Wolfenstein (1983), "Parametrization of the Kobayashi-Maskawa Matrix"

$$V_{\text{CKM}} = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}\left(\lambda^4\right) \tag{2.6}$$

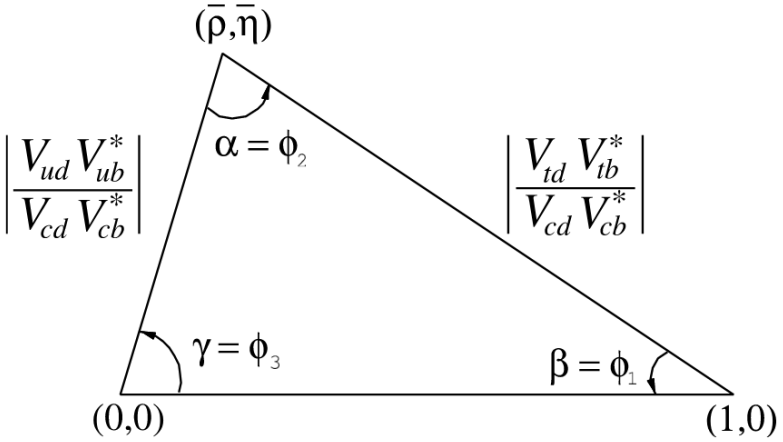


Figure 2.2: Diagram of the CKM triangle. The lower side is fixed to unit length by definition. Two conventions for labelling the three angles were used by the Belle and BaBar collaborations. Figure from [20]

The CKM matrix is assumed to be unitary in the SM; any significant deviation from this would suggest BSM physics in the quark sector. The unitarity requirement $V^\dagger V = VV^\dagger = I$ provides a set of equations which may be used to directly test this assumption. The most commonly used combination is

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0\,. \tag{2.7}$$

The second term in Equation (2.7) has been measured to the highest precision. By dividing all terms in Equation (2.7) by $V_{cd}V_{cb}^*$, we obtain an expression describing a triangle in the complex plane between the points 0, 1, and $\bar{\rho} + i\bar{\eta}$, with

$$\bar{\rho} + i\bar{\eta} = -\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\,. \tag{2.8}$$

This is known as the unitarity triangle, which is illustrated in Figure 2.2. Measuring the side lengths and angles of the triangle is a direct means of testing the unitarity of the CKM matrix.

The CKMfitter and UTFit groups regularly perform combined fits of the CKM triangle, using experimental results as constraints [21, 22]. The most recent fit is shown in Figure 2.3. The coloured bands show the constraints from each experimental measurement.

CKM matrix elements can be measured experimentally by extracting them from measurements of flavour changing processes. In the next Section, we will introduce B mesons, which provide avenues for studying four of the matrix elements.

[21]: CKMfitter Collaboration (2021), *Preliminary Results as of Spring 2021*
[22]: Collaboration et al. (2022), *New UTfit Analysis of the Unitarity Triangle in the Cabibbo-Kobayashi-Maskawa Scheme*

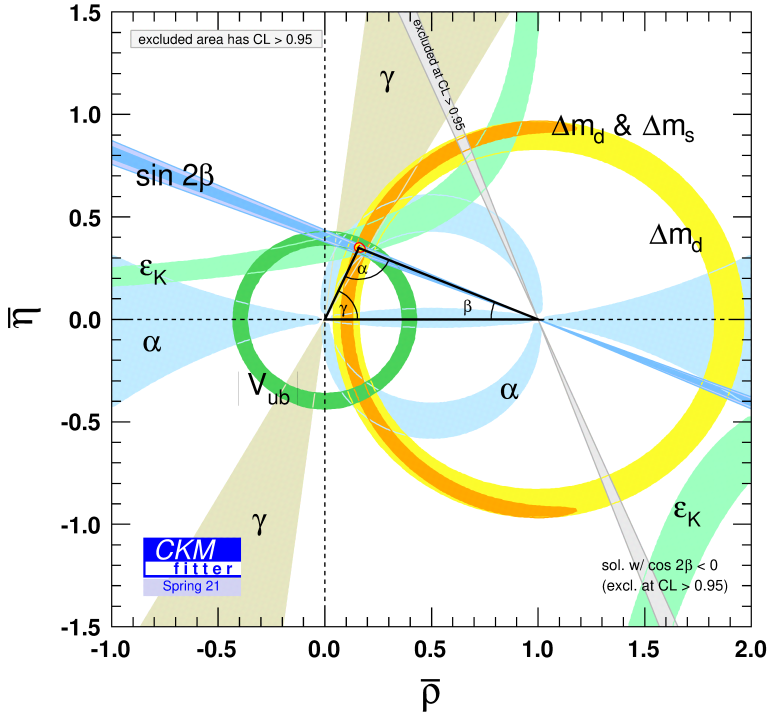


Figure 2.3: Combined fit of various experimental results, in order to test the unitarity of the CKM matrix. Results are current as of Spring 2021 [21].

2.5. B mesons

B mesons are bound states of a b quark and a lighter antiquark. The particular focus of this thesis are the $B^{\pm}$ and $B^0/\bar{B}^0$, which contain a u or d (anti-)quark, respectively. B mesons have a lifetime of approximately 10^{-12} s, and decay to a very wide range of final states. In this Section, we will discuss certain B meson processes which provide insights into the SM [23, 24].

[23]: Kou et al. (2018), "The Belle II Physics Book"
[24]: Bevan et al. (2014), "The Physics of the B Factories"

2.5.1. Neutral B mixing

Neutral B mesons undergo *neutral particle oscillation*, where a B^0 transitions to a $\bar{B}^0$ mid-flight (and vice versa). This phenomenon occurs because the mass states are linear combinations of the observable flavour eigenstates; however, unlike in the kaon system, the mass eigenstates are almost equal to the flavour eigenstates, as the degree of CP violation in this process is very small. The rate of oscillation is governed by the mass difference between the heavy and light mass eigenstates. The dominant Feynman diagrams are shown in Figure 2.4.

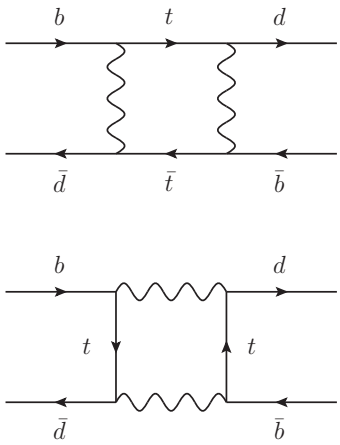


Figure 2.4: Dominant Feynman diagrams of neutral B mixing. These processes also receive contributions from diagrams where the t quark is replaced by a u or c.

This process provides access to several distinct areas of flavour physics, and is an active field of study. Firstly, it provides a means of measuring V_{td}, as this is present in the vertices of Figure 2.4. Additionally, CP violation modifies the mixing probability; this allows for CP violation to

be studied through the novel method of time-dependent B decay measurements. Finally, this process is sensitive to contributions from heavier BSM particles, as it occurs via a loop diagram.

2.5.2. Flavour changing neutral currents

Direct transitions of an up-type quark to another up-type quark are forbidden at tree-level in the SM (and likewise for down-type to down-type). Their leading order contributions come from loop diagrams, and thus they are highly suppressed. These processes are known as *flavour changing neutral currents* (FCNCs). In the context of B physics, the FCNCs of interest are: the radiative penguin processes $b \to s\gamma$ and $b \to d\gamma$; and the electroweak penguin processes $b \to s\ell^+\ell^-$, $b \to d\ell^+\ell^-$, and $b \to s\nu\bar{\nu}$. These decays allow us to measure V_{ts} using B decays. Figure 2.5 shows some of the diagrams by which these processes occur.

FCNC processes are very sensitive to BSM effects. Heavier hypothetical particles may interact via loop diagrams or even provide tree-level diagrams for the decay. Such effects could potentially boost or suppress the low event rate of these processes significantly. Furthermore, these effects can modify observables such as CP asymmetries and angular observables.

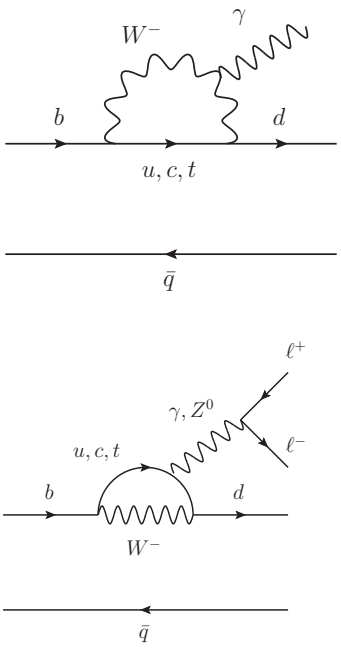


Figure 2.5: Example Feynman diagrams for radiative penguin (top diagram) and electroweak penguin (bottom diagram) processes. The term "penguin" was historically chosen for their superficial resemblance when drawn a particular way.

2.5.3. Hadronic decays

Hadronic B decays are those in which all final state particles are hadrons. There are hundreds of possible hadronic B decays, each with a very small branching fraction. For many of these modes, the QCD calculations are not yet at the point of being able to make precise predictions.

Despite this, several landmark results have been obtained for modes with simpler final states. The Belle collaboration observed direct CP violation in the mode $B^0 \to \pi^+\pi^-$ [25], and the Belle and BaBar collaborations observed direct CP violation in $B^0 \to K^+\pi^-$ [26, 27]. The decay $B \to J/\psi\, K^0_S$ is considered a "golden mode" for Belle II, because of its clean final state, small systematic uncertainty, and relatively high branching fraction [23]; this is a key mode for measuring the CKM triangle angle β.

Hadronic decays are interesting modes of study in their own right; however, their primary relevance to this thesis is their use in the analysis technique known as *tagging*. This technique will be discussed in Section 3.6.1.

[25]: Belle Collaboration (2004), "Observation of Large CP Violation and Evidence for Direct CP Violation in $B^0 \to \pi^+\pi^-$ Decays"

[26]: Belle Collaboration (2004), "Evidence for Direct CP Violation in $B^0 \to K^+\pi^-$ Decays"

[27]: BaBar Collaboration (2004), "Direct CP Violating Asymmetry in $B^0 \to K^+\pi^-$ Decays"

[23]: Kou et al. (2018), "The Belle II Physics Book"

2.5.4. Semileptonic decays

Semileptonic (SL) decays are those with a hadronic part and a leptonic part. SL decays are theoretically clean, as the contributions to the branching fractions can be factored into perturbative terms and non-perturbative QCD terms. They provide an avenue for measuring $|V_{ub}|$ and $|V_{cb}|$ via tree-level processes. SL decays also play a role in tagging.

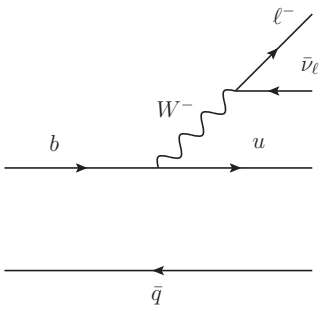


Figure 2.6: Feynman diagram for $B \to \pi\ell\nu$. This is a tree-level process which may be used to measure $|V_{ub}|$.

The SL decay $B \to \pi\ell\nu$ (with $\ell = e, \mu$) occurs via the diagram in Figure 2.6. The decay rate is governed by two form factors, $f_+^{B\pi}(q^2)$ and $f_0^{B\pi}(q^2)$, where q^2 is the invariant mass of the virtual W boson. In the limit $m_\ell \ll m_B$, the former term dominates, leaving the decay rate formula as

$$\frac{\mathrm{d}\Gamma(B \to \pi\ell\nu)}{\mathrm{d}q^2} = \frac{G_F|V_{ub}|^2}{24\pi^3}|\boldsymbol{p}_\pi|^3\left|f_+^{B\pi}(q^2)\right|^2, \tag{2.9}$$

where G_F is the Fermi constant and $\boldsymbol{p}_\pi$ is the 3-momentum of the pion [28, 29].

The quarks in Figure 2.6 are not free, but are bound inside mesons. The impact this has on the decay rate is encoded in the non-perturbative term $f_+^{B\pi}(q^2)$. This term is calculated through lattice QCD and light-cone sum rules (LCSR). Lattice QCD predictions for $B \to \pi\ell\nu$ are most reliable at high q^2; the lattice spacing impacts the range of q^2 that can be simulated, meaning that the values for higher q^2 must be extrapolated down to low q^2 [30, 31]. On the other hand, LCSR calculations are performed by expanding the correlator near the light-cone ($x^2 \simeq 0$); this condition is only fulfilled for low values of q^2, meaning the LCSR results are only obtained for low q^2 [32, 33]. Control of the theoretical uncertainties is critical for extracting $|V_{ub}|$ from this decay.

$|V_{cb}|$ may be measured at tree-level through the decays $B \to D\ell\nu$ and $B \to D^*\ell\nu$ (often denoted together via the shorthand $B \to D^{(*)}\ell\nu$.). These decays may also be factored into perturbative and non-perturbative parts [24].

V_{ub} and V_{cb} may also be extracted from measurements of the inclusive SL branching fraction. In these measurements, no explicit hadronic final state is reconstructed, but instead all hadronic final states are integrated over. The notation for these decays is $B \to X_u\ell\nu$ and $B \to X_c\ell\nu$, depending on whether the hadronic component contains a charm quark. The inclusive charmed SL branching fraction is much greater than the charmless branching fraction, largely due to the fact that $|V_{cb}| \approx 40|V_{ub}|$. Consequently, inclusive measurements of V_{ub} typically make use of kinematic endpoint information to reduce the $X_c\ell\nu$ component.

The inclusive and exclusive measurements are independent and complementary approaches, with different theory inputs. In Section 2.6, we

[28]: Aoki et al. (2019), "FLAG Review 2019"
[29]: BaBar Collaboration (2011), "Study of $B \to \pi\ell\nu$ and $B \to \rho\ell\nu$ Decays and Determination of $|V_{ub}|$"
[30]: Fermilab Lattice Collaboration et al. (2015), "$|V_{ub}|$ from $B \to \pi\ell\nu$ decays and (2+1)-flavor lattice QCD"
[31]: Flynn et al. (2015), "$B \to \pi\ell\nu$ and $B_s \to K\ell\nu$ form factors and $|V_{ub}|$ from 2+1-flavor lattice QCD with domain-wall light quarks and relativistic heavy quarks"
[32]: Khodjamirian et al. (2005), "B-Meson Distribution Amplitude from the $B \to \pi$ Form Factor"
[33]: Gubernari et al. (2019), "$B \to P$ and $B \to V$ Form Factors from B-Meson Light-Cone Sum Rules beyond Leading Twist"
[24]: Bevan et al. (2014), "The Physics of the B Factories"

will see how aspects of these measurements are in tension with one another.

2.5.5. Leptonic decays

The purely leptonic decay $B^+ \to \ell^+ \nu_\ell$ occurs via the Feynman diagram shown in Figure 2.7. This is a tree-level process with a relatively simple branching fraction expression [34],

$$\mathcal{B}(B^+ \to \ell^+ \nu_\tau) = \frac{G_F^2 m_B m_\ell^2}{8\pi} \left(1 - \frac{m_\ell^2}{m_B^2}\right)^2 f_B^2 |V_{ub}|^2 \tau_B \,, \tag{2.10}$$

where m_B and m_ℓ are the masses of the B and lepton, and τ_B is the B lifetime—all of these are well measured quantities. f_B is the decay constant, which is calculated using Lattice QCD.

The mode $B^+ \to \tau^+ \nu_\tau$ provides an independent method for directly measuring $|V_{ub}|$. Additionally, this process could be significantly enhanced or suppressed by the presence of a charged Higgs, which several BSM models predict.

[34]: UTfit Collaboration et al. (2010), "An Improved Standard Model Prediction Of BR($B \to \tau\nu$) And Its Implications For New Physics"

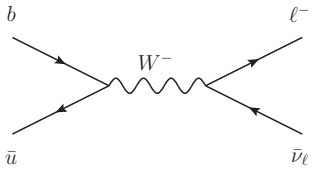


Figure 2.7: Feynman diagram for $B^- \to \ell^- \nu_\ell$. This process provides a very clean way to measure $|V_{ub}|$. It is sensitive to contributions from BSM physics.

2.6. Flavour anomalies

In the last decade, LHCb and the B factories have made various measurements of the flavour sector which disagree with the SM prediction at a 2–4σ significance. These have come to be collectively known as the *flavour anomalies* [35]. The current status of a selection of the anomalies is summarised in Figure 2.8.

[35]: London et al. (2022), "B Flavour Anomalies"

The flavour anomalies are currently the most tantalising hints of physics beyond the SM. Thus far, direct searches have not led to discoveries of any BSM particles. However, the processes in which the flavour anomalies were discovered are sensitive to effects from much higher mass scales than the LHC can reach. In this way, the anomalies provide indirect evidence of the structure and interactions of the higher mass physics.

In this Section, we primarily discuss the anomalies discovered in B decays, as these are the anomalies that the Belle II experiment is most suited to investigating. We will also discuss the anomalous magnetic moment of the muon, as it has important impacts on the BSM models discussed in Section 2.7.

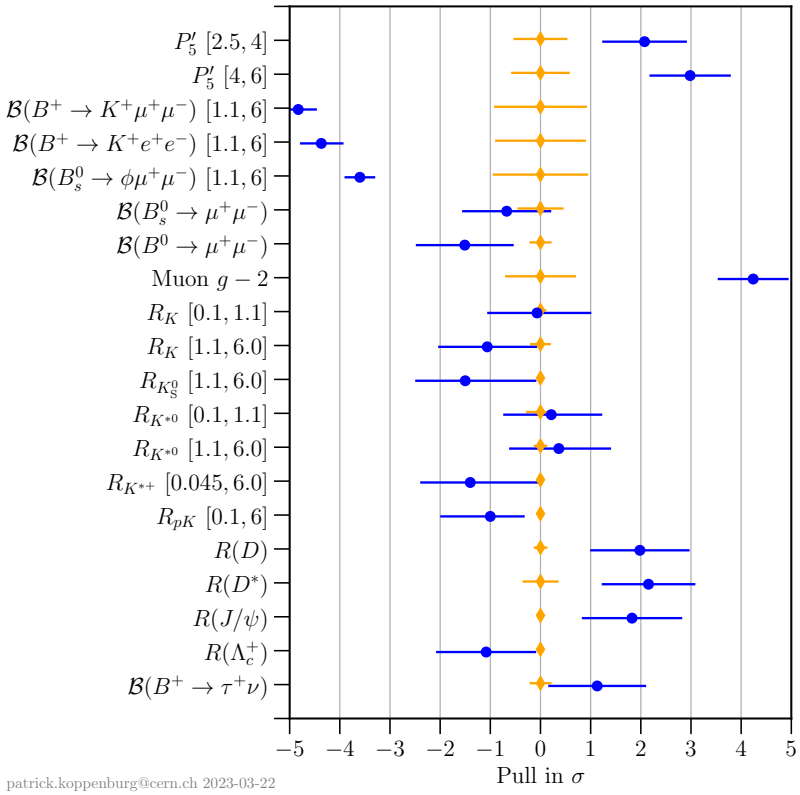


Figure 2.8: Summary of current status of a selection of flavour anomalies. From Reference [36] (see for full explanation and citations of each anomaly plotted). The orange dots are the SM predictions with attached theoretical uncertainties (with nominal value shifted to zero), and the blue dots are the experimental measurements, plotted as a signed significance. This Section will primarily discuss the results $(g-2)_\mu$, $R(D^{(*)})$, and $R(K^{(*)})$.

2.6.1. Ratio measurements

The SM gauge bosons couple equally to all charged leptons, so the couplings themselves do not prefer one lepton flavour over the others. The only factor impacting the rates of (semi)leptonic processes is the lepton mass, and consequently the phase space available to the decay. This property of the SM is known as *lepton flavour universality* (LFU). B decays provide several methods for testing LFU, and some of these tests have produced anomalous results. These may be evidence of BSM models with flavour-dependent couplings.

Semileptonic decays provide an ideal environment for studying LFU, as their decay rates are factored into hadronic and leptonic parts. By looking at two decays differing only in lepton flavour and taking their ratio, we can effectively reduce the theoretical uncertainty due to hadronic effects. The ratios $R(K^{(*)})$ and $R(D^{(*)})$ are two ratios for testing LFU, and are defined in Equations (2.11) and (2.12).

$$R(K^{(*)}) = \frac{\mathcal{B}(B \to K^{(*)}\mu\mu)}{\mathcal{B}(B \to K^{(*)}ee)} \tag{2.11}$$

$$R(D^{(*)}) = \frac{\mathcal{B}(B \to D^{(*)}\tau\nu)}{\mathcal{B}(B \to D^{(*)}\ell\nu)}, \quad (\ell = e, \mu) \tag{2.12}$$

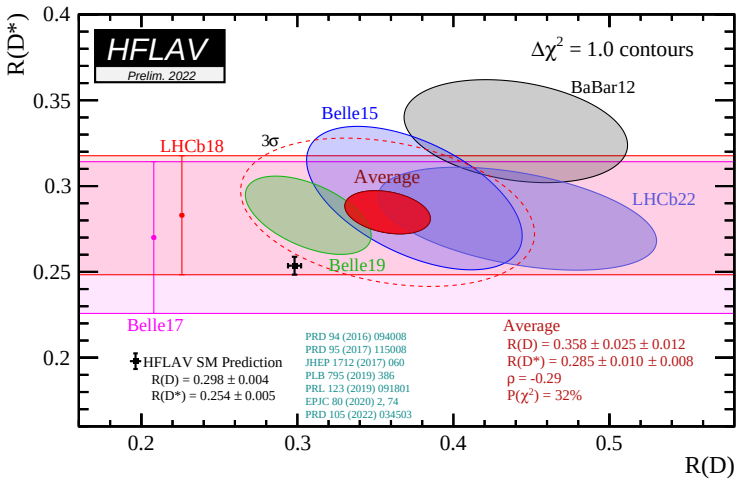


Figure 2.9: Global fit of $R(D^{(*)})$ measurements, performed by HFLAV [38]. The combined significance of the $R(D)$ and $R(D^*)$ anomalies is 3.2σ.

The electron and muon masses are both much less than the kaon mass, so the SM prediction for $R(K^{(*)})$ is very close to one. The theoretical uncertainty of these ratios is low, as they do not receive large higher-order corrections. LHCb, Belle and BaBar have measured these ratios, with the former observing a deficit in the muon modes. However, a recent LHCb analysis identified a previously unaccounted-for background, and obtained a result consistent with the SM [37]. The Belle II experiment is well suited to measuring $R(K^{(*)})$, as it has a similar reconstruction efficiency for electrons and muons.

[37]: LHCb Collaboration (2022), "Measurement of lepton universality parameters in $B^+ \to K^+\ell^+\ell^-$ and $B^0 \to K^{*0}\ell^+\ell^-$ decays"

$R(D^{(*)})$ is measured from tree-level processes, and provides a test of the coupling of BSM physics to third generation leptons compared to light leptons. The processes in the numerator and denominator have very similar final states, so techniques are required to distinguish the decays using their different kinematics. $R(D^{(*)})$ measurements at the B factories commonly utilise the tagging technique discussed in Section 3.6.1.

[38]: HFLAV Collaboration (2023), *Averages of b-hadron, c-hadron, and τ-lepton properties as of 2021*

The most recent $R(D^{(*)})$ global fit performed by HFLAV is shown in Figure 2.9 [38]. Both $R(D)$ and $R(D^*)$ have been found to exceed the SM prediction, with significances of 2.16σ and 2.26σ, respectively. Taken together, the combined significance of the two anomalies is 3.2σ.

Numerous models have been proposed to explain these anomalies, including leptoquark models and charged Higgs bosons from two-Higgs-doublet models. Leptoquark models will be discussed in Section 2.7, and are the focus of Chapter 6. Additional measurements can provide insight into the properties of the BSM effects; these include the polarisation of the τ and D^* [39], and the q^2 spectrum and angular distributions of the $B \to D^*\tau\nu$ decay.

[39]: Belle Collaboration (2017), "Measurement of the τ lepton polarization and $R(D^*)$ in the decay $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$"

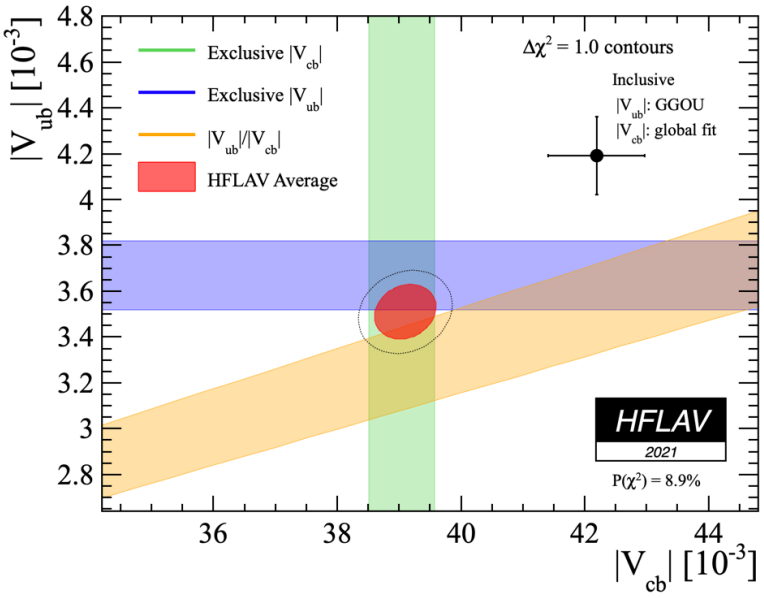


Figure 2.10: Global fit result of V_{ub} and V_{cb}. Coloured bands show constraints from exclusive measurements, and the black point with error bars shows the inclusive measurements [38].

2.6.2. Inclusive and exclusive $|V_{ub}|$ and $|V_{cb}|$

$|V_{ub}|$ and $|V_{cb}|$ may be cleanly measured using semileptonic B decays. The two experimental approaches to this are:

- *exclusive*, where a particular final state is explicitly reconstructed, and
- *inclusive*, where a lepton is reconstructed and all other final state particles are assigned to an X system.

These approaches are complementary, as they differ both in experimental technique as well as theoretical inputs. There are long-standing disagreements between the inclusive and exclusive determinations of $|V_{ub}|$ and $|V_{cb}|$, with current tensions of 3.5σ and 3σ, respectively [40, 41]. The most recent HFLAV global fit of $|V_{ub}|$ and $|V_{cb}|$ is shown in Figure 2.10. It remains to be seen whether these discrepancies will persist as the theoretical understanding improves.

The theoretical description of inclusive semileptonic decays uses the fact that only the long-range effects of the B meson impact these decays, allowing for *operator product expansion* to be employed [20]. The theoretical limitations of this approach stem from the lack of knowledge of non-perturbative parameters such as the b quark mass and the form of the shape functions. The theoretical description of inclusive $|V_{ub}|$ measurement is further complicated by the experimental necessity to separate out the $X_c\ell\nu$ background using kinematic endpoints. However, measurements of differential distributions of inclusive decays can provide independent information to reduce this theoretical uncertainty. These measurements are currently statistics-limited, but are one of the focuses of the Belle II experiment.

[40]: Petrella (2009), "Inclusive and Exclusive $|V_{ub}|$"
[41]: Bouchard et al. (2019), *Summary of the 2018 CKM working group on semileptonic and leptonic b-hadron decays*

[20]: Particle Data Group (2022), "Review of Particle Physics"

In the exclusive case, the theoretical description of non-perturbative hadronic effects relies on the calculation of hadronic form factors. The lattice QCD calculations of form factors are performed using heavy quark effective theory, which exploits the large mass difference between the light quarks and b quark to expand in terms of $\Lambda_{\mathrm{QCD}}/m_b$. The matching procedure required by the effective field theory treatment introduces a leading systematic uncertainty—this uncertainty may be reduced in the future with calculation methods that do not require perturbative matching [42]. The large amount of data to be collected at Belle II will allow for differential measurements which may shed further light on this puzzle.

[42]: McLean et al. (2019), $B_s \to D_s^{(*)} l\nu$ *Form Factors with Heavy HISQ Quarks*

2.6.3. Muon magnetic moment

We will conclude this Section with a recent high-profile anomaly which has implications for the models discussed in Section 2.7. This is the anomalous magnetic moment of the muon, quoted as the quantity $a_\mu = (g-2)_\mu/2$. In 2006, an experiment at Brookhaven National Laboratory measured an a_μ value higher than the SM expectation, with a significance of 2.2–2.7σ [43]. This anomaly was confirmed by Fermilab in 2021, with a measurement that increased the tension to 4.2σ [44].

The measurement of a_μ is a sensitive test of the SM, as its precise value is dependent on the contributions from virtual particles. These include electroweak processes and hadronic vacuum polarisation. Several BSM models which provide loop contributions have been put forward as explanations for the anomaly [45, 46]. The calculation of hadronic contributions is particularly challenging, and currently relies on measurements of the $e^+e^- \to$ hadrons cross-section. However, independent determinations from lattice QCD conflict with these measurements [47]; it remains to be seen whether improving the accuracy of the SM prediction will resolve the anomaly.

[43]: Muon g–2 Collaboration (2006), "Final Report of the Muon E821 Anomalous Magnetic Moment Measurement at BNL"

[44]: Abi et al. (2021), "Measurement of the Positive Muon Anomalous Magnetic Moment to 0.46 Ppm"

[45]: Crivellin et al. (2019), *Combined explanations of* $(g-2)_\mu$, $(g-2)_e$ *and implications for a large muon EDM*

[46]: Athron et al. (2021), "New physics explanations of a_μ in light of the FNAL muon $g-2$ measurement"

[47]: Borsanyi et al. (2021), "Leading Hadronic Contribution to the Muon Magnetic Moment from Lattice QCD"

2.7. Physics beyond the Standard Model

The present formulation of the SM was arrived at in the 1970's. Since then, numerous extensions have been proposed to address particular shortcomings of the SM. These extensions are collectively termed "BSM models," and include supersymmetry, two-Higgs-doublet models, and leptoquarks, to name just a few. The search for evidence of these models is one motivation for HEP experiments including Belle II and the experiments at the LHC.

The SM cannot be a complete theory of nature, as there are observed phenomena that it does not account for: the presence of dark matter, the

presence of large *CP* violating sources, and the incompatibility with general relativity. Furthermore, there are questions raised by the seemingly arbitrary structure of the SM: why are there three generations of quarks and three generations of leptons, and does the SM gauge group arise from a larger group? Further motivation for this claim comes from fact that quark and lepton charges seemingly conspire to allow for neutral atoms, which is not a requirement of the underlying theory. Finally, in the last decade, there has been renewed interested in BSM models which can explain the anomalies discussed in Section 2.6. At the time of writing, there has been no experimental confirmation of any BSM model.

In this Section we will introduce leptoquark models, and the general framework of effective field theory, which is a model-independent way of studying the effects of BSM physics in low-energy experiments.

2.7.1. Leptoquark models

The history of modern physics is marked by unification, the idea that two seemingly distinct phenomena are in fact manifestations of a single more fundamental phenomenon. Electricity and magnetism were discovered to be aspects of electromagnetism, and electromagnetism and the weak force were found to be low-energy manifestations of the electroweak force. We naturally ask why this trend should not continue, unifying the electroweak and strong forces in a Grand Unified Theory (GUT), and ultimately a GUT with gravity. On the other hand, we may ask whether the SM quarks and leptons are also manifestations of a single field, which would give us *matter unification*.

Leptoquarks (LQs) are not a model on their own, but are an inevitable outcome of almost all attempts at unification. One of the first attempts at SM unification was the SU(5) model, proposed in 1974 by Georgi and Glashow [48]. In this model, the quarks and leptons of the SM are placed in a single representation. This leads to a direct coupling of quarks to leptons, which is an interaction that does not exist in the SM. The coupling is mediated by a new set of gauge bosons, known as leptoquarks. Processes which can only occur at loop level in the SM are possible at tree level with LQs. Furthermore, LQs allow for processes such as proton decay which are not possible in the SM. Since FCNCs have been measured to be very rare and proton decay has not been observed at all, there are stringent constraints on the allowed couplings of LQs. The bare SU(5) model has been ruled out by measurements of the proton lifetime. However, other GUTs are still viable candidates for unification, including SO(10) with an SU(4) colour symmetry, SU(15), and superstring-inspired E_6 [49].

[48]: Georgi et al. (1974), "Unity of All Elementary-Particle Forces"

[49]: Witten (2002), "Quest For Unification"

Despite their origin in a wide range of BSM models, LQs can be studied in a relatively model-independent way [50, 51]. This is done via an

[50]: Hewett et al. (1997), "Much Ado About Leptoquarks"

[51]: Doršner et al. (2016), "Physics of leptoquarks in precision experiments and at particle colliders"

Label	Spin	F	SM group representation
R_2	0	0	$(\mathbf{3}, \mathbf{2}, 7/6)$
$\tilde{R}_2$	0	0	$(\mathbf{3}, \mathbf{2}, 1/6)$
S_1	0	-2	$\left(\bar{\mathbf{3}}, \mathbf{1}, 1/3\right)$
$\tilde{S}_1$	0	-2	$\left(\bar{\mathbf{3}}, \mathbf{1}, 4/3\right)$
S_3	0	-2	$\left(\bar{\mathbf{3}}, \mathbf{3}, 1/3\right)$
U_1	1	0	$(\mathbf{3}, \mathbf{1}, 2/3)$
$\tilde{U}_1$	1	0	$(\mathbf{3}, \mathbf{1}, 5/3)$
U_3	1	0	$(\mathbf{3}, \mathbf{3}, -1/3)$
V_2	1	-2	$\left(\bar{\mathbf{3}}, \mathbf{2}, 5/6\right)$
$\tilde{V}_2$	1	-2	$\left(\bar{\mathbf{3}}, \mathbf{2}, -1/6\right)$

Table 2.1: All possible LQ species, assuming only the SM fermion content. They may be broadly categorised by spin, fermion number ($F = 3B + L$), and their representation under the SM group. The subscript index indicates the dimension of the $\mathrm{SU}(2)_L$ representation.

effective Lagrangian which minimally extends the SM by the addition of a small number of LQ fields. We construct the most general renormalisable Lagrangian which obeys the SM gauge symmetry, using only the SM fermion content[4]. The resulting Lagrangian contains five scalar LQ fields and five vector LQ fields; these are tabulated in Table 2.1. This table includes each LQ's fermion number $F = 3B + L$ (where $B = \frac{1}{3}$ for quarks, and $L = 1$ for leptons) and its representation under the SM group, $\mathrm{SU}(3)_C \otimes \mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$. A given BSM model will typically only predict a subset of these, so the discovery of any one LQ would be evidence in favour of particular models.

4: If we include additional fermion fields, such as right-handed neutrinos, then more LQ couplings would arise.

The quantum numbers of each of these fields are determined requiring that contractions of LQs with two fermions be gauge invariant. This requires that every species of LQ be a colour triplet. Consequently, some LQs have diquark couplings (namely, the $|F| = 2$ species), but there are no gauge-invariant lepton-lepton-LQ couplings. Quark-lepton contractions may be singlets, doublets, or triplets under $\mathrm{SU}(2)_L$, and so the $\mathrm{SU}(2)_L$ representation of LQs may be any of these. The hypercharge simplest to determine, since it is additive. The quantum numbers are used to distinguish leptoquarks, and inform the notation of labels shown in Table 2.1.

The analysis performed in Chapter 6 focuses on R_2, S_1, and S_3. The Lagrangian terms describing the interactions of each of these LQs to SM fermions are shown in Equations (2.13) to (2.15). In these equations, τ^k are the Pauli matrices; y and z are arbitrary Yukawa coupling matrices, with the superscript indicating the chirality of the coupling; i, j and k are flavour indices; a and b are SU(2) indices; and Q_L, L_L, e_R, u_R, and d_R are the SM fermion multiplets.

$$\mathscr{L}_{R_2} = -y_{2\,ij}^{RL}\bar{u}_R^i R_2^a \epsilon^{ab} L_L^{j,b} + y_{2\,ij}^{LR}\bar{e}_R^i R_2^{a\,*} Q_L^{j,a} + \text{h.c.} \tag{2.13}$$

$$\begin{aligned}\mathscr{L}_{S_1} = {} & y_{1\,ij}^{LL}\bar{Q}_L^{C\,i,a} S_1 \epsilon^{ab} L_L^{j,b} + y_{1\,ij}^{RR}\bar{u}_R^{C\,i} S_1 e_R^j \\ & + z_{1,ij}^{LL}\bar{Q}_L^{C\,i,a} S_1^* \epsilon^{ab} Q_L^{j,b} + z_{1,ij}^{RR}\bar{u}_R^{C\,i} S_1^* d_R^j + \text{h.c.}\end{aligned} \tag{2.14}$$

$$\mathscr{L}_{S_3} = y_{3\,ij}^{LL}\bar{Q}_L^{C\,i,a}\epsilon^{ab}(\tau^k S_3^k)^{bc} L_L^{j,c} + z_{3\,ij}^{LL}\bar{Q}_L^{C\,i,a}\epsilon^{ab}((\tau^k S_3^k)^\dagger)^{bc} Q_L^{j,c} + \text{h.c.} \tag{2.15}$$

The free parameters are the LQ masses and the fermionic Yukawa couplings. The vector LQs have only gauge couplings, which are not free parameters, but are derived from the effective field theory. Low-energy measurements place strong constraints on the Yukawa couplings, since LQs can mediate many processes which are rare or forbidden in the SM.

For example, LQs contribute to the decay $\pi^+ \to e^+ \nu_e$. The low measured branching fraction of this decay gives rise to constraints of the form $M_{\mathrm{LQ}} > C\sqrt{\lambda_R \lambda_L}$, where λ_L and λ_R are the couplings to the left- and right-chiral lepton fields. In order for the mass to be low enough and LQ-induced processes to be prevalent enough to be of experimental interest, one of λ_L or λ_R must vanish. That is, the Yukawa couplings must be chiral.

The absence of proton decay places strong constraints on LQ models. In the SM, protons are absolutely stable, since lepton number is conserved exactly and baryon number is conserved in perturbative calculations. This is confirmed by measurements of the proton half-life, which place a lower bound greater than the age of the universe [52]. LQs can violate baryon and lepton number individually, but this is not necessarily problematic. These can only lead to proton decay if the LQ has diquark couplings, and has a low mass and large coupling. The R and U species only have quark-lepton couplings, and so cannot contribute to proton decay in this way. The lower bound on V masses places them far beyond the reach of present and planned colliders, as their gauge couplings are not free parameters. The S species do possess diquark couplings, but the Yukawa couplings can be tuned to avoid proton decay while leaving their masses experimentally accessible.

[52]: Super-Kamiokande Collaboration (2020), "Search for proton decay via $p \to e^+\pi^0$ and $p \to \mu^+\pi^0$ with an enlarged fiducial volume in Super-Kamiokande I-IV"

Much of the current interest in LQ models is due to their ability to account for the flavour anomalies. These anomalies were found in semileptonic decays, FCNCs, and CP-violating processes. Such processes are rare in the SM, and some are only allowed at loop level. LQ models can mediate these processes at tree level, and so give significant enhancements to their rates. Global fits have been performed to test single LQ explanations for the flavour anomalies [53–55]. U_1, S_1 and R_2 can each individually explain the $R(D^{(*)})$ anomalies [56], and S_3 individually explained the previously-anomalous $R(K^{(*)})$ result [57]. If we allow for combinations of LQs, then processes can receive multiple contributions from different LQs.

[53]: Angelescu et al. (2018), "Closing the window on single leptoquark solutions to the B-physics anomalies"

[54]: Bauer et al. (2015), "One Leptoquark to Rule Them All"

[55]: Capdevila et al. (2018), "Patterns of New Physics in $b \to s\ell^+\ell^-$ transitions in the light of recent data"

[56]: Sakaki et al. (2013), "Testing leptoquark models in $\bar{B} \to D^{(*)}\tau\bar{\nu}$"

[57]: D'Amico et al. (2017), "Flavour anomalies after the R_{K^*} measurement"

Leptoquarks can also be searched for directly in colliders, either in single production or pair production process. Proton-proton colliders are well suited to LQ direct searches, since LQs carry colour charge. However, pair production can also be achieved in e^+e^- experiments such as Belle II through LQ coupling to either electroweak gauge bosons or to electrons. The physics program of Belle II includes searches for lepton flavour violating final states, which could be evidence of LQs [23].

[23]: Kou et al. (2018), "The Belle II Physics Book"

A common approach in LHC searches for LQs is to analyse a *simplified model*, where a LQ is assumed to couple to only a single generation, and have two decay modes—frequently a quark and charged lepton, and a quark and neutrino. The free parameters of these models are the LQ mass and the branching fraction between the two modes. The measured lower limits on LQ masses in these models are on the order of 1 TeV [58–61]. The observation of LQs through direct searches would be a revolutionary discovery that may shed light on other outstanding puzzles in the SM.

[58]: ATLAS Collaboration (2021), *Search for pair production of third-generation scalar leptoquarks decaying into a top quark and a τ-lepton in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*

[59]: ATLAS Collaboration (2021), "Search for new phenomena in pp collisions in final states with tau leptons, b-jets, and missing transverse momentum with the ATLAS detector"

[60]: ATLAS Collaboration (2021), *Search for new phenomena in final states with b-jets and missing transverse momentum in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector*

[61]: CMS Collaboration (2021), *Search for singly and pair-produced leptoquarks coupling to third-generation fermions in proton-proton collisions at $\sqrt{s} = 13$ TeV*

[62]: Neubert (2006), "Effective Field Theory and Heavy Quark Physics"

2.7.2. Effective field theory

Effective field theory (EFT) is a framework for obtaining low-energy experimental predictions from a theory without a complete high-energy description [62]. As such, it is a powerful tool for probing physics beyond the reach of current experiments. Here we will introduce the EFT formalism, and describe how precision measurements in the flavour sector can provide insight into physics beyond the SM.

We start with a QFT with a large fundamental mass scale M—this mass could be the mass of a heavy particle in the theory, or the typical scale of some interaction. We wish to derive a description of interactions at energy scales E greatly below M. We choose a cutoff $\Lambda < M$ and split the fields into high- and low-frequency modes, $\phi = \phi_L + \phi_H$. All of the low-energy physics is described by the low-frequency modes, so the EFT will ultimately be written only in terms of ϕ_L.

In general, all the information we wish to obtain from any QFT can be calculated from functional derivatives of the generating functional, $Z[J]$, where J is a source field. Having split the fields into high- and low-frequency modes, we can write the generating functional for the low-energy physics as

$$Z[J_L] = \int \mathscr{D}\phi_L \mathscr{D}\phi_H e^{iS(\phi_L,\phi_H)+i\int \mathrm{d}^D x J_L(x)\phi_L(x)} \,. \tag{2.16}$$

Here we only include source fields for the low-frequency modes, J_L, as this is sufficient to describe the low-energy physics. We then perform the path integral over the high-frequency fields, ϕ_H, to obtain the Wilsonian effective action,

$$e^{iS_\Lambda(\phi_L)} = \int \mathscr{D}\phi_H e^{iS(\phi_L,\phi_H)} \,. \tag{2.17}$$

Having done this, the high-energy physics beyond Λ has been "integrated out" of the theory. As a consequence, S_Λ is non-local at distance scales below $\sim 1/\Lambda$.

The final step of the EFT construction is to expand the non-local effective action as a series of local operators—this is known as *operator product expansion*. We rewrite the effective action as the integral over the effective Lagrangian, $\mathscr{L}_\Lambda$, which we then express as

$$\mathscr{L}_\Lambda(x) = \sum_i C_i \mathcal{O}_i(\phi_L(x)) \,, \tag{2.18}$$

where $\mathcal{O}_i$ are local operators composed of the light fields, and C_i are known as the *Wilson coefficients*. The operators describe the long-distance (low-energy) interactions, and the short-distance (high-energy) effects are entirely factored out into the Wilson coefficients.

The effective action is dependent on the cutoff scale. However, the operators are composed of the low-frequency fields, and so are scale-independent. Any change in the cutoff must be absorbed into the Wilson coefficients. Consequently, the Wilson coefficients are running couplings that depend on the scale of a given interaction.

Equation (2.18) is in principle an infinite sum over all gauge-invariant field combinations. However, in practice only a finite number of these are of interest to us. As mentioned in Section 2.2, operators are only renormalisable if their mass dimension is four or less. However, in the EFT framework, non-renormalisable terms are not explicitly forbidden, but provide small contributions to the low-energy physics. By performing precision measurements of the Wilson coefficients, we can learn about the structure of the high-energy physics from low-energy interactions. The infinite sum in Equation (2.18) is reduced to a finite series of terms by retaining the renormalisable terms and a small number of non-renormalisable terms. The decision of which higher order terms to keep is informed by the target precision—in flavour physics, the non-renormalisable terms that are studied have mass dimension six.

Now that the effective Lagrangian has been reduced to a finite sum, we are able to calculate the Wilson coefficients. This is done by a procedure known as "matching": we calculate the Wilson coefficients by requiring that all matrix elements of the effective Lagrangian are equal to the matrix elements of the full theory. If the theory is weakly coupled at the cutoff scale, then calculation of the matrix elements can be greatly simplified by evaluating them perturbatively at Λ. We then run the Wilson coefficients back down to low energies, using the renormalisation group.

Flavour EFT

The flavour EFT is an EFT of the SM that is useful for studying rare B decays and FCNCs [24, 62]. The cutoff is set above the bottom quark mass so that the electroweak bosons and top quark are integrated out.

[24]: Bevan et al. (2014), "The Physics of the B Factories"
[62]: Neubert (2006), "Effective Field Theory and Heavy Quark Physics"

The effective Lagrangian for FCNC processes is written in terms of the following operators:

- $\mathcal{O}_1$ and $\mathcal{O}_2$: the current-current operators,
- $\mathcal{O}_3$–$\mathcal{O}_6$: the QCD penguin operators, and
- $\mathcal{O}_7$–$\mathcal{O}_{10}$: the electroweak penguin operators.

The expected values of the associated Wilson coefficients, C_1–C_{10}, are calculated by matching to the SM.

Following the operator product expansion, the high-energy contributions to B decays are separated out into the Wilson coefficients. FCNCs are forbidden at tree-level in the SM, so the only contributions to C_1–C_{10} come from loop diagrams. In particular, $\mathcal{O}_9$ and $\mathcal{O}_{10}$ occur in $b \to s\ell^+\ell^-$ and $b \to s\nu\bar{\nu}$ decays; these are shown in Equations (2.19) and (2.20).

$$\mathcal{O}_9^{sb\ell\ell'} = \frac{e^2}{16\pi^2}(\bar{s}_L\gamma_\mu b_L)(\bar{\ell}\gamma^\mu\ell') \tag{2.19}$$

$$\mathcal{O}_{10}^{sb\ell\ell'} = \frac{e^2}{16\pi^2}(\bar{s}_L\gamma_\mu b_L)(\bar{\ell}\gamma^\mu\gamma_5\ell') \tag{2.20}$$

For a given BSM model, we can calculate the contributions to the Wilson coefficients by matching the matrix elements of the effective flavour theory to the full BSM model. Since the same Wilson coefficients appear in multiple processes, different measurements can be combined to constrain their values. This then places constraints on the types of BSM models that are compatible with precision measurements. This will be the focus of Chapter 6.

2.8. Conclusion

In this Chapter, we have introduced the Standard Model of Particle Physics, and discussed its particle content and mathematical basis. We discussed how the flavour structure of the SM arises from electroweak symmetry breaking, and covered several areas of B meson physics, which is the major focus of the Belle II experiment. In the field of flavour physics, and B physics especially, there have been recent measurements that disagree with the SM prediction. We described a selection of these anomalies, and introduced leptoquark models, which have been put forward as possible explanations of the anomalies.

Belle II 3.

The Belle II experiment is a particle detector built at the site of the SuperKEKB accelerator, in Tsukuba, Japan. The analysis in Chapter 5 uses data collected from the detector. In this Chapter, we will introduce the accelerator and detector, and describe their design and operation. We will then describe the software used to analyse the detector data, which is known as the Belle II Analysis Software Framework. Finally, we will describe analysis techniques relevant to this thesis—in particular, the Full Event Interpretation [63, 64], which is an novel algorithm used in many Belle II analyses.

[63]: Keck (2015), *The Full Event Interpretation for Belle II*
[64]: Keck et al. (2018), "The Full Event Interpretation – An exclusive tagging algorithm for the Belle II experiment"

3.1. The B factories

B mesons are particularly interesting objects of study in modern particle physics, for several reasons. B meson decays provide methods of directly measuring the CKM matrix elements $|V_{ub}|$ and $|V_{cb}|$, and the CKM triangle angle γ/ϕ_3. Neutral B mesons oscillate flavours between B^0 and $\overline{B}^0$ during flight, which gives rise to time-dependent CP violation effects. High-precision measurements of certain B decays are also sensitive to contributions from BSM physics. In order to take advantage of this wide variety of physics available for study, two experiments were built in the 1990's: Belle and BaBar.

These experiments were referred to as the "B factories," as their goal was to produce B mesons in large numbers. They achieved this by leveraging a very particular experimental setup: an asymmetric electron-positron beam operating at a centre of mass energy of 10.58 GeV. This energy is the mass of an excited $b\bar{b}$ state known as the $\Upsilon(4S)$ resonance[1].

1: Υ being the Greek letter Upsilon, not y from the Latin alphabet.

The $\Upsilon(4S)$ resonance is somewhat of a miracle of nature, providing the perfect laboratory for studying B mesons. It decays almost 100% of the time to a pair of B mesons, either B^+B^- or $B^0\overline{B}^0$. The $\Upsilon(4S)$ mass is only slightly above the mass of two B mesons, and hence they are produced almost at rest in the $\Upsilon(4S)$ frame. Finally, we know that all final state particles of an $e^+e^- \to \Upsilon(4S) \to B\bar{B}$ event originate from one of the two B mesons[2]. This is an extremely useful condition, and it stands in contrast with the much "messier" decays of high energy proton-proton collisions, such as those at the LHC.

2: As we will see in Section 3.3, this is complicated by the presence of *beam background*.

The B factories were designed with these properties of the $\Upsilon(4S)$ in mind. They have two key design aspects: asymmetry in the beam ener-

gies and detector design, and near-total angular coverage. These properties are necessary for the following reasons.

- *Asymmetry.* The B mesons have very little momentum in the $\Upsilon(4S)$ frame. If the $\Upsilon(4S)$ frame was at rest with respect to the lab frame, they would be produced and decay at roughly the same point. Instead, the beams are tuned to different energies so that the $\Upsilon(4S)$ has a forward boost, and the B mesons decay some distance from their point of origin. As a result, more of the final state decay products are emitted in the "forward" direction. To account for this, the B factory detectors are designed asymmetrically, with more detector components in the forward region. The boosted $\Upsilon(4S)$ frame is a prerequisite for measuring time-dependent CP violation in neutral B decays, as the time difference between the decays of a $B^0\overline{B}^0$ pair can only be measured via the spatial separation of their decay vertices.
- *Angular coverage.* Many analyses rely on the knowledge that all non-beam-background final state particles originate from the $\Upsilon(4S)$. This fact is most useful if the detector is able to detect all particles produced. The B factory detector designs achieve this by almost entirely enclosing the solid angle around the $\Upsilon(4S)$ production point[3]. This property is referred to as *hermeticity*.

3: Neutrinos are an exception to this, as they require entirely different experimental setups to detect. Consequently, analyses of decays involving neutrinos are not able to take full advantage of this fact.

The Belle and BaBar experiments collected data from 1999 until 2010 and 2008, respectively. In that time, the collaborations published many landmark results, including the observation of CP violation in previously unmeasured modes, and the discovery of the flavour anomalies[4]. After Belle was shut down, the accelerator and detector underwent major upgrades, which we will now discuss.

4: See Section 2.6.

3.2. SuperKEKB

The SuperKEKB accelerator is an upgrade of the KEKB accelerator used by the Belle experiment [65, 66]. The accelerator consists of two side-by-side storage rings with a 3 km circumference. One ring contains a 7 GeV electron beam and the other a 4 GeV positron beam. These are injected into the rings at full energy by a linear accelerator. The beams have a single crossing point, referred to as the *interaction point* (IP). The Belle II detector is built around the IP. This structure is illustrated in Figure 3.1.

[65]: Abe et al. (2010), *Belle II Technical Design Report*
[66]: Kurokawa et al. (2003), "Overview of the KEKB Accelerators"

In high energy physics, the quantity used to describe amounts of collider data is *luminosity*. For a given process, the number of expected events per second is the process cross-section multiplied by the instantaneous luminosity. Consequently, luminosity has units of inverse area per unit time (commonly $\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$). The *integrated luminosity* is the time

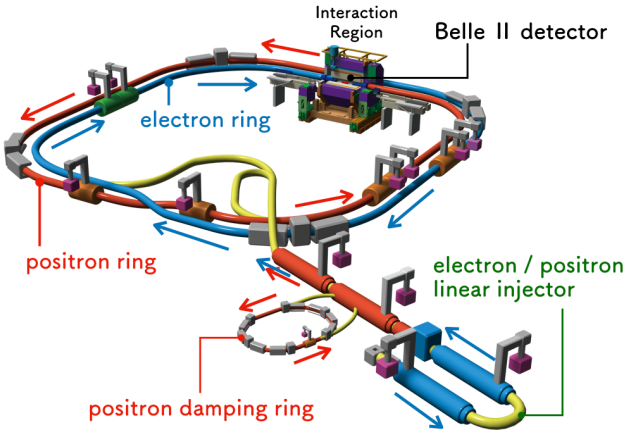


Figure 3.1: Diagram of the SuperKEKB accelerator. The oppositely-travelling electron and positron beams cross at the location of the Belle II detector. Figure adapted from Reference [67].

integral of instantaneous luminosity, and has units of inverse area (commonly fb^{-1} or ab^{-1}).

The primary goal of the SuperKEKB upgrade was to increase the instantaneous luminosity of its predecessor by a factor of forty. To understand how this was achieved, consider Equation (3.1), which describes the luminosity (L) of a collider with two oppositely-travelling beams.

$$L = \frac{\gamma_{\pm}}{2er_e}\left(\frac{I_{\pm}\xi_{\pm}}{\beta^*_{y\pm}}\right)\left(\frac{R_L}{R_{\xi_y}}\right) \tag{3.1}$$

The terms contributing to Equation (3.1) are as follows, where "+" and "−" refer to the positron and electron beams, respectively.

- $\gamma_{\pm}$: the Lorentz factor of the beam,
- e: the electron charge,
- r_e: the classical electron radius,
- $I_{\pm}$: the beam current,
- $\xi_{y\pm}$: the beam-beam parameter,
- $\beta_{y\pm}$: the vertical beta function,
- R_L and R_{ξ_y}: reduction factors of the luminosity due to the crossing angle and hourglass effect; their ratio is close to one.

From the above list, the parameters which can be modified to increase the luminosity are the beam current, the beam-beam parameter and the vertical beta function. SuperKEKB improved the luminosity of KEKB by doubling the current and decreasing the vertical beta function by a factor of twenty. The latter of these was achieved using the "nano-beam" scheme, which minimises the beam overlap at the IP via focusing quadrupole magnets. SuperKEKB also employs the "crab waist" scheme [68], which twists the beams at the IP to further reduce the beam overlap size [69].

[68]: Zobov (2016), "Crab Waist collision scheme"
[69]: Ohnishi et al. (2021), "SuperKEKB Operation Using Crab Waist Collision Scheme"

SuperKEKB surpassed the world record instantaneous luminosity in June 2020, with the record previously held by the PEP-II accelerator at

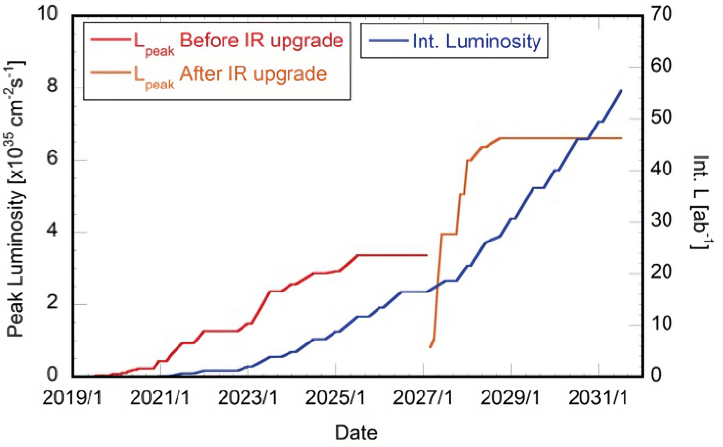


Figure 3.2: Luminosity projection, current as of November 2022 [71]. The two long shutdown periods are in 2022 and 2026. The dip in the instantaneous luminosity at the end of 2026 corresponds with the second long shutdown, in which the interaction region will be upgraded in order for the accelerator to achieve the target luminosity.

[70]: KEK (2022), "SuperKEKB Collider Achieves the World's Highest Luminosity"

SLAC [70]. Its target instantaneous luminosity is $6.5 \times 10^{35}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$—at the time of writing, the highest instantaneous luminosity it has achieved is $4.7 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. The goal of the Belle II collaboration is to record $50\,\mathrm{ab}^{-1}$ by 2030, which is fifty times larger than the full Belle dataset. The luminosity projection is shown in Figure 3.2. Belle II has currently entered its first "long shutdown" period for upgrades, having collected $427.79\,\mathrm{fb}^{-1}$ [71].

[71]: Onuki (2022), "Belle II Status and Prospect"

3.3. Beam background

The upgraded luminosity of SuperKEKB brings with it an increase in background processes arising from the beam operation. This increased background is unavoidable, as the rates of these processes are dependent on the beam size, current, and luminosity. The trigger system[5] plays an important role in suppressing these backgrounds. This Section will introduce the five main categories of beam background.

5: See Section 3.4.6.

3.3.1. Touschek scattering

The Touschek effect is Coulomb scattering of particles within a bunch. It predominantly occurs in the 4 GeV positron beam, as the rate is inversely proportional to the third power of the beam energy. The rate is also inversely proportional to bunch size; because of the nano-beam scheme employed by SuperKEKB, the expected Touschek scattering rate at SuperKEKB is twenty times that of KEKB.

In the scattering, one particle gains energy and the other loses energy, moving them both away from the nominal beam energy. This energy difference causes them to collide with the beam pipe and magnet walls, and produce showers of particles. If these showers occur close to the Belle II detector, then the particles can interact with the subdetector components.

SuperKEKB includes two components to mitigate beam background: collimators and shielding. The collimators are placed at several points around the rings, and capture particles deviating from the main beam trajectory. KEKB only collimated particles in the horizontal direction on the inner side of the ring, while SuperKEKB also collimates on the outer side and in the vertical direction. Furthermore, the innermost detector components are protected from shower particles by heavy metal shielding.

3.3.2. Beam-gas scattering

Beam-gas scattering occurs when beam particles interact with residual gas molecules in the beam pipe. These interactions occur via bremsstrahlung and Coulomb scattering. Similar to Touschek scattering, the beam particles' energies are changed from the nominal energy, and they shower when hitting the beam pipe walls. The two main factors governing the rate of this process are the vacuum pressure and the beam current. There was no change made to the vacuum pressure at SuperKEKB from the KEKB design, but the beam currents are doubled, so a moderate increase in this background is expected.

The systems in place for suppressing Touschek scattering are also effective at suppressing beam-gas scattering processes. The horizontal collimators are particularly effective at reducing bremsstrahlung backgrounds, and the vertical collimators are essential for suppressing the Coulomb scattering background. However, the width and placement of the vertical collimators must be carefully chosen to avoid introducing beam instabilities.

3.3.3. Synchrotron radiation

Synchrotron radiation is the emission of photons from beam particles as they undergo a change in direction. The radiation power is proportional to the square of the beam energy, so the 7 GeV electron beam is the primary source of this background.

Photons from synchrotron radiation can be very damaging to the inner detector components, so it is important to prevent them from reaching the IP. This is achieved by designing the beam optics and IP chamber to avoid direct hits. Additionally, the inside of the beam pipe is gold-coated, and the inner pipe surface has a ridged design to capture radiated photons.

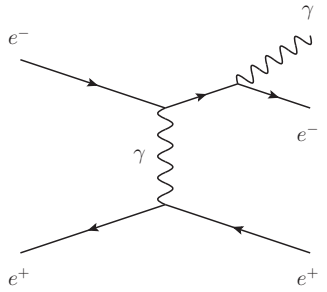


Figure 3.3: Feynman diagram of the radiative Bhabha process.

3.3.4. Radiative Bhabha scattering

In a radiative Bhabha event, an electron and positron scatter off each other and emit a photon (see Figure 3.3). The rate of this process is directly proportional to the luminosity. Photons produced in this way typically travel along the beam pipe and interact with the iron of the magnets, producing a large number of neutrons. These neutrons are the largest background for the K_L^0 and Muon Detector. The beam pipe contains shielding to prevent neutrons from reaching this detector.

Furthermore, the electron and positron may both lose energy but continue to travel in the direction of the beam pipe. In the KEKB design, these outgoing lower-energy particles could be over-bent by the focusing magnets and shower in the magnet walls. SuperKEKB uses a different magnet setup to KEKB, with quadrupole magnets for the incoming and outgoing beams. This design is effective in mitigating the over-bending of beams after energy loss.

3.3.5. Two-photon process

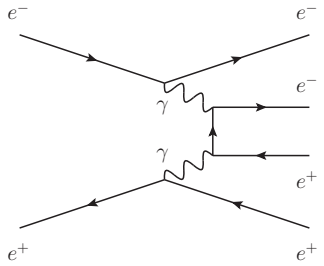


Figure 3.4: Feynman diagram of pair production via the two-photon process.

The final beam background source is electron-positron pair production through the two-photon process $e^+e^- \to e^+e^-e^+e^-$ (see Figure 3.4). The particles produced through this interaction have very low momentum. They then spiral very close to the IP, causing multiple hits in the Pixel Detector. Additionally, the primary beam particles lose momentum and can scatter into the detector, similar to the radiative Bhabha case.

3.4. The Belle II detector

The Belle II detector is an asymmetric particle detector built around the crossing point of the SuperKEKB beams. It shares some of the design of the original Belle detector, and is built in the same location, but every component of it is either upgraded or entirely new. The detection capabilities are improved in several areas: greater vertex resolution due to upgrades to inner pixel detector layers, improved particle identification due to new detector subsystems, and faster electronic readout to handle higher event rates [23, 24, 65].

[23]: Kou et al. (2018), "The Belle II Physics Book"
[24]: Bevan et al. (2014), "The Physics of the B Factories"
[65]: Abe et al. (2010), *Belle II Technical Design Report*

The detector is composed of several subsystems, which are illustrated in Figure 3.5. Its components cover approximately 90% of the solid angle around the IP. A longitudinal magnetic field is imposed on the entire detector, in order to curve the trajectories of charged particles. In this Section, we will discuss each subdetector system, before moving on to the data acquisition and trigger system, and finally the particle identification method.

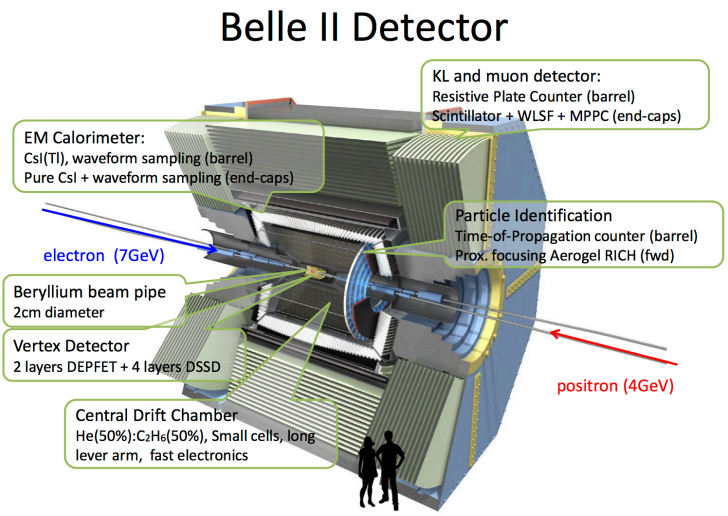


Figure 3.5: Diagram of the Belle II detector, with subdetector components labelled. Figure from https://belle2.jp/detector.

3.4.1. Vertex detectors

The vertex detector subsystem is composed of the silicon Pixel Detector (PXD) and the Silicon Vertex Detector (SVD), arranged in six layers [72, 73]. The first two layers are pixelated sensors, and are arranged at radii of 14 mm and 22 mm (relative to the beam line). This is significantly closer than Belle's innermost vertex layer, which was located at a radius of 88 mm. The remaining four layers are two-layer silicon strip sensors, located at radii of 38 mm, 80 mm, 115 mm, and 140 mm. This extends to a greater radius than in the Belle design. The outer layers were chosen to be strips rather than pixel detectors, in order to reduce the number of output channels. The upgrades in the vertexing system have allowed for significant improvements in vertex resolution and $K_S^0 \to \pi^+\pi^-$ reconstruction efficiency.

[72]: Ye et al. (2021), "Commissioning and Performance of the Belle II Pixel Detector"
[73]: Forti (2022), "The Design, Construction, Operation and Performance of the Belle II Silicon Vertex Detector"

3.4.2. Central Drift Chamber

The Central Drift Chamber (CDC) is a volume drift chamber for measuring the momentum of charged particles [74]. The chamber contains a gaseous mixture of equal parts helium and ethane, and 14,336 sensor wires arranged in 56 layers. These layers are either aligned with or skewed with respect to the longitudinal magnetic field. By combining the information from these two layer alignments, the CDC information can be used to reconstruct full three-dimensional helical tracks. The drift cell size in the Belle II CDC is smaller than that of Belle, which is important for handling the higher event rate. The CDC also extends to a greater radius than in the Belle design.

[74]: Taniguchi (2017), "Central Drift Chamber for Belle-II"

3.4.3. Particle identification systems

The Belle II detector contains two particle identification subdetectors: the time-of-propagation counter (TOP), and the Aerogel Ring Imaging Cherenkov detector (ARICH) [75, 76]. Both subdetectors record Cherenkov radiation, but differ substantially in their mechanism and design. These components completely replaced earlier subdetectors in Belle.

[75]: Fast (2017), "The Belle II Imaging Time-of-Propagation (iTOP) Detector"
[76]: Nishida et al. (2014), "Aerogel RICH for the Belle II Forward PID"

The TOP is located in the barrel region and is composed of 16 quartz bars, each 2.6 m long, 45 cm wide and 2 cm thick. When a particle enters the quartz, it emits a cone of Cherenkov radiation, which internally reflects in the bar. The three-dimensional shape of the cone is reconstructed by using precise measurements of the propagation times of photons. This method relies on precise knowledge of the particle's production time; this is obtained through the synchronisation of the waveform-sampling electronics with the accelerator radio-frequency clock [77].

[77]: Starič (2023), "The TOP Counter and Determination of Bunch-Crossing Time at Belle II"

The ARICH is located in the forward endcap region. It consists of two 2 cm thick layers of aerogel, and an array of photon detectors. Similarly to the TOP, particles emit cones of Cherenkov radiation when passing through the aerogel. However, in the ARICH, these photons are allowed to propagate through an expansion volume and then form a ring on the photon detectors, from which the cone is reconstructed. The choice of expansion volume size is a compromise between two constraints:

1. A sufficient number of photons must be detected in order to reconstruct the cone, which limits the length of the volume.
2. The required resolution of the cone's angle can only be achieved if the photons are allowed to propagate a sufficient distance.

Belle II uses a novel design with two aerogel layers of different refractive indices. This effectively focuses the photons, thus boosting the photon yield without sacrificing angular resolution.

3.4.4. Electromagnetic Calorimeter

The Electronic Calorimeter (ECL) is an array of 8,736 thallium-doped caesium iodide scintillator crystals in the barrel and endcap regions, covering 90% of the solid angle around the IP [78]. The purpose of the ECL is to detect photons and distinguish electrons from hadrons (particularly pions). The individual crystals and support structure are reused from Belle, but the readout electronics have been entirely replaced to handle the higher event rate. Studies are currently in progress to investigate the replacement of the crystals in the forward region with pure caesium iodide crystals, which are more radiation tolerant and provide a faster readout.

[78]: Aulchenko et al. (2015), "Electromagnetic Calorimeter for Belle II"

3.4.5. K_L^0 and Muon Detector

The final subdetector component is the K_L^0 and Muon Detector (KLM), which is composed of 4.7 cm thick iron plates located outside the superconducting solenoid [79]. The iron plates also serve the purpose of magnetic flux return. The KLM design of Belle included resistive plate chambers, which performed well in the environment of KEKB. However, this design was deemed untenable for Belle II, due to the large neutron background from SuperKEKB's increased beam background. Instead, the resistive plate chambers have been replaced by layers of scintillator strips, which have a shorter recovery time after electric discharge.

[79]: Aushev et al. (2015), "A scintillator based endcap K_L^0 and muon detector for the Belle II experiment"

3.4.6. Trigger system

The trigger system is a critical part of the data taking scheme [80]. It is essentially impossible for Belle II to record the entire readout of every event, due to the high event rate and number of output channels. In particular, beam backgrounds make up the vast majority of events, and are significantly more prevalent in SuperKEKB than in KEKB. The role of the trigger system is to identify and record only the events of interest. There are two components to the system: the low level trigger (Level-1, or L1) and the high level trigger (HLT).

[80]: Yamada et al. (2015), "Data Acquisition System for the Belle II Experiment"

The L1 trigger is implemented in the hardware of the Belle II detector, and is based on a set of sub-triggers. The sub-triggers summarise the readouts from the CDC, ECL, KLM and particle identification systems. These summaries use low-level information such as track and cluster multiplicity, total energy deposited, hit topology, and timing. The CDC sub-trigger performs 3D track reconstruction at the hardware level, in order to provide kinematic and vertex information. The summaries are then passed to the Global Decision Logic, which makes the decision on whether to trigger on the event. The total latency of the L1 trigger is 5 µs. It is designed for a maximum trigger rate of 30 kHz, which is sufficient for handling the event rate at the highest target luminosity.

Once the L1 trigger issues a positive decision, the event data is passed to a nearby CPU farm, containing 6000 cores. The HLT is run at this point to reduce the number of events written to disk. The HLT is entirely software-based, and is a part of the same software used for offline reconstruction[6]. In its first step, the HLT performs a fast ECL cluster reconstruction and a fast track reconstruction using only CDC data. It then applies a selection using this information, reducing the event rate to around 15 kHz.

6: That is, basf2. See Section 3.5.4.

In the final step of the HLT, the result of the fast event reconstruction is used to select events based on broad physics criteria. Events are

identified as containing hadrons if they have at least three tracks satisfying basic quality cuts and zero tracks satisfying the Bhabha track selections. This HLT decision is known as `hlt_hadron`, and is used for all B physics analyses. The efficiency of this trigger on $B\overline{B}$ and hadronic continuum events is close to 100%. The physics program of Belle II also includes τ and dark sector analyses, so a separate HLT selection exists to record these types of events. Finally, another HLT decision selects Bhabha and $\mu\mu$ events, with a pre-scale factor on the order of 1%, as these events are used for luminosity measurements and monitoring.

3.4.7. Data acquisition system and data taking

[80]: Yamada et al. (2015), "Data Acquisition System for the Belle II Experiment"

The data acquisition system (DAQ) is responsible for transferring the detector output through several processing steps and ultimately to disk storage [80]. The DAQ builds on the existing Belle design, with upgrades in the digitisation system and the implementation of a unified data link system known as Belle2Link.

The limitations of the DAQ are of primary concern in the trigger design. In particular, the limitation of disk write speed require that the full HLT processing brings the event rate down to 10 kHz. The HLT also identifies interesting regions of the PXD to record, as the full channel readout cannot be feasibly processed or stored.

7: See Section 3.5.5.

In the final step of online processing, the raw detector data is written to storage. The event size in raw data is approximately 300 kB. The raw data is then later reconstructed on the computing grid[7], and saved to the Data Storage Tape (DST) format. DST files contain the raw data as well as higher-level reconstruction objects, but analysts typically only use the latter. For this reason, the DST files are later condensed into mini-DST (mDST) files, which only contain the higher-level objects. The event size in mDST files is approximately 10 kB. mDST files may be further condensed to micro-DST (uDST) files, in a process known as *skimming*, which is covered in Chapter 4.

Data taking at Belle II is divided into *experiments* and *runs*. An experiment is a period of data collection over several weeks in which the accelerator and detector conditions are not expected to change significantly. A run is a shorter period of data collection, lasting at most a few hours. Data taking is divided in this way because the calibration and reconstruction of events depends on accurate monitoring of the operating conditions, which may change between runs. Furthermore, entire runs may be excluded from the analysis dataset if they do not meet certain quality criteria.

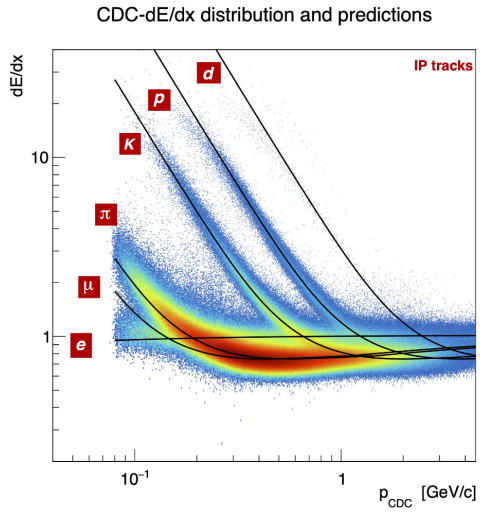


Figure 3.6: Simulated $\mathrm{d}E/\mathrm{d}x - p$ distribution of tracks in the CDC [81]. For momenta below 1 GeV, particles of different species lie within distinct bands.

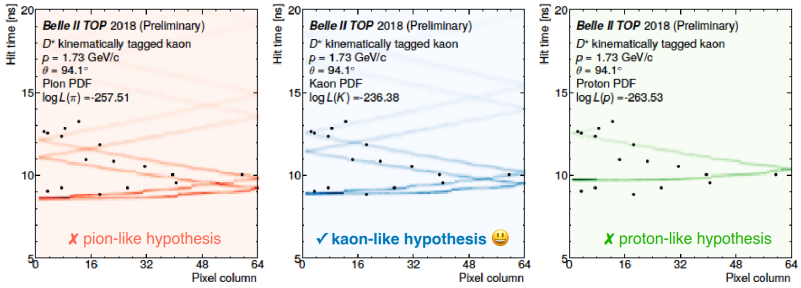


Figure 3.7: Example of fits to TOP readouts under different particle hypotheses [82].

3.4.8. Charged particle identification

Physics analyses rely on accurate identification of the charged particles seen by the detector. The types of charged particles that reach the detector are electrons, muons, kaons, pions, protons, and deuterons. The classification of tracks is referred to as *particle identification* (PID). This Section will briefly outline the PID procedure employed by Belle II.

Each type of particle leaves a distinct signature in the Belle II subdetector components, which can be combined to classify a given track. The rate of energy loss over distance ($\mathrm{d}E/\mathrm{d}x$) of a particle in the CDC and SVD is directly proportional to the ratio of its momentum and mass. When plotting $\mathrm{d}E/\mathrm{d}x$ against momentum, each particle type traces out distinct bands; this distribution is shown in Figure 3.6. The placement of a track in the $\mathrm{d}E/\mathrm{d}x - p$ plane is a main source of PID information for hadrons with momenta below about 1 GeV. The main sources of information for electron and muon PID are the ECL and KLM, respectively. Finally, the new ARICH and TOP subdetectors are instrumental in ensuring the PID capabilities of Belle II cover a wider kinematic range; Figure 3.7 shows example fits to a TOP readout under different particle hypotheses.

The goal of PID is to provide a probability that a given particle hypothesis is true. To construct this probability, we first take a single hy-

pothesis, h, and calculate a log likelihood from each subdetector, given that hypothesis and the pattern of hits in that subdetector. We then sum each subdetector log likelihood, to produce the log likelihood of that hypothesis, $\mathscr{L}_h$. The probability of h being the correct identification is given by this log likelihood normalised to the log likelihoods of all other hypotheses (see Equation (3.2)). In analyses, these probability variables are often referred to by their variable names in code, *e.g.* `pionID`, `muonID`.

$$P(h) = \frac{\log \mathscr{L}_h}{\log \mathscr{L}_e + \log \mathscr{L}_\mu + \log \mathscr{L}_\pi + \log \mathscr{L}_K + \log \mathscr{L}_p + \log \mathscr{L}_d} \tag{3.2}$$

3.5. Software

Having described the hardware aspects of the Belle II experiment, we now turn our attention to the software. The software ecosystem of Belle II encompasses a wide range of data processing and analysis tasks. This Section will explain the role of software in simulation and reconstruction, and introduce the software framework basf2.

3.5.1. Simulation

The goal of any HEP experiment is to compare the collected data to a corresponding theoretical prediction. To perform this comparison, experimental collaborations create Monte Carlo (MC) samples. In MC, we generate decays using our best understanding of the decay rates and kinematics of each particle. We then model the response of the detector as the final state particles pass through it. This means that the MC contains a "true" record of the origin of each object. Statistical inference in HEP is often performed by fitting data distributions to MC.

[83]: Torben Ferber et al. (2015), *Overview of the Belle II Physics Generators*

Belle II uses a specialised software stack to simulate specific processes [83]. EvtGen is the most relevant generator for B physics analyses. It simulates decays of B mesons and their daughters according to a table listing the branching fractions and theoretical parameters of each decay mode. To name just a few of the other primary generators that are used to simulate specific processes: Babayaga.Nlo for Bhabha scattering, Kkmc for continuum events, and MadGraph for simulating BSM processes. Several secondary generators are used in conjunction with the others: Tauola is specialised for simulating τ decays, Photos simulates final state electromagnetic radiation emitted from outgoing particles, and Pythia is used to model quark fragmentation and inclusive decays. The final state particles are passed through a complete model of the detector in Geant4. The response of the simulated detector is digitised to correspond to the read out of the electronic components of the actual detector.

Finally, beam background is simulated separately and later added to the MC in the form of "overlay files."

Belle II MC is generally categorised as either "generic" or "signal." The definition and purpose of each is as follows.

- *Generic MC*. In generic MC samples, a specific outcome of the primary e^+e^- interaction is simulated, but those decay products are then allowed to decay in all possible modes. The four main categories of generic samples are: $\Upsilon(4S) \to B^+B^-$ (denoted "`charged`"); $\Upsilon(4S) \to B^0\overline{B}^0$ including neutral B mixing (denoted "`mixed`"); continuum $e^+e^- \to q\bar{q}$ and $e^+e^- \to \tau^+\tau^-$; and other low-multiplicity continuum processes such as $e^+e^- \to \mu^+\mu^+\mu^-\mu^-$. Each analysis may require a different set of generic samples, depending on the kinds of backgrounds expected to be present after all selections are made.
- *Signal MC*. Signal MC samples simulate a small number of specific decay processes; for example $\Upsilon(4S) \to B^+B^-$, with one of the B mesons required to decay via $B^+ \to \pi^0\ell^+\nu_\ell$. Signal samples are important for studies of rare decays, as they allow the analyst to substitute the small number of signal decays in the generic sample with a large number of decays of that type. Doing so reduces the statistical uncertainty in the shape of the signal component.

It is important to note that the "real" and simulated data are both in the same mDST file format, despite their differing origins. When an analyst performs their reconstruction, they process the two types of data using the same code. The only difference between the two is that the MC data contains additional information; namely, the full simulated decay chain, and the relations between the detector hits and the particles that caused them. This information is referred to as "MC truth information." We will now examine how the truth information can be related to the reconstructed particle information.

3.5.2. MC truth matching

The procedure of relating generator-level information to the analyst's reconstruction is known as *MC-matching* or *truth-matching*. MC-matching is a powerful tool for optimising selections and understanding the types of backgrounds that can arise from misreconstruction. Belle II has developed an algorithm for recursively truth-matching entire decay chains. The two goals of the algorithm are (1) the creation of relations objects[8] between the reconstructed `Particle` objects and the `MCParticle` objects, and (2) the evaluation of those relations to categorise failures of the reconstruction [84].

8: See Section 3.5.4.

[84]: Sato et al. (2021), "Monte Carlo Matching in the Belle II Software"

The creation of MC relations proceeds in several steps, according to the type of particle being truth-matched. Final-state particles are assigned their match based on whether they are reconstructed from the hits/clusters produced by a single generated particle. A track is considered a match if 66% of its hits are from one generated particle, and the reconstruction includes at least 5% of the hits of that generated particle [85]. For each cluster, a weighted sum of the ECL crystal energy deposits, w (with units GeV^2), is calculated; a cluster is considered a match if this $w/E_{\text{reconstructed}} > 0.2\,\mathrm{GeV}$ and $w/E_{\text{generated}} > 0.3\,\mathrm{GeV}$ [86]. If there is no appropriate match, then no `MCParticle` is assigned. Composite particles are matched by first matching all daughter particles, and calling the algorithm recursively on composite daughters until the entire decay chain is matched. If any daughter is not truth-matched, then no `MCParticle` is assigned. If all daughters are matched, then the `MCParticle` assigned is the nearest common mother of all particles used in the reconstruction. In cases where daughter particles from one B are mis-assigned to the other B, the nearest common mother of that B candidate will be the $\Upsilon(4S)$.

[85]: Bertacchi et al. (2021), "Track Finding at Belle II"

[86]: Belle II Collaboration (2022), *Belle II Analysis Software Framework (Basf2)*

In the second step of truth matching, the algorithm evaluates the MC relations and assigns error flags for each category of misreconstruction. These scenarios may involve a particle that was missed in the reconstruction, or the erroneous inclusion of a particle from some other part of the event. Alternatively, a final state candidate could be misidentified and erroneously used in later reconstruction. For example, a track produced by a K^+ can misidentified as a π^+ track, and that fake K^+ might then be used to reconstruct a D^0 candidate. As another alternative, a particle may have no `MCParticle` match, in which case an "internal error" flag is assigned. Throughout this step, the error flags of daughter particles are propagated to the reconstructed mothers. The outcome of this process is the integer variable `mcErrors`, which contains all the error flags encoded in its bits.

The error flags can also be used to create boolean variables which indicate whether a candidate was correctly reconstructed. The variable `isSignal` is one if the candidate has no error flags, and zero otherwise. This criterion is too harsh for particular use cases, but the error flag scheme allows for individual conditions to be loosened. For example, consider a reconstructed semileptonic decay, which will always be missing a neutrino. The variable `isSignalAcceptMissingNeutrino` loosens one error flag to allow for missing massless particles, making this a more appropriate variable for evaluating the correctness of the reconstruction.

3.5.3. Event weights

Every event in data and MC has an associated "event weight," which governs how much that event contributes to histograms and event sums. For

detector data, all event weights are one, as these events are ultimately the basis of any measurement. MC event weights are commonly modified to achieve some purpose, in a process known as "reweighting." In the main analysis of this thesis, event weights are modified for three purposes, which we will outline here.

The first purpose is normalisation of MC samples to match the integrated luminosity of the detector dataset. For example, in 100 fb^{-1} of data, we expect to see 54M $\Upsilon(4S) \to B^+B^-$ events, assuming an $e^+e^- \to \Upsilon(4S)$ cross-section of 1.1 nb. Using these values, we can then scale all event weights of any `charged` sample to match the integrated luminosity of the data. A similar calculation can be done with signal samples, using the expected branching fraction of the decay.

The second purpose of reweighting is to update MC to incorporate new theoretical understanding which was not available or implemented at the time of production. Branching fractions and form factor parameters can be updated in this way. Generating MC is very expensive (particularly the detector simulation), so it is unreasonable to regenerate it every time a new theory result is published. Instead, tools can be written to compare the generator-level information to the new theory predictions, and modify the event weights accordingly. Furthermore, the analyst is able to derive sets of weight variations by varying the theoretical parameters within the range of uncertainty. These variations allow them to estimate the impact of theoretical uncertainty on their final measurement.

The final purpose of reweighting is correct mismodelling of the detector response. The detector simulation includes inherent assumptions regarding the behaviour of each type of particle. These assumptions must then be tested against the actual response in each detector region—this is the domain of "performance studies." For example, PID classification variables are functions of the signature left in the detector by a given particle. However, given the differences between the real and simulated detector, a cut on a PID variable may perform differently on data and MC. The PID working group provides tables which correct the differences in efficiency and fake-rate arising from such a selection. Analysts then apply these correction factors to the MC event weights, in order to calibrate the MC to be more true to the data. Each correction factor has an uncertainty attached, and we can again propagate this uncertainty to the final measurement via a set of weight variations.

3.5.4. The Belle II Analysis Software Framework

All simulation and reconstruction in Belle II is performed using The Belle II Analysis Software Framework (basf2, pronounced *bass-eff-two*)[87]. This framework has been developed over more than ten years, improving on

[87]: Kuhr et al. (2019), "The Belle II Core Software"

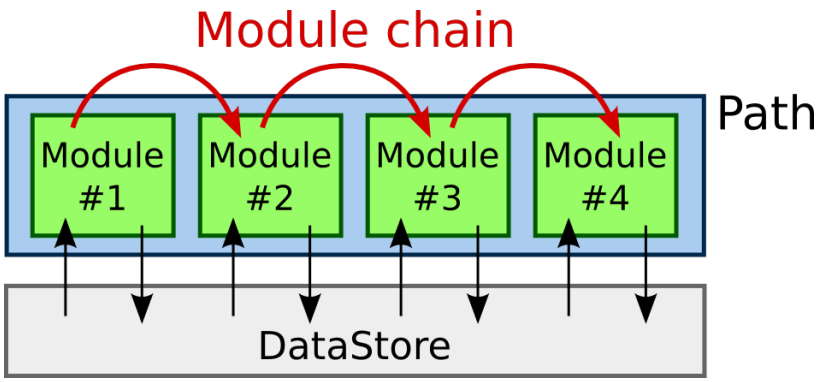


Figure 3.8: Diagram of data processing during a basf2 run. For each event, the path runs a sequence of modules. All modules may read from and write to the Data Store. Figure adapted from [88].

the software framework of Belle. Here we will introduce the framework, and explain how it is used for particle reconstruction.

basf2 achieves the dual goals of high performance and user accessibility by having a core written in C++ that is controlled by the user from Python. Python is an accessible scripting language with an extensive library of third-party packages. The decision to steer basf2 using Python allows users with very little programming experience to write performant code. Furthermore, basf2 can be run in Jupyter notebooks, allowing for interactive analysis development. This combination of C++ and Python is appropriate for a collaboration consisting of a small number of software experts and a large number of physics analysts of widely differing software ability.

A "steering file" is a Python script in which the user defines the event processing in a high level manner. This contrasts to the traditional HEP approach, where analysts write their own event loop, and directly access and manipulate data objects. The two fundamental concepts to steering files are *paths* and *modules*. Modules are objects which perform some particular processing task for each event, for instance, event generation or decay reconstruction. A path is an object which stores an ordered sequence of modules to be run. This relationship is illustrated in Figure 3.8.

Code 1 shows a basic example of a steering file. We first create a path, then add modules to the path, and then process the path. This call to `basf2.process` is what starts the event loop. All of the event processing is handled internally in the modules; the only thing the analyst needs to do is configure them and add them to the path in the desired order. basf2 contains modules for reading and writing to ROOT files, reconstructing decays, applying kinematic fits, and writing histograms, to name just a few of the many available modules. The Python module `modularAnalysis` contains convenience functions for including common modules in steering files.

Modules have five methods that are called at different points in processing: methods called at the start and end of all processing (for initialisation and completion), methods called at the start and end of each run

```python
# Import core objects and methods
import basf2 as b2
# Import python module containing convenience functions
import modularAnalysis as ma

# Instantiate path object
path = b2.Path()

# Read mDST, and write variables to flat ntuple
ma.inputMdst("some_file.mdst.root", path=path)
ma.variablesToNtuple(
    decayString="",  # write event-level ntuple
    variables=["nTracks"],
    filename="ntuples.root",
    path=path,
)

# Start event processing
b2.process(path)
```

Code 1: Example of a basic steering file. First a path object is created, then modules are added to the path, and then the path is processed. No event processing is performed until `basf2.process` is called.

(*e.g.* to produce run summary histograms), and a method called for each event. Most modules have optional parameters, which configure how a module is run. For example, the parameters of the vertex fitter module specify the vertices to be fitted, and the constraints to be applied to the fit. Most modules are written in C++, but modules can also be written in Python; this option is provided for fast prototyping, and for cases where performance is not critical.

During event processing, modules pass data between each other by accessing the "Data Store," as shown in Figure 3.8. All modules have access to the Data Store, and are able to create and modify objects stored there. The design of the Data Store consists of objects and lists of objects, and "relations" objects which can express arbitrary relationships between objects. Relations are utilised to keep track of which particles are used to reconstruct intermediate particles, and are also a key aspect of the MC truth matching procedure detailed in Section 3.5.2. If one side of a relation is removed (*e.g.* by applying an analysis cut), then the entire relation is removed—this allows objects to be safely dropped without compromising data integrity.

basf2 stores information about particles as arrays of `Particle` objects in the Data Store, referred to as "particle lists." In a steering file, these are accessed by a string identifier, formatted as the particle name, a colon, and a user-supplied label (*e.g.* `"pi+:tight"`). This allows the user to manage multiple particle lists of the same particle type, with different selections on each. Whenever a particle list is accessed by name, the charge conjugate list is also accessed. It is important to note that the analyst never directly manipulates the particle list object, but instead passes its string identifier to modules.

Code 2: Example of decay reconstruction in basf2. Here, all photons of energy greater than 0.1 GeV are used to reconstruct a neutral pion, to which a mass cut is applied.

```
ma.fillParticleList(
    "gamma:loose",
    cut="E > 0.1",
    path=path,
)
ma.reconstructDecay(
    "pi0:analysis -> gamma:loose gamma:loose",
    cut="0.120 < M < 0.145",
    path=path,
)
```

All particle reconstruction in basf2 is performed by passing "decay strings" to modules. Decay strings take the form of `"MOTHER -> CHILD [CHILD ...]"`. A "cut string" may be used to apply some selection to each reconstructed candidate. Code 2 illustrates a simple $\pi^0 \to \gamma\gamma$ reconstruction, with cuts applied to both the photons and the reconstructed π^0. When combining multiple particles of the same type, the reconstruction modules prevent the same particle being used twice in the same candidate. Furthermore, if the same physical track or cluster exists in several particle lists, the modules again ensure it is not used twice in the same candidate.

The final part of the basf2 design is the Conditions Database (CDB). The CDB exists to handle settings and calibrations that may change between runs. Data is uploaded to the CDB in payloads, which each contain metadata listing the range/s of runs to which the payload applies. There is no restriction on the types of objects that can be uploaded as payloads. There are two uses of the CDB that are particularly relevant to this thesis. The first is the magnetic field map, which is frequently recorded and uploaded to the CDB—this map is crucial for valid PID functioning. Secondly, as we will see in Section 3.6.3, the weights of trained classifiers are made available to the collaboration by uploading them to the CDB.

3.5.5. Grid processing

The Belle II experiment relies on a global computing grid to process large amounts of data. The grid is utilised for raw data processing and MC generation (performed by the data production group during a "campaign"), and physics analyses (performed by individual analysts). Each of these tasks is performed using basf2 steering files. The computing group developed the gbasf2 program, which submits steering files to the grid [89]. The required inputs to gbasf2 are a steering file and an input dataset to be processed. We will briefly cover the aspects of these two inputs that are relevant to Chapter 4 and the analysis.

[89]: Miyake et al. (2015), "Belle II Production System"

The key benefit of gbasf2 is that a steering file can be tested locally and then run on the grid without any modifications. This minimises the

onboarding required for non-expert grid users. A necessary restriction is that gbasf2 *only* accepts basf2 steering files, and does not allow for arbitrary scripts to be run on the grid. With this restriction in place, gbasf2 takes full responsibility for the environment setup on each job site.

Dataset locations on the grid are referred to by their Logical Path Name (LPN). This is a directory-like string which maps to a dataset in the metadata catalogue. When processing an LPN containing multiple files, gbasf2 will submit individual jobs for each file, and ensure that the output filenames are unique. In addition to the gbasf2 command, the computing group has written a suite of tools for querying and managing data storage on the grid. These aspects of gbasf2 greatly simplify data handling for the user.

3.6. B physics analysis techniques

In the previous Sections, we examined the hardware used to produce and collect data, and the software for simulating data and reconstructing events. We will now turn our attention to specific aspects of B physics analyses which are relevant to the upcoming Chapters. First, we will discuss the technique known as *tagging*, which improves the sensitivity of analyses by taking advantage of the unique experimental setup of the B factories. We will then define several variables which are useful for analyses of semileptonic B decays, and explain the properties of these variables. Finally, we will introduce the Full Event Interpretation, which is a novel tagging algorithm developed by the Belle II collaboration.

3.6.1. Tagging

As discussed in Section 3.1, the Belle II experimental setup has several useful properties; these properties allow for a technique known as *tagging*. The centre of mass frame (CMS) of the two beams is precisely known, and beam energies are tuned so that this is the rest frame of the $\Upsilon(4S)$. In an $\Upsilon(4S)$ event, every final-state particle seen by the detector originates from one of the two B mesons (barring beam background). Finally, the near-hermeticity of Belle II ensures that all particles in the event are detected (modulo detector efficiency). The key idea behind tagging is that if we can reconstruct one B, then all the other particles in the event must originate from the other B.

In practice, we refer to the first B as the "tag side B" (B_{tag}), and other as the "signal side B" (B_{sig}). B_{tag} is reconstructed in one or more "tag modes," which are typically chosen for their high branching fraction and cleanness of reconstruction. The signal B is then reconstructed in whichever decay mode we are interested in studying. This is illustrated

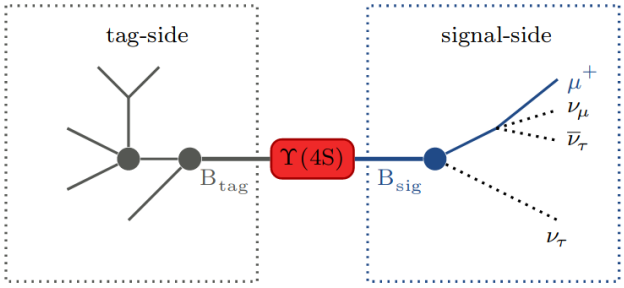


Figure 3.9: Illustration of tag and signal B reconstruction. Figure adapted from Reference [64].

in Figure 3.9. It is important to note that the decay products of the two B mesons overlap in the detector, so we cannot rely on their spatial separation (despite what is shown in the illustration). With the full $\Upsilon(4S)$ decay chain reconstructed, we may then use conservation laws to place strong constraints on the signal B. This is particularly useful for studying signal modes with missing energy due to neutrinos or possible dark matter particles.

Tagging efficiency is an important quantity when discussing tagged analyses. It is defined as

$$\text{tagging efficiency} = \frac{\text{number of events with a tag}}{\text{total number of events}} . \tag{3.3}$$

For the purpose of this discussion, we will also define the *purity* as

$$\text{purity} = \frac{\text{number of correctly reconstructed events}}{\text{total number of events}} . \tag{3.4}$$

Note that the tagging efficiency does not account for the correctness of the B_{tag} reconstruction. There is a trade-off between increasing the tagging efficiency and increasing the purity, and this decision is highly analysis-dependent. It is important for an analysis to have a sufficiently high tagging efficiency, because this maximises the number of opportunities to reconstruct a B_{sig}. The tagging efficiency can be increased by loosening the requirements on B_{tag}, and by reconstructing B_{tag} in a larger number of modes.

There are three broad categories of tagged approaches in B physics analyses: hadronic tagging, semileptonic tagging, and inclusive tagging.

- In hadronic tagging, the decay products of the tag modes are all hadrons. For example, the mode $B^- \to D^0\pi^-$, where the D^0 is reconstructed from a kaon and a pion. All final state particles of B_{tag} are seen by the detector, so its 4-momentum can be precisely determined. Using conservation of momentum, we can then infer the momentum of B_{sig}. This knowledge enables very useful selections on the B_{sig}, such as angular variables in the B_{sig} frame. The precise reconstruction of the B_{tag} momentum also allows for the use of a pair of variables which are highly effective for discriminating

background: the energy difference from the beam energy ΔE, and the beam-constrained mass M_{bc},

$$\Delta E = E_B^* - E_{\text{beam}}^* \tag{3.5}$$

$$M_{\text{bc}} = \sqrt{\left(E_{\text{beam}}^*\right)^2 - \left(\boldsymbol{p}_B^*\right)^2}, \tag{3.6}$$

where the superscript asterisks indicate variables calculated in the CMS frame, and $E_{\text{beam}}^* = \sqrt{s}/2$.

The drawback to hadronic tagging is its low tagging efficiency. B mesons have a large number of hadronic decay modes, each with very small branching fractions. The modes suitable for use in tagging each have branching fractions of at most around 0.1%. To address this issue, hadronic tagging requires that a large number of tag modes are reconstructed.

- Semileptonic tagging instead uses B decays which have a hadronic part and a leptonic part. For example, $B^- \to D^0 e^- \bar{\nu}_e$, where the D^0 is reconstructed from a kaon and a pion. Semileptonic tagging produces a higher tagging efficiency than hadronic tagging. The decays $\overline{B}^0 \to D^{*+} e^- \bar{\nu}_e$ and $\overline{B}^0 \to D^{*+} \mu^- \bar{\nu}_\mu$ alone make up over 10% of the total $\overline{B}^0$ branching fraction (the same is true for similar decays of the B^-). Thus, with just a small number of tagging modes, semileptonic tagging can achieve a large tagging efficiency.
  However, the final state of the tag-side contains a neutrino, which does not interact with the detector. This missing 4-momentum means that we cannot make use of M_{bc} and ΔE selections as effectively as in hadronic tagging. Consequently, this approach has a lower purity. However, there are still methods of using the kinematics of the semileptonic-tagged B to constrain the signal B, as we will see in Section 3.6.2.
- Inclusive tagging is the simplest tagging method with regard to reconstruction. We first reconstruct B_{sig}, and then combine all remaining final state particles into a B_{tag} candidate. That is, no individual B_{tag} decay chains are reconstructed, and all particles not on the signal side are assumed to have come from the tag side. This approach is not reliant on selecting high branching fraction modes, and so has the highest tagging efficiency of all three methods. However, this gives us the least amount of information for constraining the signal side, making this the lowest purity approach.

These three complementary approaches each offer benefits and drawbacks. The general trend is a trade-off between tagging efficiency and purity, as illustrated in Figure 3.10. In Section 3.6.3, we will discuss an important algorithm used in Belle II for reconstructing hadronic and semileptonic tag modes.

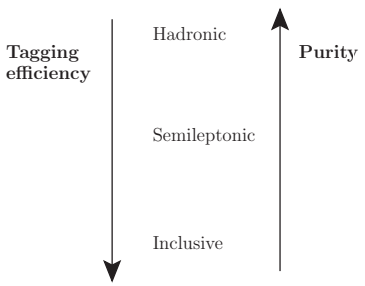


Figure 3.10: The trade-off of tagging efficiency and purity between different tagging approaches.

3.6.2. Discriminating variables for semileptonic analyses

We will now discuss several variables which are useful in analyses of semileptonic decays. These variables take advantage of our knowledge of the experimental setup, and of the particular decays we are reconstructing. In the main analysis of this thesis, we use these variables to discriminate signal and background candidates. In this Section, we will look at their definitions and properties.

Second Fox-Wolfram moment

[90]: Fox et al. (1978), "Observables for the Analysis of Event Shapes in e^+e^- Annihilation and Other Processes"

Fox-Wolfram moments are quantities which describe the "shape" of an event [90]. They are particularly useful for identifying continuum events, which contain a $q\bar{q}$ or $\tau^+\tau^-$ pair, rather than a $B\bar{B}$ pair. The ℓ^{th} moment is defined as

$$R_\ell = \frac{1}{s}\sum_{i,j}|\boldsymbol{p}_i||\boldsymbol{p}_j|P_\ell(\cos\theta_{ij})\,, \tag{3.7}$$

where P_ℓ is the ℓ^{th} Legendre polynomial, $\boldsymbol{p}_i$ is the 3-momentum of the i^{th} particle, θ_{ij} is the angle between two particles' 3-momenta, s is the beam centre-of-mass energy, and the indices i and j range over all tracks and clusters in the event. Note that R_ℓ is calculated in the CMS frame, and has no preferred axis. From energy and momentum conservation, Equation (3.7) can be used to show that $R_0 = 1$ and $R_1 = 0$. The factor of $1/s$ is in place so that all other R_ℓ are in the range $[0, 1]$.

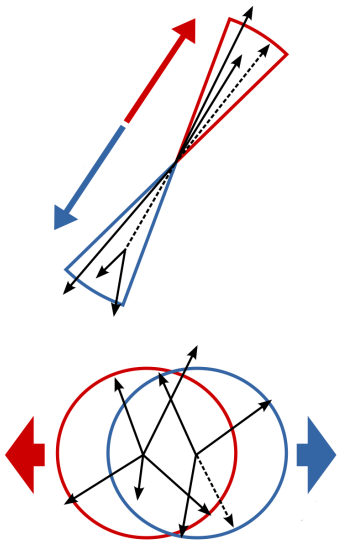


Figure 3.11: Difference in distribution of decay products in $q\bar{q}$ events (upper diagram) and $B\bar{B}$ events (lower diagram). $q\bar{q}$ events tend to be jet-like, while $B\bar{B}$ events are more spherical. This shape difference is quantified in the variable R_2. Figure adapted from documentation of basf2 [86].

The second Fox-Wolfram moment, R_2, describes how jet-like the event is. This is a useful property for identifying continuum events. In a $B\bar{B}$ event, each B has a very low momentum in the CMS frame. Consequently, the decay products have no preferred direction, so $B\bar{B}$ events are typically "spherical". In contrast, the decay products of $q\bar{q}$ pairs have very large momenta, so the final-state particles will tend to be distributed in two oppositely-pointing jets. This is illustrated in Figure 3.11. R_2 is defined so that $B\bar{B}$ events have low values ($\lesssim 0.4$), and $q\bar{q}$ events have values closer to one.

Angular variables

When we reconstruct decays involving a missing neutrino, we are limited in the constraints that we can apply to the B candidate. The presence of missing 4-momentum means that we cannot determine the precise flight direction of the B. Despite this restriction, we are able to construct discriminating angular variables, on the assumption that there is only

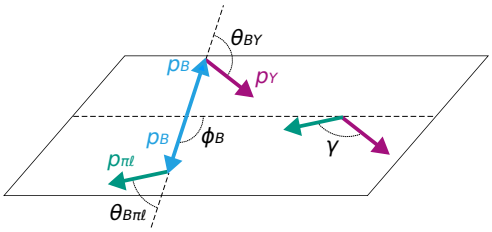


Figure 3.12: Illustration of definition of $\cos^2 \Phi_B$. The reconstructed momenta p_Y and $\mathrm{p}_{\pi\ell}$ define a plane, and the angle $\cos \Phi_B$ is the cosine of the angle between the $\mathscr{B}$ momenta and that plane.

one source of missing momentum in the decay chain of the B. We will discuss two variables which are relevant to the upcoming thesis.

The first variable is $\cos \theta_{BY}$, which is the angle between the momentum of the nominal B and that of the reconstructed Y system[9]. Using the fact that the neutrino is massless, and under the assumption that there is only one neutrino in the decay chain, we can express this angle entirely in terms of known quantities.

$$\cos \theta_{BY} = \frac{2E_B^* E_Y^* - m_B^2 - m_Y^2}{2|\boldsymbol{p}_B^*||\boldsymbol{p}_Y^*|} \tag{3.8}$$

In Equation (3.8), E_B^* and E_Y^* are the energies of the B meson and Y system, $|\boldsymbol{p}_B^*|$ and $|\boldsymbol{p}_Y^*|$ are magnitudes of their 3-momenta, m_Y is the invariant mass of the Y system, and m_B is the mass of the B meson. The superscript asterisks indicate variables calculated in the CMS frame. m_B is a constant, E_B^* and $|\boldsymbol{p}_B^*|$ can be calculated from energy-momentum conservation, and E_Y^* and $|\boldsymbol{p}_Y^*|$ are variables of the reconstructed visible system.

The equality in Equation (3.8) only holds true when the candidate is correctly reconstructed. This gives the variable the useful property that correctly reconstructed candidates have values in the interval $[-1, 1]$. In practice, this window is loosened slightly to account for detector resolution and reconstruction effects. If we calculate the right-hand side using a misreconstructed candidate, the result does not correspond to a cosine value, and so is more broadly distributed. This gives us a powerful discriminant between correctly reconstructed and misreconstructed candidates for semileptonic decays.

The second angular variable relevant to this thesis is used in the case where B_{tag} and B_{sig} are *both* reconstructed in semileptonic modes. For the sake of this example, we will denote the reconstructed tag system as Y, and the signal system as $\pi\ell$. In the CMS frame, the two B mesons decay back-to-back, and their flight direction defines an axis. Using the 3-momenta of the Y and $\pi\ell$ systems, we can define a plane, as shown in Figure 3.12. The Φ_B is defined as the angle between the B axis and this plane [91]. $\cos^2 \Phi_B$ can be expressed in terms of $\cos \theta_{BY}$ of each B and the angle γ between the Y and $\pi\ell$ 3-momenta. This is shown in Equation (3.9).

9: If we are interested in the decay $B \to D^* \ell \nu$, then Y denotes the $D^* \ell$ system.

[91]: BaBar Collaboration (2006), "Measurement of the $B \to \pi \ell \nu$ Branching Fraction and Determination of $|V_{ub}|$ with Tagged B Mesons"

$$\cos^2 \Phi_B = \frac{\cos^2 \theta_{BY,\text{sig}} + \cos^2 \theta_{BY,\text{tag}} + 2 \cos \theta_{BY,\text{tag}} \cos \theta_{BY,\text{sig}} \cos \gamma}{\sin^2 \gamma} \quad (3.9)$$

Similar to the $\cos \theta_{BY}$ case, the equality in Equation (3.9) only holds for correctly reconstructed $\Upsilon(4S)$ candidates. From this we can conclude that correctly reconstructed candidates will have $\cos^2 \Phi_B$ values in the interval $[0, 1]$. Again, the right-hand side can be calculated for misreconstructed candidates, and the resulting values take a much broader range.

These two variables are not independent, as the former is an input to the latter. This means that making a stringent selection on $\cos \theta_{BY}$ will lessen the usefulness of $\cos^2 \Phi_B$. However, loose selections on $\cos \theta_{BY}$ can still be made to remove low quality B candidates without biasing $\cos^2 \Phi_B$.

$\cos^2 \Phi_B$ was used in $B \to \pi \ell \nu$ analyses by BaBar [91], and the closely related variable x_B^2 was used by Belle [92].

[91]: BaBar Collaboration (2006), "Measurement of the $B \to \pi \ell \nu$ Branching Fraction and Determination of $|V_{ub}|$ with Tagged B Mesons"

[92]: Belle Collaboration (2007), "Measurements of branching fractions and q^2 distributions for $B \to \pi \ell \nu$ and $B \to \rho \ell \nu$ Decays with $B \to D^{(*)} \ell \nu$ Decay Tagging"

Rest-of-event variables

Once we have reconstructed an $\Upsilon(4S)$ candidate, all other objects not used during the reconstruction are referred to as the *rest-of-event* (ROE). Because of the unique experimental setup of the B factories, backgrounds can be greatly reduced by making selections on variables calculated from the ROE. The two variables relevant to this thesis are:

- $n_{\text{tracks,ROE}}$: the number of charged tracks in the ROE. If we have reconstructed two B mesons correctly, then we should expect no remaining tracks. A cut of $n_{\text{tracks,ROE}} = 0$ removes any candidate which does not account for all tracks in the event.
- $E_{\text{extra,ROE}}$: the sum of energy deposits in the ECL which are not used in either B reconstruction. Signal events should peak around zero in this variable. Background events tend to have higher values, due to unassigned neutral clusters.

In practice, dealing with ROE variables is complicated by the presence of beam background and detector resolution effects. These give rise to additional tracks and photons not originating from the $\Upsilon(4S)$, as well as spurious energy deposits in the ECL. The approach used in Belle II analyses is to define an *ROE mask*, which is a set of cuts to determine which tracks and clusters should be considered as part of the ROE. For tracks, a typical mask includes loose cuts on the vertex parameters and momentum, to remove tracks likely due to beam background. For clusters, a typical mask may include a lower bound on the cluster energy, the value of which is different for the barrel and endcap regions. The effect of the mask is to remove certain ROE particles from consideration, and

bring the $n_{\text{tracks,ROE}}$ and $E_{\text{extra,ROE}}$ distributions more in line with the above outlined expectations.

3.6.3. The Full Event Interpretation

Analyses using exclusive tagged approaches suffer from low reconstruction efficiencies, as explained in Section 3.6.1. This limitation can be addressed by reconstructing a large number of exclusive tag modes, thus increasing the opportunities to reconstruct a signal-side B. To this end, the Belle II collaboration developed the Full Event Interpretation (FEI). The FEI is an algorithm which is able to reconstruct $\mathcal{O}(10,000)$ exclusive decay modes. It also provides a final output variable indicating the probability of the reconstruction being correct. The FEI is an important tool for Belle II analyses, and forms the basis of the main analysis in this thesis.

The FEI is a successor to the Full Reconstruction (FR) algorithm developed by Belle. Its design is largely inspired by the FR, but improves on it in several regards. The FR was based on NeuroBayes [93], whereas the FEI is implemented using the FastBDT package [94]. The FEI was designed with user customisation and flexibility in mind, and as such, provides many configuration options for customising the algorithm setup at a granular level. The FEI also implements far more modes than the FR—in particular, the FEI implemented semileptonic tagging, whereas the FR only performed hadronic tagging. Finally, training the FR algorithm typically took several weeks, whereas the FEI can be trained in days.

[93]: Feindt et al. (2006), "The NeuroBayes Neural Network Package"
[94]: Keck (2016), *FastBDT*

The building blocks of the FEI are *multivariate classifiers*. These are functions which take multiple input values and map them to either a finite set of categories or a continuous range of values. The individual classifiers in the FEI are Boosted Decision Trees. Their purpose is to take several variables from a given particle, and return a value between zero and one that the particle is correctly reconstructed. The output value can be interpreted as a probability, and is referred to as the "signal probability."

When the FEI analyses a particular decay, it executes the following steps.

1. *Loading/reconstructing candidates.* The first step depends on whether the classifier is for a final-state or intermediate particle. The former is loaded from the detector object, and the latter is reconstructed in the given mode. The reconstruction may also include a vertex fit, to provide additional information about the candidate for the classification step.
2. *Pre-classifier cuts.* A set of loose pre-classifier cuts are then applied to the candidates. These are basic quality selections to remove candidates which are unlikely to have been correctly reconstructed.

For example, each charged tracks is required to have the closest-point-of-approach to the IP within 2 cm radially and 4 cm longitudinally; and neutral clusters must have an energy of at least 100 MeV, 90 MeV, or 160 MeV for the forward, barrel, and backward regions, respectively. This reduces both the computing time and memory footprint of the algorithm.

3. *Classification.* The classifier is run using some chosen set of variables from the particle as inputs. For final state tracks, the inputs are PID variables, momentum, track parameters, and the track fit result. For photons, the inputs are the number of ECL hits, cluster energy, cluster timing, and cluster shape. For intermediate particles, the inputs include the invariant mass, momentum, vertex fit result, decay angles, and the signal probabilities of the daughter particles.
4. *Post-classifier cuts and candidate selection.* After classifying each candidate, the FEI then applies tighter post-classifier cuts. Additionally, a candidate selection is applied here, keeping only 10–20 candidates based on some ranking criterion (often the signal probability). This is a necessary step, as the number of candidates would otherwise increase exponentially, quickly exhausting the available memory. It is important that the post-classifier cuts on final-state and intermediate particles are sufficiently tight, so that correctly reconstructed candidates are not lost in later steps due to high multiplicities of background candidates from low-purity modes. No post-classifier cuts or candidate selection are applied to the final B candidates.

The FEI is composed of classifiers arranged in a hierarchical network. Figure 3.13 provides a schematic overview of the algorithm. When running the FEI, the final state particles are first reconstructed from the tracks and clusters in data. The next steps are to reconstruct intermediate particles and ultimately B mesons via various decay modes. Each classifier in this network is responsible for analysing a single reconstructed decay mode. The signal probability output of each classifier is used as an input to later classifiers, so these variables are propagated through to the final B candidates. Consequently, this gives an indication of the quality of the entire reconstruction. In this thesis, the signal probability of the B candidates is denoted $\mathscr{P}_{\text{tag}}$.

When developing any classifier, we first train it using examples with the "correct" labels to learn, and then later use the trained classifier to classify inputs it hasn't seen before. In the context of the FEI, these are referred to as the *training* and *evaluation* stages. In the training stage, the FEI is run on MC, using an appropriate `isSignal` variable for the "correct" categories for each candidate[10]. The "official" trainings of the FEI are done with $100\,\text{fb}^{-1}$ of generic $B^0\,\bar{B}^0$ and $B^+\,B^-$ MC, and the resulting weights are uploaded to the CDB. The analyst may now use these

10: Depending on the decay mode reconstructed, the labels may either be taken from `isSignal` or `isSignalAcceptMissingNeutrino`.

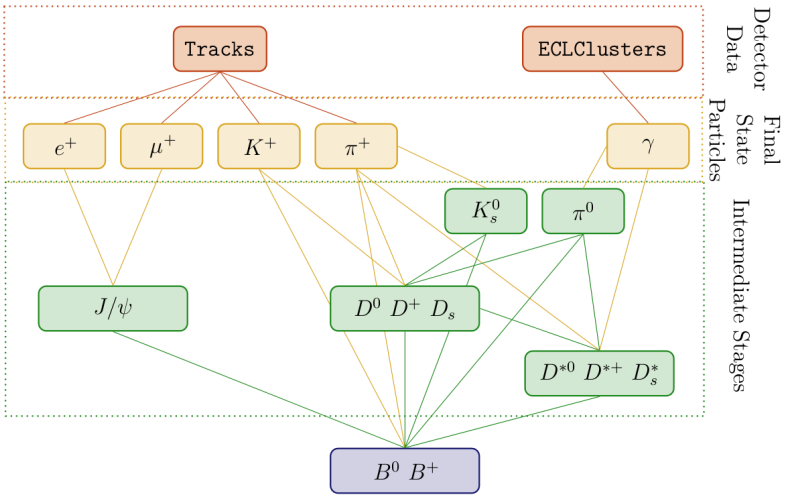


Figure 3.13: Diagram of the FEI heirarchy. The final state particles are first reconstructed from detector-level objects, and these are then used to reconstruct intermediate particles and eventually B mesons. Figure adapted from Reference [64].

weights to run the evaluation stage of the FEI, at which point it can classify MC and detector data. It is important that the training dataset is not used for later data analysis, as this will give a biased view of the FEI performance on MC. The official training is currently performed on the KEK on-site supercomputer KEKCC. Attempts have been made to train it on the grid, but have thus far run into issues regarding computing resource restrictions.

The FEI is written in a way that makes it straightforward to add additional modes. However, not all modes are able to be trained well. For example, hadronic modes involving a J/ψ may have very clean final states, but their branching fraction is low. Since the FEI is trained on a finite amount of MC, the classifiers for these modes do not see enough examples to achieve an adequate training. A larger training dataset could address this, but this also increases the resource usage of the training stage.

In the evaluation stage, the FEI may be configured to produce the following four particle lists.

- `B0:generic` and `B+:generic`: B mesons reconstructed in fully-hadronic decay modes.
- `B0:semileptonic` and `B+:semileptonic`: B mesons reconstructed in semileptonic decay modes. These also includes modes where the B decays hadronically and the D decays semileptonically, though these tend to be poorly trained due to low statistics.

In this thesis, these configurations are simply referred to as the hadronic FEI and semileptonic FEI, respectively.

The FEI weights are trained on MC, and this can lead to differences in the FEI performance on MC and detector data. These differences may arise if the input variables to a classifier have a degree of mismodelling.

For example, consider the PID variables, which are inputs to the final-state particle classifiers. If we were only applying a fixed cut on a PID variable, then the efficiency difference could be addressed using a correction table. However, the classifiers essentially learn the full PID variable distribution. Since the response of each classifier is fed into later classifiers, the propagation of any efficiency differences becomes highly non-trivial.

A common approach to correcting efficiency differences in classifiers is to *calibrate* the final output. In the case of the hadronic FEI, calibration tables have been derived, which apply scaling factors of $\epsilon_{\text{data}}/\epsilon_{\text{MC}}$, binned in the B decay mode ID and the signal probability. This binning is chosen because the $\mathscr{P}_{\text{tag}}$ output of each B classifier is independent. This approach has been successful for the hadronic FEI [95], but efforts to apply the same technique to the semileptonic FEI have not yet converged. Although calibrations have been performed [96], they have not been found to be robust to different signal sides.

[95]: Belle II Collaboration (2021), *Exclusive $B \to X_u \ell \nu_\ell$ Decays with Hadronic Full-event-interpretation Tagging in 62.8 fb^{-1} of Belle II Data*

[96]: Alina Manthei et al. (2021), *A first calibration of the Belle II semileptonic tag-side reconstruction algorithm using $B \to X \ell \nu$ decays with 62.8 fb^{-1}*

3.7. Summary

In this Chapter, we have introduced the experimental setup of SuperKEKB and the Belle II detector. In particular, we examined the individual sub-detector components and their particle detection capabilities. Throughout this, we touched on how the accelerator and detector were upgraded from their predecessors.

We then looked at the software used by the Belle II collaboration. These Sections described the simulation of MC, and the procedure for matching reconstructed MC to the underlying "truth" information. We introduced the concept of event weighting, which is important to the evaluation of systematic errors in the analysis in Chapter 5. We also introduced the software framework used for all event generation, simulation and reconstruction in the Belle II experiment. The following Chapter is entirely focused on improving one package in the framework.

Finally, we described several aspects of B physics analyses which are relevant to the analysis in Chapter 5. We defined a set of useful variables for discriminating signal and background candidates in semileptonic B decays. We described the process of *tagging*, and introduced the Full Event Interpretation, which forms a core aspect of the upcoming analysis.

Part II.

Belle II & GAMBIT work

4. Skims

5. Measurement of $|V_{ub}|$ with semileptonic FEI tagging

6. Global fit of leptoquark models with flavour constraints

4. Skims

4.1. Introduction

One of the major challenges faced by the Belle II experiment is the large amount of data produced by the detector. The lifespan target is 50 ab^{-1}, and data production campaigns on the current 200 fb^{-1} dataset amount to hundreds of terabytes of data. The ever-increasing size of the dataset poses problems for both individual analysts and for the collaboration. An analyst is typically interested in only a very small subset of the recorded events, so the full dataset contains many events which will never be included in their final analysis data files. Running over all these events lengthens the turn-around time on their reconstruction jobs. Doing so would also mean that the collaboration's computing resources are being used inefficiently.

To address these concerns, the Belle II collaboration decided to incorporate *skimming* as a part of the data processing. The key realisation is that analysts working on similar areas of physics (be that semileptonic decays, charmless decays, or tau physics) will tend to have similar basic selections. The skim production group defines a set of *skims*, which implement common selections for each physics working group. These skims are centrally produced on the grid, and analysts are intended to run on the skimmed datasets, rather than the full dataset. Once an analyst has performed their analysis reconstruction on the skimmed dataset, the final output file size is small enough to be handled on a desktop machine (on the order of gigabytes). This data flow is illustrated in Figure 4.1.

In practice, skimming is performed in the following way. A skim is defined in a steering file[1], similar to how a user would define their

1: See Section 3.5.4.

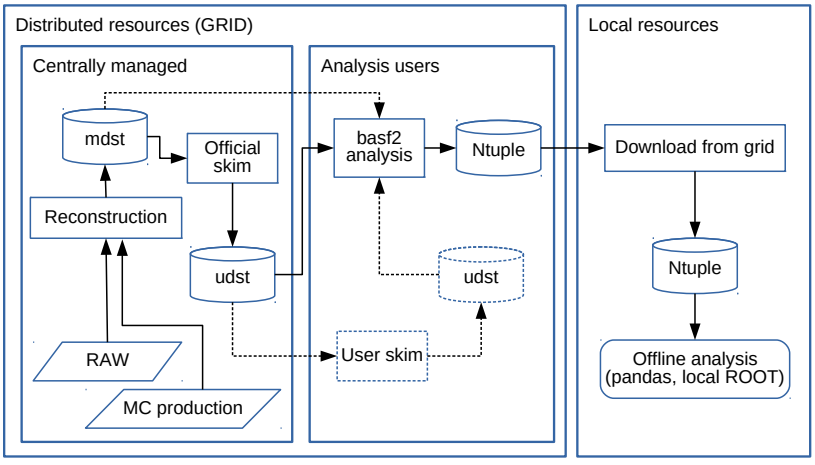


Figure 4.1: Schematic of the Belle II computing workflow. Analysts read in mDST or skimmed uDST files and produce flat ntuples for offline analysis. Dashed lines denote optional steps. Adapted from Reference [97].

own analysis reconstruction. The steering file loads the mDST, performs some basic reconstruction and selections, and then writes out a uDST. The only events written to the uDST are those for which at least one reconstructed particle list is non-empty. For events which pass the skim, the uDST contains all the original data contained in the mDST, in addition to information about the reconstructed particles. This information is saved so that the analyst is able to use the skim's reconstruction in their own analysis—this is particularly important in the FEI skims, where the reconstruction is computationally expensive. The resulting file contains more information per event, but a much smaller number of events, netting an overall smaller file size.

Skimming is fundamentally a compromise, balancing the concerns of the analysts and the computing resource managers. The skimmed file size is typically required to be between 1% and 10% of the input file size. This target file size reduction ensures that the skims do not take up as much grid storage space as the raw data itself, and also reduces the time required for the analyst to run over the full dataset. Skims can be expensive to run and store, so in order to be worthwhile, each skim should be applicable to multiple analyses. It is also critical that the selections have a minimal impact on the direction taken by analyses utilising them.

All code related to skimming is stored in the `skim` package of basf2 [86]. The package is a collection of skims and a framework for running and testing them. In particular, the package contains a registry of all available skims, a list of standard test files, and command-line tools for measuring performance statistics and producing production requests.

[86]: Belle II Collaboration (2022), *Belle II Analysis Software Framework (Basf2)*

This Chapter documents work I performed to improve the `skim` package in several ways. I will first discuss my refactoring of the core of the package, and then my work on the tools for testing and producing skims. These changes were first implemented in Release 5, so I will refer to Release 4 and Release 5 when talking about the old and new designs. The original design of the `skim` package was developed when the detector dataset was small, the number of skims was low, and very few analysts used skims. However, this design became cumbersome as more skims were developed. My work helps to future-proof the package for the development of more skims, which will become critical as the Belle II dataset continues to grow.

4.2. Defining a skim in Release 4

To understand the necessity of the changes made to the `skim` package in Release 5, we must first understand how a skim is defined. In Release 4, every skim is defined in two parts: a *list-building function* and a *skim steering file*.

The list-building function defines the main "business logic" of the skim—namely, the reconstructed decays and selections. The function takes a `basf2.Path` as an argument, adds the reconstruction to this path, and returns a list of particle list names. It may also add other modules to the path, in order to construct additional variables used in the skim. The list-building functions are organised in modules named for each physics working group (WG). Code 3 illustrates the typical features of a list-building function.

Code 3: Example list-building function for a skim in Release 4.

```
import modularAnalysis as ma

def InclusiveLeptonicList(path):
    # Apply continuum suppression cut, which requires
    # event shape builder module
    ma.fillParticleList(
        "pi+:clean", cut="pt > 0.1", path=path
    )
    ma.fillParticleList(
        "gamma:clean", cut="E > 0.1", path=path
    )
    ma.buildEventShape(
        inputListNames=["pi+:clean", "gamma:clean"],
        foxWolfram=True,
        path=path,
    )
    ma.applyEventCuts("foxWolframR2 < 0.5", path=path)

    # Fill e- list 'skim' from pre-filled 'all' list
    ma.cutAndCopyList(
        "e-:skim", "e-:all",
        "useCMSFrame(p) > 2.0", path=path
    )
    # Reconstruct B candidates using just electron
    ma.reconstructDecay(
        "B-:skim -> e-:skim", "nTracks > 2", path=path
    )

    # Return a list of skim names
    return ["B-:skim"]
```

The skim steering file is an executable script that imports and runs the list-building function. The file also performs several other tasks which are identical for every skim: adding the ROOT input module to the path, running a collection of helper functions from the `skimExpertFunctions` module, and then executing the path. Each skim in Release 4 requires its own steering file, so the `standalone/` directory exists to store them. A typical steering file is illustrated in Code 4.

The one aspect which does differ between skims is the loading of *standard particle lists*. In Code 3, the `e-:skim` particle list is copied from the `e-:all` list, which is itself loaded in the steering file by the call to `stdCharged.stdE`. If this standard list is not filled prior to calling the

Code 4: Example steering file for a skim in Release 4.

```
import basf2 as b2
import modularAnalysis as ma
import skimExpertFunctions as expert
from skim.leptonic import InclusiveLeptonicList
from stdCharged import stdE

# Create a new path and add RootInput module
path = b2.Path()
input_files = expert.get_test_file("MC12_mixedBGx1")
ma.inputMdstList(filelist=input_files, path=path)

# Load the standard particle lists required by skim
stdE("all", path=path)

# Run the skim's list-building function
skim_lists = InclusiveLeptonicList(path)

# Add RootOutput module for skimmed uDST output
decay_mode = expert.encodeSkimName("InclusiveLeptonic")
expert.skimOutputUdst(
    skimDecayMode=decay_mode,
    skimParticleLists=skim_lists,
    path=path
)

# Start the event loop
expert.setSkimLogging(path)
b2.process(path)
```

list-building function, basf2 would crash when trying to access the undefined list. There is no signposting of this dependency in `InclusiveLeptonicList`, except for a close reading of the source code. Thus, the full definition of a skim is split across the list-building function and the steering file.

4.3. Defining a skim in Release 5

The refactoring of the `skim` package was brought about by the realisation that each skim has some associated metadata, and common functions that are executed according to that metadata. The metadata of a skim includes the skim name, an eight-digit skim code, the required standard particle lists, and the sample label for testing the skim. The functions influenced by the skim metadata are all the functions in the skim steering file which add modules to the path. Whenever there is a coupling between data and functions, we can consider whether a class should be written to package both together. I made the decision to create the `BaseSkim` class, which is the basis of the `skim` package redesign.

`BaseSkim` is defined as an *abstract base class* [98], and all skims are defined in Release 5 as subclasses of it. The base class defines particular

[98]: Van Rossum et al. (2009), *Python 3 Language Reference*

basic behaviours, such as the order of functions to be run when adding modules to the path—however, certain behaviours are deliberately left unimplemented, such as the content of those functions. These unimplemented methods are defined in `BaseSkim` as *abstract methods*, using the `@abstractmethod` decorator. If a subclass attempts to instantiate without implementing all abstract methods, then an error is thrown. Defining the base class in this way sets up a common framework for writing skims, with formalised requirements for the subclasses.

We will take the example skim from Section 4.2 and show how to rewrite it in Release 5. We start by defining a class in the relevant WG module, naming it with the skim name and subclassing from `BaseSkim`.

```
class InclusiveLeptonic(BaseSkim):
    ...
```

Code 5: Example of writing a skim in Release 5: subclassing from `BaseSkim`.

We then define a set of variables which describe the author/s and purpose of the skim. All of these are defined in `BaseSkim` as abstract properties, so `InclusiveLeptonic` is required to override them.

```
class InclusiveLeptonic(BaseSkim):
    __authors__ = "Phil Grace"
    __contact__ = "Phil Grace
     ↪ <philip.grace@adelaide.edu.au>"
    __category__ = "physics, leptonic"
    __description__ = "Fictional skim for illustration."
    ...
```

Code 6: Example of writing a skim in Release 5: adding required properties.

We now move the standard particle lists from the steering file to the method `load_standard_lists`, which we override from the `BaseSkim` definition. Overriding this method is optional, as not all skims require standard lists. As such, `load_standard_lists` is simply given an empty definition in `BaseSkim`, rather than being declared as an abstract method.

```
class InclusiveLeptonic(BaseSkim):
    ...
    def load_standard_lists(self, path):
        stdE("all", path=path)
```

Code 7: Example of writing a skim in Release 5: adding standard lists calls to `load_standard_lists`.

Finally, we move the content of the list-building function to the method `build_lists`.

Code 8: Example of writing a skim in Release 5: moving the skim list-building function to `build_list`.

```
class InclusiveLeptonic(BaseSkim):
    ...
    def build_list(self, path):
        ma.fillParticleList(
            "pi+:clean", cut="pt > 0.1", path=path
        )
        ma.fillParticleList(
            "gamma:clean", cut="E > 0.1", path=path
        )
        ma.buildEventShape(
            inputListNames=["pi+:clean", "gamma:clean"],
            foxWolfram=True,
            path=path,
        )
        ma.applyEventCuts("foxWolframR2 < 0.5",
         ↪  path=path)
        ma.cutAndCopyList(
            "e-:skim", "e-:all",
            "useCMSFrame(p) > 2.0", path=path
        )
        ma.reconstructDecay(
           "B-:skim -> e-:skim", "nTracks > 2", path=path
        )

        # Return a list of skim names
        return ["B-:skim"]
```

With these attributes and methods defined, this is now a fully functional skim. There are other additional methods and attributes that can be added to the class to alter the way the skim is run, but these are optional.

`BaseSkim` subclasses are callable, because `BaseSkim` defines the special "dunder method" `__call__`. This method runs `load_standard_lists` and `build_lists`, and uses the returned value of `build_lists` to set up the uDST output module. It also customises the logging level of any modules listed in the skim's `NoisyModules` attribute, in order to reduce the size of the log files. All of the functions previously run in every skim steering file are now factored into this single method, and the code required to run the skim is greatly reduced. Code 9 illustrates the simplicity of a skim steering file in Release 5, in contrast to Code 4.

We will now examine the ways in which this new structure solves a large number of issues with the original `skim` package design.

```python
import basf2 as b2
import modularAnalysis as ma
from skim.WGs.leptonic import InclusiveLeptonic

path = b2.Path()
ma.inputMdstList(filelist=[], path=path)

# Initialise Skim object
skim = InclusiveLeptonic()
# Use __call__ method to add modules to path
skim(path)

b2.process(path)
```

Code 9: Example steering file for a skim in Release 5.

4.4. Issues addressed by Release 5 refactoring

The refactoring solved a large number of problems relating to the organisation, readability and maintainability of the `skim` package. Many of these issues were not problematic when the package was first written, but became unavoidable as more skims were written. In addition to solving existing issues with the package, the refactoring also allowed for the implementation of new functionality which would have been difficult in the original framework.

The primary goal achieved by the `BaseSkim` abstraction is the localisation of each skim's definition to a single class. Previously, the code defining the reconstruction and selections was split between the steering file and WG module. Any user wishing to understand a skim's definition would have to read both source files. To further complicate things, there was not necessarily any relation between the skim name and the name of the associated list-building function. Some authors used the quasi-standard convention of *[skim name]*`List`, but this was not enforced. The `BaseSkim` design stands in contrast to this: the list-building code is always found in the `build_lists` method of the class with the skim's name, so the user knows exactly where to find it. As a result, the new design reduces the mental overhead for reading and understanding a skim.

The localisation also made each skim's dependencies more explicit, and packaged them into a single function call. In the original design, the required standard particle lists are only listed in the steering file. In the `BaseSkim` design, the required particle lists are always found in the `load_standard_lists` method, so the dependencies are very clearly signposted. Furthermore, `BaseSkim.__call__` runs `load_standard_lists` before building the skim particle lists. This means that the responsibility to include the correct dependencies is shifted away from the user.

The original package design led to a large amount of duplicated code that was challenging to maintain. The `standalone/` directory contained upwards of seventy steering files. These all shared the same basic code required run a skim, and differed in only a few lines. If it was decided that skim processing should be modified in some slight way, then that change needed to be propagated to every single steering file—needless to say, this was highly error-prone.

All of this duplicated code was made redundant by the Release 5 refactoring. The primary reason was that any shared elements of processing were factored out into the `BaseSkim.__call__` method, and any unique elements were implemented in each skim's `load_standard_lists` and `build_lists` methods. I developed two command-line tools which completely supplanted the need for the directory of steering files:

- `b2skim-run` is a tool which takes a skim name as an argument, imports the appropriate skim class, adds the modules to the path, and processes the path. This performs all the functions previously performed by the individual steering files, but in one generalised script.
- `b2skim-generate` is available for cases where a user requires a steering file for a single skim, such as submitting the steering file to the grid. This tool takes a skim name as an argument, and writes a steering file via to a template written in the Jinja2 templating language [99]. Writing a template was possible because steering files in the new format differ only in which skim they import[2].

[99]: Ronacher (2022), *Jinja*

2: See Code 9, for example.

Thus, by refactoring all skims into `BaseSkim` subclasses, we have a framework for running skims and generating steering files on-demand, while completely removing the need to maintain a directory of over seventy nearly-identical scripts.

Factoring the common parts of the processing into a single function also allows for more complex features to be implemented, and have those features immediately available to all skims. One such feature requested by users was *skim flags*, which are boolean variables that indicate whether or not an event passed the skim. The main difficulty in implementing this feature was the possibility of conditional paths (a subtlety of combined skims that will be explained in Section 4.5). The solution ended up involving adding two custom basf2 modules to the path in a carefully handled way: one to initialise the skim flag, and the other to update it after running the skim by checking the skim list contents on the appropriate path. By modifying `BaseSkim.__call__` to include this change, skim flags are now standard functionality in all skims. This kind of complex extension to skim processing would not have been feasible in the old framework, as the amount of copy-and-pasting would have led to numerous mistakes.

Finally, the refactoring into `BaseSkim` subclasses improved the organisation of more complex skims which utilise multiple "sub-list" functions to reconstruct intermediate particles. In Release 4, the sub-list functions existed on the same level as the main list-building functions in the WG modules. Consequently, there was no clear distinction between the two function varieties, nor was there a clear indication of which skim the sub-list functions belonged to. In Release 5, each skim was refactored into its own class, so there was an obvious place for the sub-list functions to go: inside the skim class to which they belonged. This grouping together of related functions provides a natural way of organising complex skims without compromising readability.

4.5. Combined skims

Due to the limits of the data production system, the most efficient way to produce all the skims in a campaign is to group them into *combined skims*, which each run between two and ten skims in the same steering file. Each combined skim of N skims produces N output files from one input file, which are then catalogued on the grid under a directory for each skim. Skims with similar computing resource requirements are typically combined, since the production system performs best with these kinds of groupings. Combined skims are a critical aspect of skim production, and my work on the package improved them in several ways.

In Release 4, the combined skims were defined in a set of steering files in the `combined/` directory, similar to the individual steering files in `standalone/`. The combined steering files were far more cumbersome to write and maintain than those of individual skims, despite there being fewer combined steering files. To write a combined skim, the author needed to copy all prerequisite standard particle lists from each individual steering file, manually delete duplicate calls, copy the list-building functions, and then write additional boilerplate lines to process each skim. Since all of this was done by hand, the combined steering files were incredibly error-prone.

The combined skim design in Release 4 was also very inflexible, since the set of combined skims was hard-coded into the package. A skim campaign can be a chaotic time for the skim production team, with the list of requested skims changing throughout. The hard-coded set of combined skims made it difficult to quickly add a new skim to an existing steering file or reorganise the combined skims at production time. What was needed was a solution that could take individual skims and combine them automatically without manual intervention.

The solution I developed was the `CombinedSkim` class, which I defined as a subclass of `BaseSkim`. It takes a list of `BaseSkim` subclass ob-

jects, and provides a `__call__` method which runs the methods of all constituent skims. Simply running each skim one after another would not have been a sufficient solution, as this would lead to duplicate standard list calls and skims possibly interfering with each other. As such, `CombinedSkim.__call__` is carefully written to avoid potential interference. Code 10 demonstrates how straightforward it is for a user to write a steering file with `CombinedSkim`. Like in the `BaseSkim` case, responsibility for correctly loading prerequisite lists is shifted away from the user.

Code 10: Example steering file for a combined skim in Release 5.

```
import basf2 as b2
import modularAnalysis as ma
from skim import CombinedSkim
from skim.WGs.foo import OneSkim, TwoSkim
from skim.WGs.bar import RedSkim, BlueSkim

path = b2.Path()
ma.inputMdstList([], path=path)

# Initialise CombinedSkim object with BaseSkim objects
skim = CombinedSkim(
    OneSkim(), TwoSkim(), RedSkim(), BlueSkim(),
)
# Add all skim modules to path
skim(path)

path.process()
```

The key consideration in writing a general method of combining skims is ensuring that the running of one skim does not interfere with that of another. During development, it was recognised that skims could interfere with each other if they include *event-level cuts*, which remove entire events from the processing. Event-level cuts use variables which are independent of any particle reconstruction, such as the event shape or charged track count, and are a useful tool to reduce the computing time of a skim. Prior to Release 4, event-cuts in skims were applied with `modularAnalysis.applyEventCuts`, which would create a "dead-end" path branch for events not passing the given cut; events on the dead-end path would not be used in any further processing. This stands in contrast to `modularAnalysis.applyCuts`, which removes candidates from a particle list, but does not remove the entire event from processing. Consequently, the event-level cuts of one skim would inadvertently be applied to all subsequent skims in the path.

The work-around to this issue was challenging to implement, but ultimately possible due each skim having an associated object, and `CombinedSkim` handling path manipulation in its `__call__` method. The solution was to allow each skim to potentially create its own *conditional path*. At the start of `BaseSkim.__call__`, the path passed as an argument would be set to the attribute `._MainPath`[3]. The `BaseSkim` method

3: The leading underscore suggesting that this is a variable not intended for users to access directly.

Module	Parameters
`ParticleLoader`	`decayStrings: ["e+:good"]` `writeOut: True`
`ParticleListManipulator`	`outputListName: ["e+:good"]` `writeOut: True`
`ParticleSelector`	`cut: <`e^+_{good} *selection*`>` `decayString: "e+:good"`

Table 4.1: Modules and parameters added to the path by running `stdCharged.stdE`. These module-parameter combinations are inspected by `CombinedSkim.load_standard_lists` when building the combined skim path.

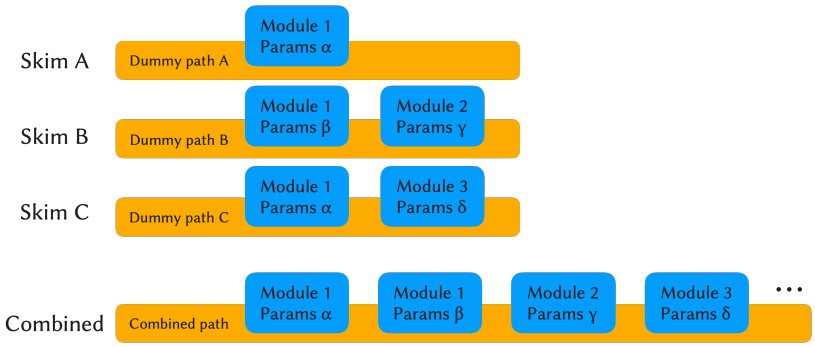


Figure 4.2: Example of how `CombinedSkim.load_standard_lists` adds modules to the path. Module 1 with parameter set α is repeated in Skims A and C, so it is only added to the main path once.

`skim_event_cuts` then creates a branch in the path for events passing the given cut, and assigned this conditional path to the attribute `._ConditionalPath`. Any modules to be included after that point must be added to the conditional path. While it would have been technically possible to implement this scheme in Release 4, the manual writing of combined steering files would have made the chance of error very high. This solution maintains the computational benefits of event-level cuts, while ensuring that cuts in one skim cannot impact the processing of another.

The second major challenge to writing `CombinedSkim` was the handling of standard particle lists. In Release 4, if two skims both loaded the same standard particle list, the author of the combined steering file would need to make sure to copy only one call. This is clearly unacceptable if we want an automatic method of combining skims. In the solution I developed, `CombinedSkim.load_standard_lists` creates a set of dummy paths for each skim, which are not intended to be processed. It then runs each skim's `load_standard_lists` method on one of those dummy paths. For example, running `stdCharged.stdE` adds the modules listed in Table 4.1 to the path.

`CombinedSkim` then inspects the dummy paths and creates a list of unique module-parameter combinations, which it then adds to the main skim path—this is illustrated in Figure 4.2. In this way, all the required particle lists are added to the path without any duplicate calls, and the loading order is preserved, in case one standard list requires another. This general method of combining the standard list calls of skim was only possible due to the refactoring of each skim's standard list calls into the `load_standard_lists` method.

Since writing `CombinedSkim` two years ago, I have learned more about object-oriented design and have had time to reevaluate the design choices I made. It was clearly a mistake to define `CombinedSkim` as a subclass of `BaseSkim`. Inheritance should be reserved for "is-a-kind-of" relations, and no such relation exists here. `CombinedSkim` is an entirely different kind of object to `BaseSkim`, and acts more like a container class. In fact, the very act naming it `CombinedSkim` may have contributed to the confusion; a more suitable noun like `SkimCombiner` would have suggested an different design. Furthermore, the `CombinedSkim.__init__` method took many of the same arguments as `BaseSkim`, and would use these to override the corresponding parameters of the constituent skims. While this reduces the amount of typing the user needs to do, it needlessly complicates `CombinedSkim`'s design, and these arguments should have been left as the responsibility of `BaseSkim`. Despite these non-optimal decisions, the current implementation is entirely functional, and has been used in numerous successful skim campaigns.

4.6. Other changes to `skim` framework

We will now look at two other aspects of the framework which were improved in Release 5: the package documentation and the registry of available skims.

4.6.1. Documentation

The documentation of skims in Release 4 was very inconsistent, largely due to the fragmented nature of the skim definition. This aspect of the package was also substantially improved by the refactoring of skims into `BaseSkim` classes.

Firstly, the reorganisation of skim definitions contributes to *self-documentation*. Each skim class is named for the skim it defines, and all the "business logic" of the skim is defined in a small number of standard methods. These two aspects of the Release 5 `skim` package aid users by placing information about a skim in a location they can expect to find it.

I also wrote the `@fancy_skim_header` decorator, which modifies the class docstring with a section containing key skim information. This information is taken from the required properties such as `__authors__` and `__category__` shown in Code 6. By simply adding the decorator as shown in Code 11, the compiled Sphinx documentation [100] will display the information block shown in Figure 4.3. The skim code is retrieved from the registry, so that it is included in the documentation without the author needing to manually copy it, as was the case in Release 4. The decorator also converts a properly-formatted contact name and email

[100]: Brandl et al. (2022), *The Sphinx Documentation Generator*

```python
@fancy_skim_header
class InclusiveLeptonicList(BaseSkim):
    ...
```

Code 11: Example usage of the `@fancy_skim_header` decorator.

Figure 4.3: Example of an auto-generated information block in the compiled Sphinx documentation.

address into a clickable `mailto:` link. This information block sets a minimum standard of documentation for all skims, which authors may then build on.

Finally, I extended the existing documentation to give detailed instructions on writing and running skims, and using the available tools. The command-line tools that existed were undocumented, so I added documentation and usage examples; I also did the same for the new tools I wrote. This was performed with the help of the Sphinx extension `sphinx-argparse` [101], which generates formatted documentation directly from a tool's argument parser. By including these tools in the documentation, users are made more aware of them and thus more likely to use them. All of these documentation improvements assist skim authors and make the package more accessible to newcomers.

[101]: Rudakov et al. (2022), *sphinx-argparse*

4.6.2. Skim registry

The skim registry is a required part of the skim package, as each skim must have an associated unique eight-digit *skim code*. Skim codes are necessary because of the LPN length restriction of the production system. The digits follow a convention which encodes information about the WG and reconstructed decays. The registry exists to convert between names and codes, and I refactored it in order to implement further functionality.

In Release 4, the registry existed as a list of (*code*, *skim name*) pairs parsed by look-up functions; this was another case where data and functions are packaged together, and hence an opportunity to unify them in a class. I defined a registry class which stores the list as an attribute, and moved the look-up functions inside the class as methods. Doing so allowed for additional methods to be written to act on the registry data; in

Code 12: Example usage of the `Registry` object to convert between skim names and skim codes.

```
>>> from skim.registry import Registry
>>> Registry.encode_skim_name("feiHadronicB0")
"11180100"
>>> Registry.decode_skim_code("11180100")
"feiHadronicB0"
```

particular, the `Registry.names` attribute which provides a list of all registered skims, and the method `Registry.get_skims_in_module` which lists all skims in a given WG module.

A crucial improvement I made was a set of unit tests to ensure that the registry always held correct information about the existing skims. The tests check that no two skims have the same name or skim code, and check that each code has the correct format. Additionally, the tests inspect the WG modules to ensure that all skims listed in the registry are defined in the modules, and *vice versa*. These tests are run as a part of basf2's Continuous Integration, so that code cannot be merged to the `main` branch if the skim registry is incorrect.

An instantiated registry object is provided in `skim/registry.py`, so users can import the object and directly use it as shown in Code 12.

The `Registry` object provides a way of retrieving information about the `skim` package at runtime. This functionality was very useful in extending the performance testing tools, which we will now discuss.

4.7. Skim performance testing

Skims are ultimately run on the data production system, and hence are subject to a number of constraints, including CPU time, output file size and memory usage. In order for skims to be reliably processed, we require a standard way to measure these statistics. A set of scripts existed in Release 4 to perform these tests, but they were difficult to maintain and extend with new functionality. During my work on the `skim` package, I generalised and extended these tools, and reduced the hard-coding of inputs.

I developed two tools for measuring skim performance: `b2skim-stats-submit` and `b2skim-stats-print`. The former submits jobs to the KEKCC batch queue, running one or more skims on a set of standard test samples. The output of the $N_{\text{skims}} \times N_{\text{samples}}$ jobs is saved to a `log/` directory, with each job saved to a unique subdirectory. `b2skim-stats-print` then reads these log files and produces tables of performance statistics in a requested format.

While developing the tools, I modified their methods of retrieving data to be more reliable. In the Release 4 version of the tools, all the

required data was retrieved by parsing the human-readable log file. Manually parsing log files is generally recommended against, as such methods are liable to break when slight changes are made to text formatting. Where possible, I avoided this manual parsing by utilising the machine-readable output available. When basf2 is passed the `--job-information` argument, it writes a JSON file containing the number of processed events, output file size, and number of events in the output file. Since this JSON file conforms to a defined schema [102], retrieving the information from there was a more robust method than parsing. Certain values such as average candidate multiplicity were not written to this JSON file and still needed be obtained from parsing, but the number of such cases was kept to a minimum[4].

[102]: JSON Schema Organisation, *JSON Schema*

A major limitation of the Release 4 version of the tools was that the list of all skims was hard-coded in the tools. In order to test only the skims they wanted, the user was required to open up the script in a text editor, comment out several lines, and then run the script. My improvements to the skim registry (described in the previous Section) allowed for this hard-coded list to be removed. In its place, I accessed the list of skims via the `Registry` object, and introduced a command-line argument to the tool, for the user to select the skims they wished to test.

4: This was imposed by the structure of the basf2 data processing code. Modules do not expose all of their internal calculations, but it is straightforward to write the results to logger. Outside of skim performance testing, this information is primarily used for debugging purposes, so this did not warrant any significant rewriting of code.

A related limitation was the hard-coding of combined skim combinations. The resulting inflexibility also made it difficult to test new combinations at production time. This setup was vastly simplified with the introduction of the `CombinedSkim` class. I added a separate running mode to the tools for testing combined skims, and provided a method of defining combined skims via a simple YAML file, as illustrated in Code 13. The generalisation of combined skim testing was critical in ensuring that accurate performance statistics are passed to the data production system[5].

5: See Section 4.8

```yaml
# Combined skim name
DarkCombined:
 # List of individual skims
 - SinglePhotonDark
 - ALP3Gamma
 - TauLFV

TauCombined:
  - TauGeneric
  - TauThrust
```

Code 13: Example YAML file passed to `b2skim-stats-submit` to define combined skims at runtime.

The testing tools all read from a standard list of test sample locations, which was hard-coded in the package in Release 4. However, the fact that it existed as a part of basf2 meant that it could only be updated with the release cycle, severely restricting skim testing during the production cycle. This issue was addressed by moving the list outside of the repository altogether, into a set of YAML files (one for each production

campaign). All that remained hard-coded in the package was the location of a symbolic link to the latest file. The YAML files and the symbolic link destination could then be updated to the newest samples without requiring any changes to the `skim` package.

The standard file list in Release 4 was a script containing a mapping of labels and sample locations. This simple setup was functional, but meant that there could be no metadata attached to the samples, such as production campaign number, simulated process, simulated beam background level and beam energy. This metadata is necessary for passing appropriate statistics to the data production system, but it could only be guessed from parsing the label. To provide this ability, I changed the file list from a simple label-location mapping to a list of dictionaries containing all relevant metadata. I wrote the `DataSample` and `MCSample` classes to read from and write to this dictionary format, storing the values as attributes. These classes include the property `printable_name` for human-readable labels, and the property `encodeable_name` for indexing JSON files. This extension of the file list, and its relocation outside the package, greatly improved its utility.

The previous versions of the scripts were not transparent in how they calculated each statistic, and it was not straightforward to modify or add new statistics. I addressed this by defining all statistics in a list of `dicts`, each containing the name and description of the value, several boolean flags to toggle printing visibility, and the function to calculate the statistic [6]. Any user wishing to read the exact definition of a statistic could then look at the function, and any user wishing to add a new statistic could add another element. In hindsight, this situation would have been better served by introducing a lightweight `StatisticCalculator` class, as the expected arguments and default values are clearer in a class than in a `dict`. Regardless, my implementation is functional and improves the clarity and extensibility of the calculations.

6: Functions are "first-class objects" in Python, meaning they can be passed around like all other objects, including being stored in `dicts`.

The separation of skim performance testing into two scripts is a slight drawback to my implementation. At present, it is left to the user to allow sufficient time for the batch jobs to finish before running `b2skim-stats-print`. One option for combining these scripts would be to utilise the data pipeline package `b2luigi` [103]. `b2luigi` is able to submit and monitor batch jobs and detect when jobs fail, hence making it an ideal tool for managing all aspects of both scripts. This task has been left as future work for developers on the collaboration issue tracker, Jira.

[103]: Michael Eliachevitch et al. (2022), *b2luigi*

The final improvement I made to the testing tools was extending the printing capabilities to produce human- and computer-readable outputs. The old scripts constructed the printable string at the same time as calculating each statistic, so there was no internal object which stored all the values. I refactored the main loop to instead build a nested `dict` (indexed by skim name and sample type), and wrote functions to print this

data structure to terminal output, Markdown, or JSON. The Markdown printer found immediate use with skim developers, who share these tables in pull requests on BitBucket [104]. The JSON printer was the most impactful of these functions, as it built a bridge between the `b2skim-stats` tools and the skim production tools. The presence of this interface between tools ensure that up-to-date performance statistics are used when creating production requests.

[104]: Atlassian Corporation, *BitBucket*

4.8. Skim production tools

Skims are produced on the grid via a set of production request files. These requests are in JSON format, and set all the parameters of the production: campaign and production name, beam energy, expected file size per event, expected CPU time, list of input LPNs, and the steering file to be executed. The tool `b2skim-prod` exists to produce the files, and I made two key improvements to it, both of which help to ensure productions are reliably run.

The first major improvement to `b2skim-prod` came directly from the updates to the `b2skim-stats` tools (described in the previous Section). It is vital that accurate performance statistics are included in the requests. An incorrect CPU time estimate can cause jobs to time out, and an incorrect output file size estimate may cause the production system to merge files to sizes it can't handle. Both of these situations cause productions to crash and create a large amount of work for the skim production managers. Performance statistics can vary greatly across different skims and different sample types, so it is important that the values are taken from tests on the same sample type. Previously, outdated or approximate values were used, as there was no straightforward method of passing the results of the `b2skim-stats` tools to `b2skim-prod`. The addition of JSON printer functions provided such a method. Furthermore, my improvements to the standard test file list helped to ensure the values came from up-to-date test samples. Finally, my implementation of `CombinedSkim` simplified the process of testing new combined skim combinations with `b2skim-stats`.

The second improvement to `b2skim-prod` was how the included steering files are handled. In earlier versions of the tool, the production manager would need to manually copy the steering file from the latest basf2 version. However, my refactoring of the package completely removed the directories of steering files and replaced it with a tool to automatically generate them on-demand. I extended this tool to also generate combined steering files, which was only made possible by the straightforward interface of `CombinedSkim`. With this tool available, I included the steering file generation as a part of the main `b2skim-prod` method. Immediately after generating the file, `b2skim-prod` then runs a "dry-run"

function to check that the steering file executes properly[7]. This way, the production manager does not need to copy an up-to-date file, and can instead declaratively define the set of combined skims during production.

7: The dry-run function builds the path and runs `basf2.process` with zero events. If the file crashes, then `b2skim-prod` will inform the user.

The two improvements I made to the skim production tool were (1) integrating it with the skim performance testing tools and (2) leveraging the automatic steering file generation. These changes minimise the amount of manual human intervention required, and reduce the turnaround time on production campaigns.

4.9. Summary

My work on the `skim` package essentially left no aspect untouched. I altered the core of the package by refactoring skims into `BaseSkim` subclasses and developing `CombinedSkim` to combine them. These changes had immediate benefits for the readability and maintainability of the skim code. I made many improvements to the tools used for skim testing and production, which also benefited from the core refactoring.

Skims are an important part of the Belle II analysis workflow, particularly as ever-increasing amounts of data are recorded. My contributions to the `skim` package have streamlined the skim definitions, extended the testing capabilities of the tools, and helped to ensure that skims are produced in a timely manner for analysts.

Measurement of $|V_{ub}|$ with semileptonic FEI tagging

5.

This Chapter documents the measurement of $|V_{ub}|$ with 189.9 fb^{-1} of Belle II data, using $\Upsilon(4S)$ candidates reconstructed with SL FEI tagging. We extract $|V_{ub}|$ from a fit to the q^2 spectrum of $B \to \pi\ell\nu$ for B^+ and B^0. We normalise the $B \to \pi\ell\nu$ distribution to $B \to D^*\ell\nu$ (also reconstructed with SL FEI), as there are ongoing challenges to developing a general-purpose SL FEI calibration. This reduces the systematic uncertainty due to tagging efficiency, and provides a method for using the SL FEI in the absence of calibration.

We will first cover the theoretical motivations for this analysis, before moving on to analysis procedure. We describe the simulated samples used, the $\Upsilon(4S)$ candidate reconstruction, data processing, and corrections applied to the MC. The fit procedure is in two parts: a template fit to extract event yields from the reconstructed candidates, and a fit to ratios of the yields to extract $|V_{ub}|$. We check the fit procedure by first fitting to the MC expectation. Finally, we fit to the detector data, and derive a combined fit result of $(3.598 \pm 0.266) \times 10^{-3}$, which is consistent with the current world average for exclusive $|V_{ub}|$ determinations.

5.1. Motivation

The CKM matrix elements are free parameters of the SM which govern transitions between quark flavours. V_{ub} and V_{cb} are measurable via tree-level semileptonic B decays, which Belle II is well suited to studying. There are long-standing discrepancies between the inclusive and exclusive measurements of these two quantities. As Belle II enters its high-luminosity phase, its large dataset will provide critical insights into this puzzle. Furthermore, V_{ub} is of particular interest, as it is the least precisely measured of all CKM elements.

The most precise measurements of $|V_{ub}|$ come from studies of the decay $B \to \pi\ell\nu$[1]. The first such extraction was performed by the CLEO collaboration [105]. This measurement and similar early measurements performed by the BaBar collaboration reconstructed only the signal B decay, and used the missing 4-momentum of the whole event to infer the neutrino 4-momentum [106]. This approach was appropriate for the small datasets available, but limits the precision of the reconstruction. More recent measurements by Belle and BaBar employ the tagging approach described in Section 3.6.1 [91, 92, 107, 108]. These analyses demonstrate the use of both hadronic and semileptonic tagging for measuring $B \to \pi\ell\nu$.

1: Throughout this Chapter, ℓ refers to electrons and muons.

[105]: CLEO Collaboration (2003), "Study of the q^2-dependence of $B \to \pi\ell\nu$ and $B \to \rho(\omega)\ell\nu$ Decay and Extraction of $|V_{ub}|$"
[106]: BaBar Collaboration (2007), "Measurement of the $B^0 \to \pi\ell\nu$ Form-Factor Shape and Branching Fraction, and Determination of $|V_{ub}|$ with a Loose Neutrino Reconstruction Technique"
[91]: BaBar Collaboration (2006), "Measurement of the $B \to \pi\ell\nu$ Branching Fraction and Determination of $|V_{ub}|$ with Tagged B Mesons"
[92]: Belle Collaboration (2007), "Measurements of branching fractions and q^2 distributions for $B \to \pi\ell\nu$ and $B \to \rho\ell\nu$ Decays with $B \to D^{(*)}\ell\nu$ Decay Tagging"
[107]: BaBar Collaboration (2008), "Measurements of $B \to \{\pi, \eta, \eta'\}\ell\nu$ Branching Fractions and Determination of $|V_{ub}|$ with Semileptonically Tagged B Mesons"
[108]: Belle Collaboration (2013), "Study of Exclusive $B \to X_u\ell\nu$ Decays and Extraction of $|V_{ub}|$ using Full Reconstruction Tagging at the Belle Experiment"

Tagging allows for a much cleaner signal reconstruction, and greater background rejection. However, it also reduces the reconstruction efficiency, so such analyses can only be performed on large datasets. The reconstruction efficiency of tagged analyses is directly impacted by the branching fraction of the reconstructed tag modes. A typical approach is to explicitly reconstruct a few tag modes with the highest branching fractions. The Belle II collaboration developed the FEI to address this limitation of tagged analyses.

The two primary goals of the FEI are to reconstruct a large number of tag modes, and to provide a quantity indicating the quality of a given tag. The hadronic FEI has been successfully used in several published Belle II analyses [109, 110]. In particular, Reference [111] measures $B \to \pi e^+ \nu_e$ using hadronic FEI tagging.

The SL FEI has been comparatively underutilised, despite the possibility for SL tagging to provide access to a greater number of signal events. This has largely been due to the difficulty in calibrating the SL FEI. This is an issue which will be discussed further in Section 5.5.8.

In this analysis, we measure $|V_{ub}|$ using a normalisation mode approach. Belle II is capable of measuring absolute branching fractions. However, without calibration the SL FEI cannot yet be reliably used for this purpose. Instead, we measure the relative branching fraction of the signal mode $B \to \pi \ell \nu$ to the normalisation mode $B \to D^* \ell \nu$, and extract $|V_{ub}|$ from the ratio. The normalisation mode is chosen for its similarity to the signal mode, and both the signal and normalisation modes are reconstructed using the same SL FEI reconstruction. Measuring the ratio enables us to use the uncalibrated SL FEI to obtain meaningful results. Furthermore, in analyses which measure branching fractions using the hadronic FEI, the calibration procedure gives rise to a leading systematic uncertainty. By taking the ratio of two branching fraction, this systematic uncertainty is greatly reduced.

[109]: Belle II Collaboration (2022), *Measurement of the photon-energy spectrum in inclusive $B \to X_s\gamma$ decays identified using hadronic decays of the recoil B meson in 2019-2021 Belle II data*
[110]: Belle II Collaboration (2022), *Reconstruction of $B \to \rho \ell \nu_\ell$ decays identified using hadronic decays of the recoil B meson in 2019 – 2021 Belle II data*
[111]: Belle II Collaboration (2022), *Study of Exclusive $B \to \pi e^+ \nu_e$ Decays with Hadronic Full-event-interpretation Tagging in 189.3 fb^{-1} of Belle II Data*

5.2. Dataset and software

Here we will outline the types of datasets analysed, and the software used for simulation and reconstruction.

The detector dataset used in this analysis is $189.9\,\text{fb}^{-1}$ of Phase 3 Belle II data. The simulated datasets include generic samples and signal sample. The generic MC samples are $500\,\text{fb}^{-1}$ equivalent of: $\Upsilon(4S) \to B^0\overline{B}^0$ including neutral B mixing, $\Upsilon(4S) \to B^+B^-$, and $e^+e^- \to q\bar{q}$ continuum. The analysis selections very effectively remove other continuum processes such as $e^+e^- \to \tau^+\tau^-$, so these processes are not included in the samples. Four signal MC samples were used for the signal mode, each containing 50M events—these are resonant and non-resonant $B \to X_u \ell \nu$,

Dataset name	"Moriond 2022" dataset / proc12 and buckets 16—25
MC production campaign	MC14ri_a
FEI version	`FEIv4_2021_MC14_release_05_01_12`
basf2 version for MC skims	`release-05-02-11`
basf2 versions for data skims	`release-05-01-24` `release-05-02-11`
basf2 version for reconstruction	`light-2207-bengal`

Table 5.1: Dataset versions, FEI training version, and versions of basf2 used for each part of the analysis.

each for B^+ and B^0. The treatment of these samples will be covered in Section 5.2.1. No dedicated signal MC sample was used for the normalisation mode, as the expected number of events in 500 fb^{-1} of generic MC is comparable to the number of events in the available signal MC samples. All MC samples were produced with the nominal level of beam background applied via overlay files. The dataset labels used internally by the Belle II collaboration are listed in Table 5.1.

All the above datasets were skimmed with the SL FEI skims[2]. Table 5.1 lists the FEI version and basf2 versions used for skimming and reconstruction. The differing versions used for skimming MC and data are not a concern: basf2 follows the Semantic Versioning scheme, meaning substantial changes to the reconstruction software can only be made across major version changes (*i.e.* the first number in the version triplet). A `light` release is used for reconstruction, as this is checkpoint on the main basf2 branch, and so contains the most up-to-date analysis utilities.

2: The skim selections will be covered in detail in Section 5.3.2.

The detector data was collected with the hadron HLT selections in place, although these selections are sufficiently loose that they have almost no impact on the events which pass the FEI skim. For this same reason, the hadron HLT selections are not applied to the MC samples; the systematic uncertainty resulting from this decision is expected to be far smaller than the dominant sources of uncertainty.

From here on, the detector data and simulated data will be referred to simply as "data" and "MC", respectively.

5.2.1. Hybrid MC reweighting

The two varieties of $B \to X_u \ell \nu$ signal samples were used: resonant and non-resonant (also referred to as exclusive and inclusive, respectively). "Resonant" refers to the decays $B \to h \ell \nu$, where h is a light charmless hadron (π , ρ , η , ω). These are well measured decays, but they only make up 20% of the $B \to X_u \ell \nu$ branching fraction. "Non-resonant" refers to inclusive $B \to X_u \ell \nu$ decays, where the kinematics of X_u are parametrised to model a large number of decays with very small branching fractions which have not been precisely measured yet. The non-resonant samples used in this analysis were generated using the BLNP scheme for modelling inclusive decays [112].

[112]: Lange et al. (2005), "Theory of charmless inclusive B decays and the extraction of V_{ub}"

These two components are combined using the "hybrid" approach, which was first proposed in Reference [113]. In this method, we reweight the inclusive component in theoretically motivated bins of the following three variables:

- the invariant mass of the hadronic system, M_X,
- the 4-momentum transfer, q^2, and
- the lepton energy in the B frame, E_ℓ.

The event weights are adjusted so that the distributions of the combined samples match the operator production expansion expectation, as well as correctly matching the total $B \to X_u \ell \nu$ branching fraction.

This reweighting was performed using the EFFORT package [114]. We first calculate a table of weights using the generator level information of unskimmed MC samples. We then apply the weights to the inclusive samples which will be used in later steps of the analysis. The impact of hybrid reweighting on the M_X distribution is shown in Figure 5.1.

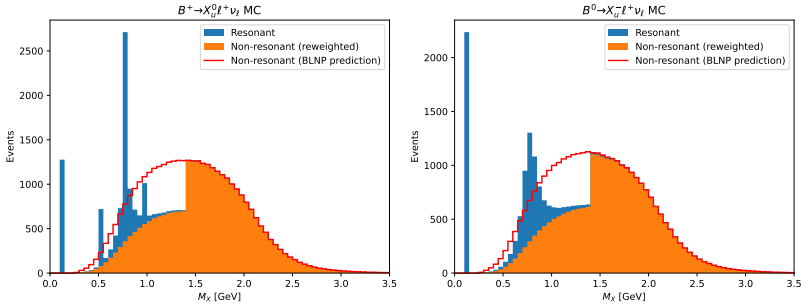


Figure 5.1: Impact of the hybrid reweighting on inclusive $B \to X_u \ell \nu$ MC. The quantity plotted is the invariant mass of the hadronic system. The discontinuity in the orange distribution arises from the finite binning of the reweighting.

5.3. Reconstruction and selections

In this Section, we will detail the reconstruction and selection of the tag-side and signal-side B mesons for each of the signal and normalisation modes. As a point of notation, it is convention that whenever we refer to the reconstruction of a particular decay, the charge conjugate decay is also implied.

We use the B_{tag} candidates reconstructed by the SL FEI, and combine them with the reconstructed B_{sig} candidates to create $\Upsilon(4S)$ candidates. The $\Upsilon(4S)$ reconstruction is done separately for $\Upsilon(4S) \to B^+B^-$ (referred to in this Chapter as "the B^+ reconstruction"), and $\Upsilon(4S) \to B^0\overline{B}^0$ (referred to as "the B^0 reconstruction"). The B^0 reconstruction also includes the neutral B mixing cases, $\Upsilon(4S) \to B^0B^0$ and $\Upsilon(4S) \to \overline{B}^0\overline{B}^0$, which provides a moderate (~20%) improvement to efficiency.

[113]: Ramirez et al. (1990), "Semileptonic $b \to u$ decay"

[114]: Prim (2020), *eFFORT*

5.3.1. Blinding of signal region

It is standard practice in HEP to deliberately avoid looking at detector data in the signal region until the very end of an analysis. This procedure is known as *blinding*. An analyst may unknowingly bias their measurement if they have access to the signal events while they are in the process of developing their analysis selections. All studies of the signal region prior to unblinding must be performed with MC only.

In this analysis, we blind the signal mode reconstruction in the region $\cos^2\Phi_B < 5$. Hereafter, this region is referred to as the "signal region", and $\cos^2\Phi_B \geq 5$ as the "sideband region". No blinding is performed for the normalisation mode, as the primary purpose of reconstructing this mode is not to measure $\mathcal{B}(B \to D^*\ell\nu)$, but to provide a means of using the uncalibrated SL FEI.

5.3.2. Tag-side reconstruction and skim cuts

The tag-side B candidates are those reconstructed by the SL FEI. The FEI is run as a part of the skimming process, and the tag B particle lists are saved to the output uDST. This reduces the computational requirements of all later data processing, but means that certain selections are not able to be modified by the analyst. The FEI skims have three main parts, which will be detailed here: a set of "pre-cuts", tag candidate reconstruction, and post-reconstruction cuts. The full set of skim selections is summarised in Table 5.2.

Skim stage	Selections
Pre-cuts	$n_{\text{clean tracks}} \geq 3$
	$n_{\text{clean clusters}} \geq 3$
	$E_{\text{clean tracks \& clusters}} > 4\,\text{GeV}$
	$2\,\text{GeV} < E_{\text{clean tracks \& clusters in ECL}} < 7\,\text{GeV}$
B_{tag} cuts	$p^*_\ell > 1\,\text{GeV}$
	$\mathcal{P}_{\text{tag}} > 10^{-2.4}$
	$-4 < \cos\theta_{BY} < 3$

Table 5.2: Summary of selections made by SL FEI skim.

The pre-cuts are in place to reduce the number of events passed through the computationally expensive candidate reconstruction step. These are event-level selections calculated using dedicated "cleaned" particle lists. For cleaned tracks, the track parameters[3] must satisfy $\mathrm{d}r < 0.5\,\text{cm}$ and $|\mathrm{d}z| < 2\,\text{cm}$, and the transverse momentum must be at least 100 MeV. For cleaned clusters, the polar angle must be between 17° and 150°, and the energy must be at least 100 MeV. These are basic selections to remove tracks and clusters not originating from an $\Upsilon(4S)$ decay. The skim pre-cuts are as follows:

- the event must have at least three cleaned tracks and at least three cleaned clusters,

3: $\mathrm{d}z$ and $\mathrm{d}r$ are the longitudinal and radial distances (in cylindrical coordinates) of the point-of-closest-approach of the fitted track to the interaction point.

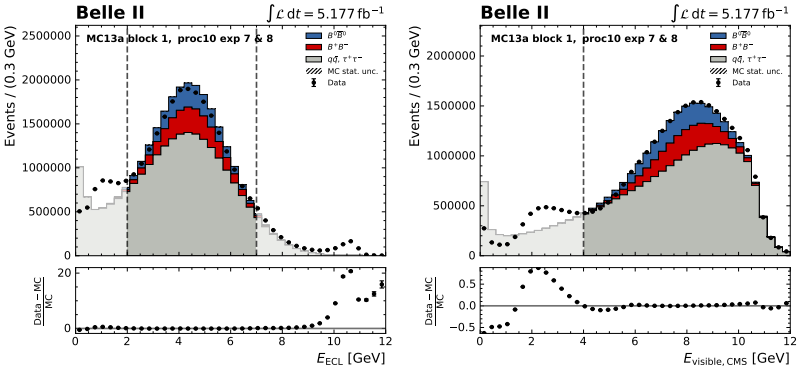


Figure 5.2: Plots of cuts on energy variables used in pre-skim selections (the "cleaned tracks and clusters" energy is here labelled as E_{visible}). The purpose of these cuts is to select $B\overline{B}$ events. The mismatches between data and MC outside of the selected ranges are due to low-multiplicity events not being included in the MC stack, and differences in trigger simulation for non-$B\overline{B}$ events.

- the sum of energy from the cleaned tracks and clusters (calculated after applying their respective mass hypotheses) must be at least 4 GeV, and
- the sum of the energy measured in the ECL must be between 2 GeV and 7 GeV.

These criteria are designed to select events which have a good chance of containing a pair of B mesons. Plots of the two energy variables are shown in Figure 5.2.

After the pre-cuts are applied, the FEI tag reconstruction is then run. Only the first four FEI tag modes are considered in this analysis; B^+_{tag} candidates are reconstructed in the modes $\overline{D}^0\ell^+$ and $\overline{D}^{*0}\ell^+$, and B^0_{tag} candidates in the modes $D^-\ell^+$ and $D^{*-}\ell^+$. Although the FEI was trained to reconstruct semileptonic D modes (where the primary B decay is purely hadronic), these modes were not included in the FEI skims, since they are comparatively rare, and computationally expensive to run. $B \to D^{(*)}\pi\ell$ modes are also not considered, as they are low purity and contribute only a ~10% increase to efficiency.

Finally, the SL FEI skims apply the following selection criteria to the tag candidates:

- The final classifier output, $\mathscr{P}_{\text{tag}}$, is required to be greater than $10^{-2.4}$.
- The lepton used in the tag reconstruction is required to have a momentum of at least 1.0 GeV in the CMS frame.
- The angle between the nominal and reconstructed B_{tag} flight direction, $\cos\theta_{BY}$, is required to be between −3 and 4. This is a relatively loose selection, given the properties of $\cos\theta_{BY}$ described in Section 3.6.2. This selection is not tightened later, as $\cos\theta_{BY}$ is an input to the signal mode fitting variable.

Since differences in the tag-side reconstruction between signal and normalisation modes may introduce biases in the efficiency ratio, the same tag reconstruction is used for both. This tag reconstruction is used for both the signal and normalisation modes. Differences in the tag-side

reconstruction may introduce biases in the efficiency ratio. The FEI produces multiple tag candidates for each event, so a candidate selection must be applied at some point. However, applying a candidate selection after the full $\Upsilon(4S)$ reconstruction is not the same as selecting a candidate before combining it with a signal-side. To maintain consistency between the tags of the signal and normalisation modes, a candidate selection is applied at this point. For each event, we keep only the tag candidate with the highest $\mathscr{P}_{\text{tag}}$.

Figure 5.3 shows the distributions of a few key variables after skim selections: the tag-side $\cos\theta_{BY}$, $\mathscr{P}_{\text{tag}}$, the tag lepton momentum in the CMS frame, the tag D mass, and the number of remaining tracks in the event (with the ROE mask detailed in Section 5.3.4 applied). The distribution of $\cos\theta_{BY}$ is highly sculpted in the MC14 FEI training, as can be seen in Figure 5.3. This is an issue that has been addressed in the most recent training, following the study detailed in Reference [115]. Furthermore, differences in the performance of the FEI on data and MC introduce slight shape differences which can be seen in Figure 5.3.

[115]: Röhrken (2021), "Retraining of the FEI: Unbiasing the $\cos\theta_{BD\ell}$ distributions"

Figure 5.3 also shows a disagreement in the lepton momentum above 3 GeV. Candidates with high lepton momentum originate from low multiplicity processes. These processes are not included in the MC stack, as these candidates do not pass the subsequent selection criteria. However, their absence does impact the data-MC agreement in Figure 5.3.

5.3.3. Signal-side reconstruction and cuts

The B_{sig} candidates for the signal mode are reconstructed as $\overline{B}^0_{\text{sig}} \to \pi^+\ell^-$ and $B^-_{\text{sig}} \to \pi^0\ell^-$. The B_{sig} candidates for the normalisation mode are reconstructed as $\overline{B}^0_{\text{sig}} \to D^{*+}\ell^-$ and $B^-_{\text{sig}} \to D^{*0}\ell^-$, with D^{*+} candidates reconstructed from $D^0\pi_s^+$ and D^{*0} from $D^0\pi_s^0$ (where the subscript s for "slow" denotes low momentum). D^0 candidates are reconstructed in the mode $K^-\pi^+$. Here we will describe the selections made on the final state particles, and the selections made throughout the B_{sig} reconstruction.

Final state particle and π^0 selections

The following selections are made on the tracks used in B_{sig} reconstruction. The track parameters must satisfy $\mathrm{d}r < 0.5\,\text{cm}$ and $|\mathrm{d}z| < 2\,\text{cm}$, the polar angle must be between 17° and 150°, and the number of hits in the CDC must be greater than 20—this final selection is not required for π_s^+. A momentum cut is placed on the charged leptons, requiring the momentum in the $\Upsilon(4S)$ frame to be greater than 0.3 GeV for electrons and 0.6 GeV for muons. Finally, selections are made on the appropriate particle ID classifier variable: `electronID` > 0.9 for electrons, `muonID` > 0.9 for muons, `pionID` > 0.6 for pions, and `kaonID` > 0.6 for kaons. The

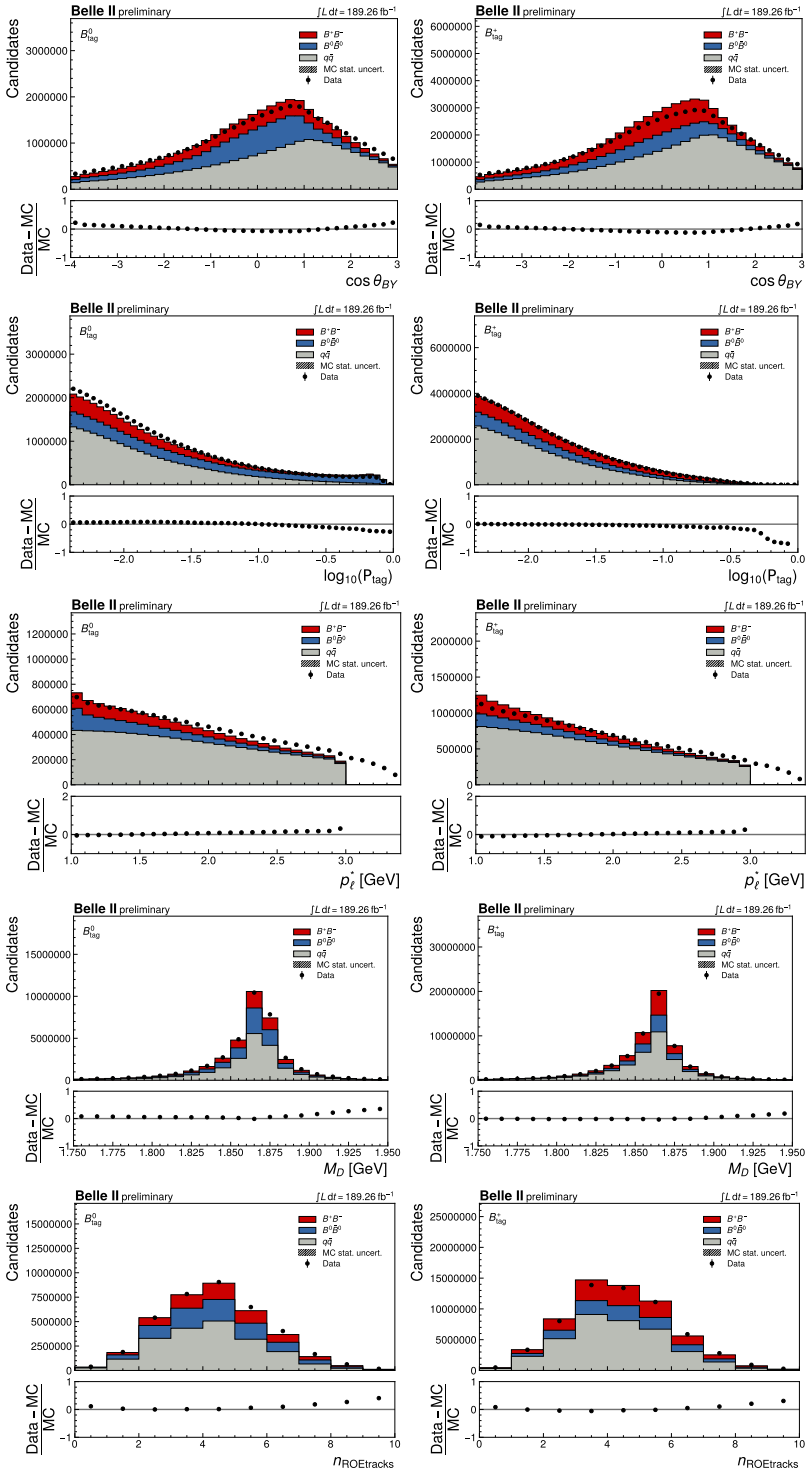


Figure 5.3: Plots of key distributions after skim selections but prior to any signal-side reconstruction. The plots are shown for the B^0 tag (left column) and the B^+ tag (right column).

charged PID variables exclude the information from the TOP and SVD systems, as the performance group discovered issues with these likelihoods. These selections are summarised in Table 5.3. Appendix B shows the distribution shapes of the variables used for charged track selection.

Table 5.3: Summary of selections made on charged particles used in $\Upsilon(4S)$ reconstruction.

Particle	Selections
Tracks	$\mathrm{d}r < 0.5\,\mathrm{cm}$
	$\lvert \mathrm{d}z \rvert < 2\,\mathrm{cm}$
	$17^\circ < \theta < 150^\circ$
	$n_{\text{CDC hits}} > 20$
e^+	Above track selections
	$p^* > 0.3\,\mathrm{GeV}$
	`electronID_noSVD_noTOP` > 0.9
μ^+	Above track selections
	$p^* > 0.6\,\mathrm{GeV}$
	`muonID` > 0.9
π^+	Above track selections
	`pionID_noSVD_noTOP` > 0.6
K^+	Above track selections
	`kaonID_noSVD_noTOP` > 0.6
π_s^+	Above track selections, except for $n_{\text{CDC hits}} > 20$

Signal-side neutral pions are reconstructed from photon pairs, using the standard particle list `pi0:eff40_May2020`. Each photon must satisfy:

- $E > 80\,\mathrm{MeV}$ if the cluster is in the forward region,
- $E > 30\,\mathrm{MeV}$ if the cluster is in the barrel, or
- $E > 60\,\mathrm{MeV}$ if the cluster is in the backward region.

The reconstructed pion is required to have a mass between 120 MeV and 145 MeV. These selections are summarised in Table 5.4.

Table 5.4: Summary of $\pi^0 \to \gamma\gamma$ reconstruction. Selections are those of the standard π^0 particle list `pi0:eff40_May2020`.

Particle	Selections
γ	$E > \begin{cases} 80\,\mathrm{MeV};\ \text{forward} \\ 30\,\mathrm{MeV};\ \text{barrel} \\ 60\,\mathrm{MeV};\ \text{backward} \end{cases}$
	$17^\circ < \theta < 150^\circ$
	$n_{\text{cluster hits}} > 1.5$
π^0	$120\,\mathrm{MeV} < M < 145\,\mathrm{MeV}$

The photons used in π_s^0 reconstruction must satisfy:

- $E > 25\,\mathrm{MeV}$ if the cluster is in the forward region,
- $E > 25\,\mathrm{MeV}$ if the cluster is in the barrel, or
- $E > 40\,\mathrm{MeV}$ if the cluster is in the backward region.

Belle II has developed a multivariate classifier of eleven Zernike moments of ECL clusters, which describe the shape of the cluster [116]; the classifier output for the photons is required to be greater than 0.3. The minimum distance from the cluster to the nearest track is required to be greater than 40 cm. The reconstructed π_s^0 is required to have a mass between 120 MeV and 145 MeV. These selections are summarised in Table 5.5.

[116]: Hershenhorn et al. (2017), *ECL Shower Shape Variables Based on Zernike Moments*

Table 5.5: Summary of $\pi_s^0 \to \gamma\gamma$ reconstruction.

Particle	Selections
γ	$E > \begin{cases} 25\,\text{MeV; forward} \\ 25\,\text{MeV; barrel} \\ 40\,\text{MeV; backward} \end{cases}$
	`clusterZernikeMVA` > 0.3
	distance to nearest track > 40 cm
π_s^0	$120\,\text{MeV} < M < 145\,\text{MeV}$

Signal mode B_{sig} selections

For the signal mode, the B_{sig} candidates are reconstructed from $\pi\ell$, and are required to have $|\cos\theta_{BY}| < 5$. This loose selection on $\cos\theta_{BY}$ is chosen because is an input to the fitting variable. The Fox-Wolfram moment R_2 is required to be less than 0.4, in order to reject continuum background. The ROE is then constructed with respect to the $\Upsilon(4S)$, and a mask is applied to the ROE tracks and clusters—the mask selections are shown in Section 5.3.4. $\Upsilon(4S)$ candidates are only kept if there are no remaining tracks in the ROE, and the sum of ROE neutral cluster energies is less than 0.3 GeV. These two cuts are highly effective at removing misreconstructed candidates. Table 5.6 summarises these selections.

Table 5.6: Summary of selections made in signal mode reconstruction $B_{\text{sig}} \to \pi\ell\nu$.

Particle	Selections
B_{sig}	$\lvert\cos\theta_{BY}\rvert < 5$
$\Upsilon(4S)$	$n_{\text{tracks,ROE}} = 0$
	$E_{\text{extra,ROE}} < 0.3\,\text{GeV}$
Event-level	$R_2 < 0.4$

Figures 5.4 and 5.5 show the "$N-1$ cut plots" of the variables in Table 5.6. In each plot, all cuts are applied *except* for the cut on the variable being plotted. These plots do not include any efficiency corrections on the MC.

Appendix D.1 includes data-MC comparison plots of kinematic variables in the signal mode sideband region, broken down by signal-side lepton flavour. These plots were used to validate the reconstruction.

Normalisation mode B_{sig} selections

D^0 candidates for the normalisation mode are reconstructed from $K^-\pi^+$. The candidate D^0 mass is required to be within 15 MeV of the nominal mass. D^{*0} and D^{*+} candidates are then reconstructed from $D^0\pi_s^0$ and $D^0\pi_s^+$. D^* candidates are only kept if the difference of the D and D^* masses is between 0.12 GeV and 0.17 GeV. The same signal mode selections on R_2 and number of ROE tracks are also used here. However, the $E_{\text{extra,ROE}}$ selection is loosened to 1 GeV, as the signal component is less sharply peaked for the normalisation mode. Additionally, no requirement is placed on $\cos\theta_{BY}$, as this variable is the chosen fitting variable for the normalisation mode. These selections are summarised in Table 5.7.

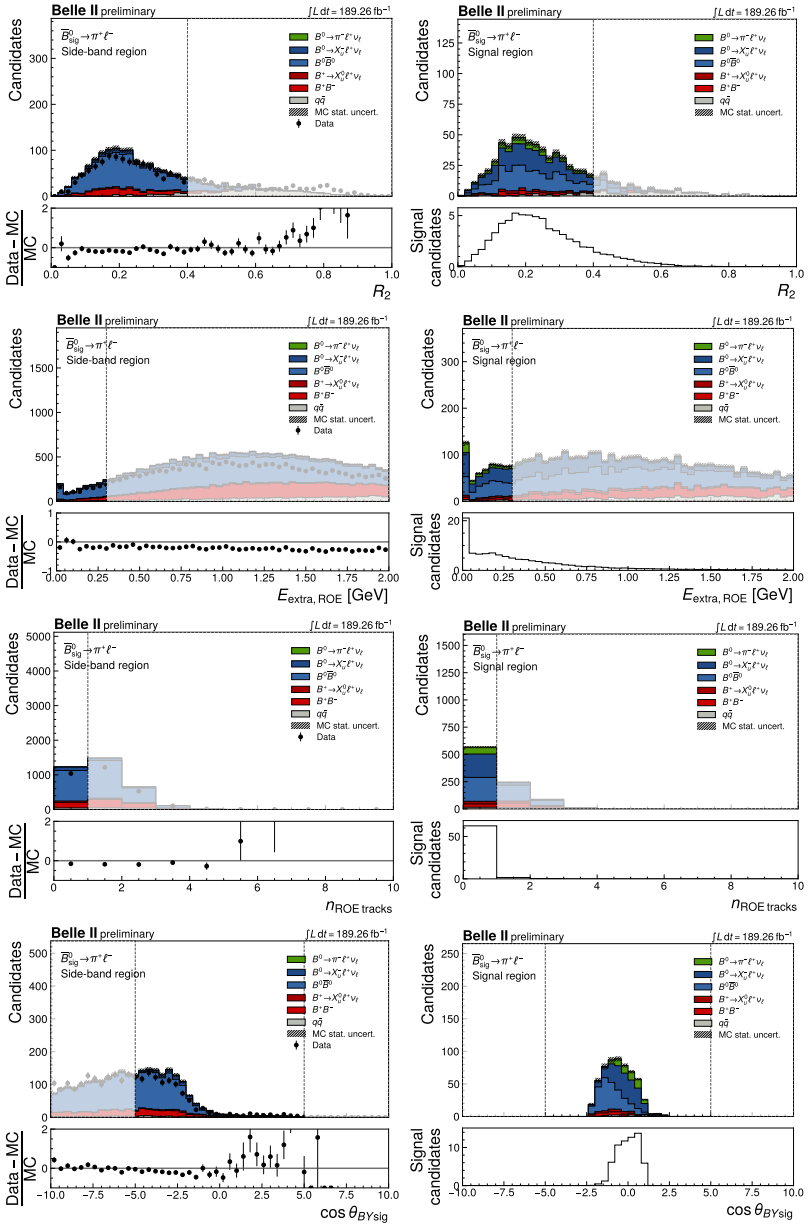


Figure 5.4: $N - 1$ plots of B^0 signal mode selections listed in Table 5.6. Selections are shown for the sideband region (left column) and signal region (right column).

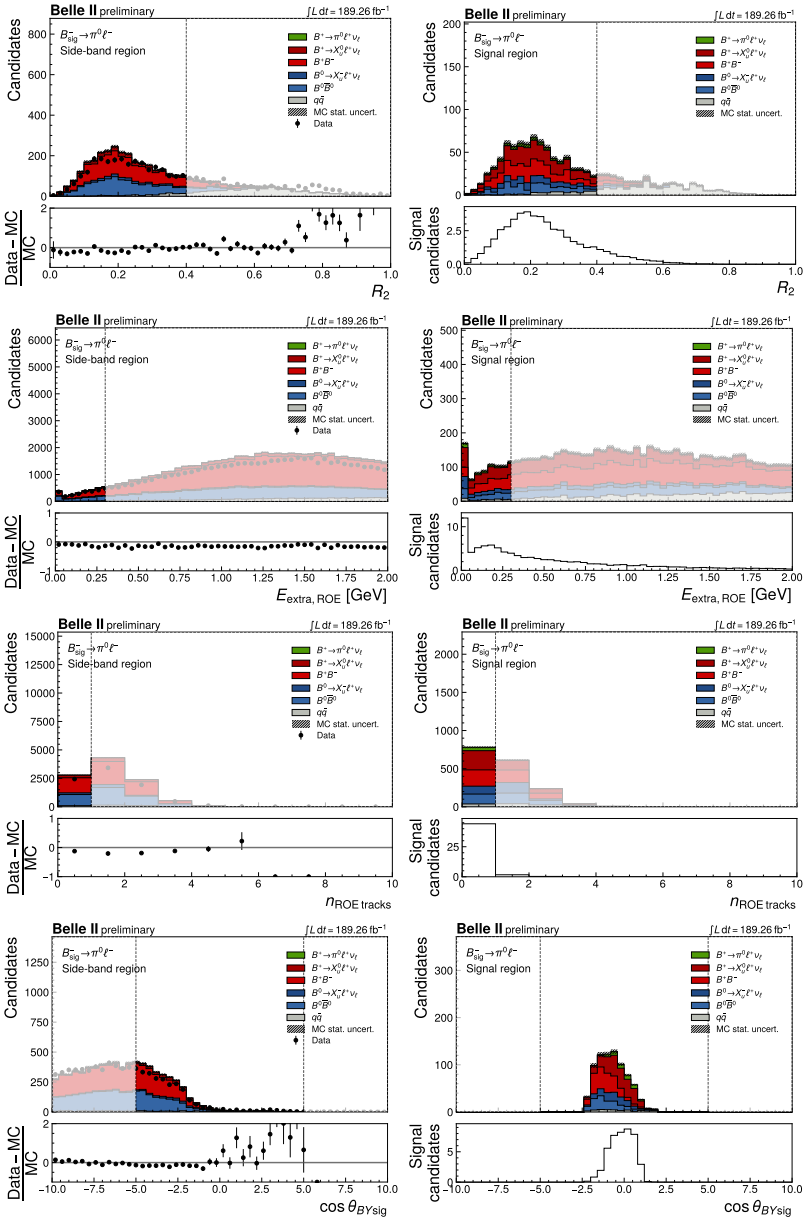


Figure 5.5: $N-1$ plots of B^+ signal mode selections listed in Table 5.6. Selections are shown for the sideband region (left column) and signal region (right column). The differences in data-MC normalisation are due to the absence of FEI calibration; this is addressed in later sections.

Particle	Selections
D^0	$\lvert M - M_{\text{PDG}}\rvert < 15\,\text{MeV}$
D^{*0}, D^{*+}	$0.12\,\text{GeV} < M_{D^*} - M_D < 0.17\,\text{GeV}$
$\Upsilon(4S)$	$n_{\text{tracks,ROE}} = 0$
	$E_{\text{extra,ROE}} < 1\,\text{GeV}$
Event-level	$R_2 < 0.4$

Table 5.7: Summary of selections made in normalisation mode reconstruction $B_{\text{sig}} \to D^* \ell \nu$.

Figure 5.6 shows the $N-1$ cut plots for the selections in Table 5.7. These plots do not include any efficiency corrections on the MC. The D mass plots show signal candidates outside selection region, but these are not of concern; these are mainly due to the looseness of the signal candidate definition explained in Section 5.4.3, which does not explicitly truth match the D reconstruction.

Appendix D.2 includes data-MC comparison plots of kinematic variables for the normalisation mode, broken down by signal-side lepton flavour. These plots were used to validate the reconstruction.

5.3.4. Rest-of-event definition

Having reconstructed the $\Upsilon(4S)$ candidates, several cuts are applied to the remaining objects in the event, termed the "rest-of-event" (ROE). To reduce the number of objects in the ROE not originating from one of the two B mesons, a *mask* is applied to all remaining tracks and clusters not associated with an $\Upsilon(4S)$ candidate. For tracks in the ROE, we require that the track parameters satisfy $\mathrm{d}r < 2\,\text{cm}$ and $|\mathrm{d}z| < 4\,\text{cm}$, and the transverse momentum be at least 200 MeV. For clusters in the ROE, we require that the energy be greater than 80 MeV, 30 MeV or 60 MeV, for the forward, barrel and backward regions, respectively. These region-dependent energy requirements are the same as those in the `gamma:eff40_May2020` list definition. We also apply a selection on two classifiers trained to identify beam background and non-photon ECL energy deposits: `beamBackgroundSuppression` and `hadronicSplitOffSuppression`. These requirements are summarised in Table 5.8. Appendix C shows the distribution shapes of the variables used for ROE photon selections.

ROE object	Selections
Charged tracks	$\mathrm{d}r < 2\,\text{cm}$
	$\lvert\mathrm{d}z\rvert < 4\,\text{cm}$
	$p_{\text{T}} > 200\,\text{MeV}$
Neutral clusters	$E > \begin{cases} 80\,\text{MeV};\ \text{forward} \\ 30\,\text{MeV};\ \text{barrel} \\ 60\,\text{MeV};\ \text{backward} \end{cases}$
	`beamBackgroundSuppression` > 0.1
	`hadronicSplitOffSuppression` > 0.1

Table 5.8: Summary of selections made on ROE objects.

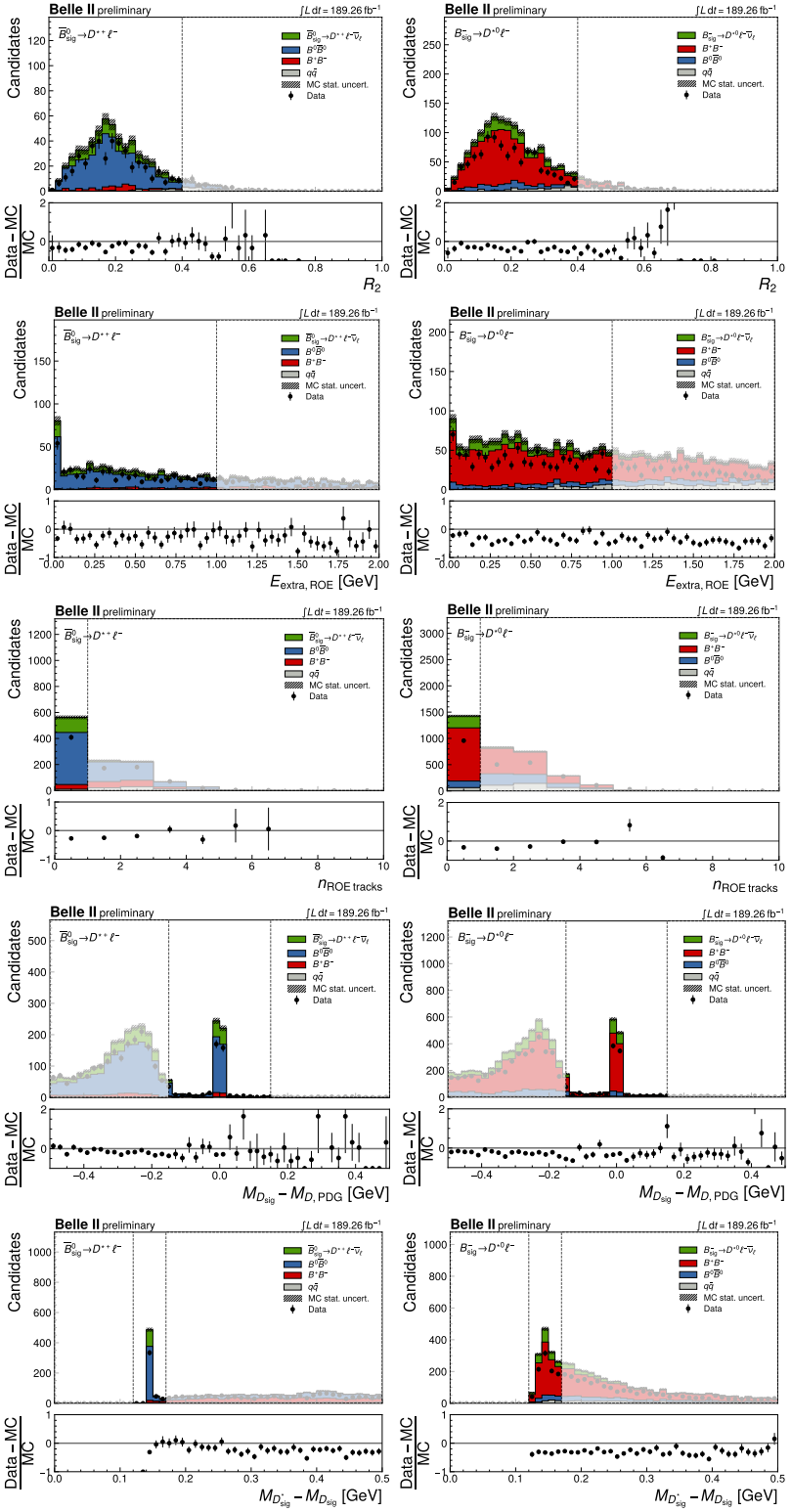


Figure 5.6: $N-1$ plots of B^0 (left column) and B^+ (right column) normalisation mode selections listed in Table 5.7.

5.3.5. Best-candidate selection

The above reconstruction procedure produces multiple candidates for each event, so we require some procedure for dealing with these [117]. In this analysis, we apply a candidate selection process, so that only one candidate is kept for each event.

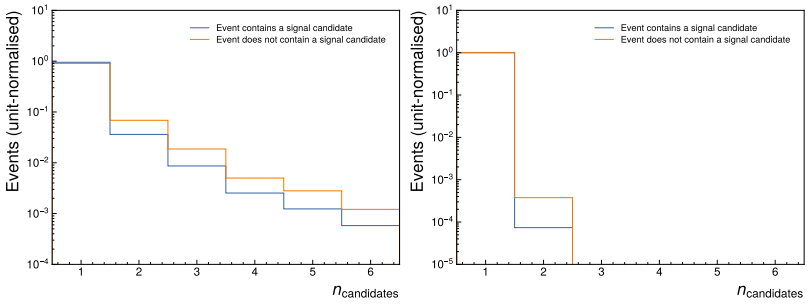
[117]: Koppenburg (2019), "Statistical biases in measurements with multiple candidates"

We described in Section 5.3.2 the first stage of the candidate selection, which keeps only a single B_{tag} candidate. This particular selection is necessitated by the normalisation mode approach taken in this analysis.

After the $\Upsilon(4S)$ reconstruction and selections, a single $\Upsilon(4S)$ candidate from each event is selected by choosing the candidate with the lowest $E_{\text{extra,ROE}}$. In the cases where multiple $\Upsilon(4S)$ candidates remain (*i.e.* two candidates have the same $E_{\text{extra,ROE}}$), then one candidate is randomly selected. Table 5.9 shows the proportion of events with multiple candidates throughout this procedure, and Figure 5.7 shows the candidate multiplicities broken down by whether the event contains a signal candidate.

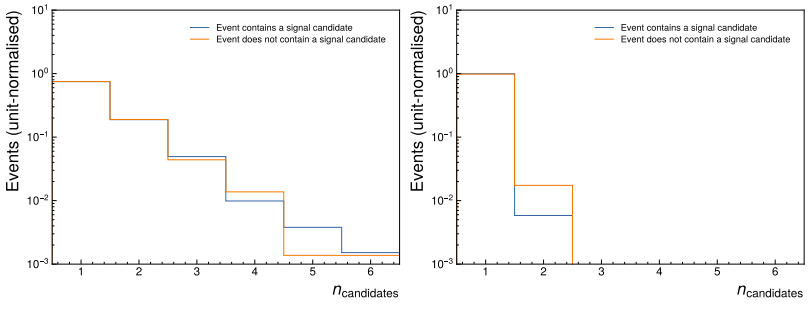
Mode	Before selection	After selection
$B^- \to \pi^0 \ell^+ \nu$	9.4%	1.1%
$B^0 \to \pi^- \ell^+ \nu$	0.02%	0.02%
$B^- \to D^{*0} \ell^+ \nu$	39.9%	3.5%
$B^0 \to D^{*-} \ell^+ \nu$	1.6%	1.6%

Table 5.9: Proportion of events with multiple $\Upsilon(4S)$ candidates before and after candidate selection using $E_{\text{extra,ROE}}$.



(a) $B^+ \to \pi^0 \ell^+ \nu$ reconstruction.

(b) $B^0 \to \pi^- \ell^+ \nu$ reconstruction.



(c) $B^+ \to D^{*0} \ell^+ \nu$ reconstruction.

(d) $B^0 \to D^{*-} \ell^+ \nu$ reconstruction.

Figure 5.7: Multiplicity of candidates for each of the reconstructions prior to $E_{\text{extra,ROE}}$ candidate selection. The distributions are broken down by whether the event contains a signl candidate, and are normalised so the shape can be compared.

The additional candidates in the $B^+ \to D^{*0} \ell^+ \nu$ reconstruction are due (in almost 100% of cases) to the π^0_s being substituted for some other

π_s^0 candidate. Likewise, the additional $B^+ \to \pi^0 \ell^+ \nu$ candidates arise from difference choices of π^0. Finally, the additional $B^0 \to D^{*-} \ell^+ \nu$ candidates are due to substituting the π_s^+ for another soft track. The final line of Table 5.9 shows that the selection on $E_{\text{extra,ROE}}$ does not impact the candidate multiplicity for $B^0 \to D^{*-} \ell^+ \nu$; these candidates only differ in the π_s^+ used to reconstruct the D^{*-}, but the ECL deposits from these pions are all below the ROE mask cuts, so the B^0 candidates have identical $E_{\text{extra,ROE}}$ values.

5.4. Additional data processing

5.4.1. MC normalisation

In preparation for fitting, the generic and signal MC samples must be normalised to the luminosity of data.

When generic MC is generated, the size of the samples is chosen to correspond with a particular luminosity. So the scaling of generic MC samples is simply $\mathscr{L}_{\text{data}}/\mathscr{L}_{\text{MC}}$.

Signal samples are not generated with a particular luminosity in mind, but are instead quoted as the number of generated events. Equation (5.1) governs the scaling of signal samples to a given data luminosity. The additional factor of two is present for the two B mesons.

$$\frac{\mathscr{L}_{\text{data}} \cdot \sigma(e^+e^- \to \Upsilon(4S)) \cdot \mathscr{B}(\Upsilon(4S) \to B\bar{B}) \cdot 2 \cdot \mathscr{B}(B \to \{\text{signal mode}\})}{N_{\text{generated events}}} \tag{5.1}$$

For example, the $B^+ \to \pi^0 \ell^+ \nu$ sample contains 50M events. The scale factor to be applied to the event weights is calculated in Equation (5.2). Here another factor of two is present for $\ell = e, \mu$.

$$\frac{(189.9\,\text{fb}^{-1}) \cdot (1.110\,\text{nb}) \cdot (0.514) \cdot 2 \cdot (2 \times 7.80 \times 10^{-5})}{50 \times 10^6} \tag{5.2}$$

5.4.2. Removal of signal events from generic samples

This analysis uses signal MC for the signal mode in addition to the generic $B\bar{B}$ MC. This is to reduce the statistical uncertainty on the shape of the signal $B \to \pi \ell \nu$ component, which has a very low branching fraction, and thus very few events in the generic samples. In order to replace these signal events in generic samples with higher-statistics samples, events must

```
(  # pi e nu_e in 0'th B
  (abs(genUps_d0_d0_PDG) == 211) and
  (abs(genUps_d0_d1_PDG) == 11) and
  (abs(genUps_d0_d2_PDG) == 12)
) or
(  # pi mu nu_mu in 0'th B
  (abs(genUps_d0_d0_PDG) == 211) and
  (abs(genUps_d0_d1_PDG) == 13) and
  (abs(genUps_d0_d2_PDG) == 14)
) or
(  # pi e nu_e in 1'th B
  (abs(genUps_d1_d0_PDG) == 211) and
  (abs(genUps_d1_d1_PDG) == 11) and
  (abs(genUps_d1_d2_PDG) == 12)
) or
(  # pi mu nu_mu in 1'th B
  (abs(genUps_d1_d0_PDG) == 211) and
  (abs(genUps_d1_d1_PDG) == 13) and
  (abs(genUps_d1_d2_PDG) == 14)
)
```

Code 14: Query matching $B^0 \to \pi^- \ell^+ \nu$ events. This query is used to remove signal events from generic MC. `genUps_di_dj_PDG` is an alias for the PDG ID of the j^{th} daughter of the i^{th} daughter of the generator-level $\Upsilon(4S)$.

be removed from the generic MC before further analysis. This removal was performed using the assigned particle codes of the generator-level $\Upsilon(4S)$ granddaughters[4]. For example, the query matching $B^0 \to \pi^- \ell^+ \nu$ events in either B is shown in Code 14.

4: These codes are systematised by the Particle Data Group, and so are often referred to as *PDG IDs*.

5.4.3. "Signal" event definition

In subsequent analysis, we require some criteria for identifying events in MC as "correctly reconstructed." The criteria was chosen to be consistent between the signal and normalisation modes.

The criteria for "correctly reconstructed" signal mode events is as follows. First, we check that the signal-side lepton is correctly truth-matched, and that its MC mother is a B meson. We then check the generator-level daughters of the lepton's MC mother: the first daughter must be a pion, and the fourth daughter must either be non-existent or a `PHOTOS` photon. No truth-matching criterion is applied to the tag-side.

The same criteria are used to identify correctly reconstructed candidates in the normalisation mode. The only difference is that the first daughter of the lepton's MC mother is required to be a D^*.

5.5. Systematic corrections and uncertainties

There are various ways in which aspects of the simulated data may disagree with the detector data. The final measurement relies on accurate

modelling of the detector data, so it is important to correct for these disagreements where possible, and to quantify their contribution to the uncertainty in the final result.

The systematic uncertainties in this analysis are of two main varieties: uncertainties due to correcting the mismodelling of the behaviour of objects in the detector, and uncertainties in the theory inputs to the simulated data. This Section will describe how these two varieties are handled. We will begin by explaining a general method for propagating the uncertainties of systematic corrections to the final measurement. We will also cover a side study investigating the possibility for bias in the efficiency ratio, which is crucial for ensuring the validity of the normalisation mode approach.

5.5.1. Procedure for estimating systematic uncertainty from correction tables

The performance group provides systematic corrections for individual physics objects. The values provided are tables of data-MC efficiency ratios, binned in detector regions and/or kinematic regions. For example, corrections for lepton ID are binned in momentum and polar angle. We apply these corrections to a particular event by multiplying the event weight by the appropriate table entry, according to the properties of the lepton for that event.

Each correction factor has an associated statistical and systematic uncertainty, and these must be propagated through the analysis in order to estimate their contribution to the systematic uncertainty of the final measurement. This is achieved by generating a set of Gaussian variations for the PID weights, and using these to produce a set of event weight variations. We will now describe how this procedure is performed.

We denote the vector of nominal efficiency corrections as $\boldsymbol{\epsilon}$, and the vector for the i^{th} variation as $\boldsymbol{\epsilon}'_i$. The variations are sampled from a multivariate normal distribution, as shown in Equation (5.3).

$$\boldsymbol{\epsilon}'_i = \boldsymbol{\epsilon} + \mathcal{N}(\mathbf{0}, \boldsymbol{\sigma}^{\mathsf{T}}_{\text{stat}} \rho_{\text{stat}} \boldsymbol{\sigma}_{\text{stat}})_i + \mathcal{N}(\mathbf{0}, \boldsymbol{\sigma}^{\mathsf{T}}_{\text{sys}} \rho_{\text{sys}} \boldsymbol{\sigma}_{\text{sys}})_i \tag{5.3}$$

$\boldsymbol{\sigma}_{\text{stat}}$ and $\boldsymbol{\sigma}_{\text{sys}}$ are the vectors of statistical and systematic uncertainties, and ρ_{stat} and ρ_{sys} are the corresponding correlation matrices. We assume that the statistical uncertainties are completely uncorrelated across all (p, θ), so that ρ_{stat} is the identity matrix. We assume that the systematic uncertainties are completely correlated, so that every entry of ρ_{sys} is 1. Some correction tables include separate up and down uncertainties,

but in this analysis we conservatively symmetrise them by choosing the larger of the two values.

Particles falling outside the coverage of the correction tables are also considered. If a particle is non-covered because it falls outside the binned range of the table, then it is assigned a weight of one, and the i^{th} variation is sampled from a 1D Gaussian,

$$\epsilon_i' = 1 + \mathcal{N}(0, \sigma_{\text{average}})_i \,, \tag{5.4}$$

where the average uncertainty across all bins, σ_{average}, is calculated as

$$\sigma_{\text{average}} = \sqrt{\bar{\boldsymbol{\sigma}}^2_{\text{stat}} + \bar{\boldsymbol{\sigma}}^2_{\text{sys}} + \overline{\boldsymbol{\epsilon} - 1}^2} \,. \tag{5.5}$$

This is a conservative approach to assigning uncertainties to non-covered particles; since Equation (5.4) is centred on unity rather than $\bar{\boldsymbol{\epsilon}}$, an additional factor of $\overline{\boldsymbol{\epsilon} - 1}^2$ is included Equation (5.5) to expand the uncertainty band.

The Gaussian variations on the correction tables are then multiplied into the event weights to produce weight variations. Where possible in this analysis, 200 weight variations were generated for each correction. These variations were used to estimate the shape uncertainty in the fitted templates, as well as the uncertainty in the yield ratios. In cases where the systematic uncertainty partially cancels in the ratio (for example, lepton ID), this procedure allows us to estimate the residual uncertainty.

5.5.2. Tracking momentum

Biases in the magnetic field map used in data reconstruction lead to biases in the measured momenta. To account for this, a global tracking scale factor of 0.99976 was applied online using the function `scaleTrackMomenta` from the `modularAnalysis` module. The scale factor was applied to tracks prior to signal-side reconstruction, and also to the B_{tag} candidate reconstructed by the FEI skims (the function updates the B_{tag} momentum as well as that of all the daughters).

Furthermore, a tracking uncertainty of 0.30% per signal-side track is assigned to MC. We assume this systematic uncertainty is fully correlated among all tracks, and so will partially cancel in the ratio.

5.5.3. Photon energy bias

Data-MC disagreements in the π^0 mass have been observed by the performance working group. These are due to the photon energy mis-calibration, and can be addressed by applying a bias-correcting factor to data. This

was done online using the `correctEnergyBias` function from the `modularAnalysis` module. This correction was applied to photon lists used in signal side reconstruction, including the photons used to reconstruct the signal-side π^0. This correction was also applied to the photon lists used in the ROE definition.

5.5.4. Particle ID corrections

The particle ID variables have slightly different efficiencies on data and MC, so cutting on these variables introduces differences between data and MC distributions. The PID working group provides tables of efficiencies and fake rates for particular *working points* (*i.e.* cut points on a PID variable). These correction tables are binned in momentum and polar angle. The method described in Section 5.5.1 is used to reweight the MC to better match the data. The treatment of non-covered particles is extended in this case; if a particle is faked by another particle for which there is no fake rate table, then it is assigned a weight of one and all variations are also one.

[118]: Sutcliffe (2021), *PIDvar*

All corrections were applied offline using the `PIDvar` package [118]. Corrections were applied to the following particles:

- Lepton ID corrections were applied to the signal-side lepton for all reconstructions. The working point is 0.9.
- Kaon ID corrections were applied to the signal-side kaon used in D^* reconstruction for the normalisation mode. The working point is 0.6.
- Pion ID corrections were applied to the signal-side pion for the neutral B signal mode, and to the pion used in D^* reconstruction for the normalisation mode. The working point is 0.6.

The lepton ID tables were provided by the performance group in CSV format, and kaon and pion ID tables were generated using the Systematic Corrections Framework [119]. No PID corrections are applied to the tag-side, for reasons discussed in Section 5.5.8.

[119]: Bilokin (2019), *Systematic Corrections Framework*

Lepton ID corrections are derived using several control channels, in order to cover different momentum regions [120].

[120]: Belle II Lepton ID Group (2021), *Muon and electron identification performance with* 189 fb^{-1} *of Belle II data*

- The mid-momentum region ($1\,\mathrm{GeV} \lesssim p_{\mathrm{lab}} \lesssim 2.5\,\mathrm{GeV}$) is covered by inclusively reconstructed $J/\psi \to \ell^+\ell^-$ decays. One of the two tracks is required to have electronID or muonID greater than 0.9. A vertex fit is performed, and only $J/\psi \to \ell^+\ell^-$ candidates passing the fit are retained. The samples are further cleaned with continuum suppression selections and bremsstrahlung corrections. The number of $J/\psi \to \ell^+\ell^-$ candidates in data and MC are obtained through fits to the dilepton invariant mass, and the binned correction factors are calculated through this.

- The region $p_{\text{lab}} > 0.4\,\text{GeV}$ is studied through the two-photon process $e^+e^- \to (e^+e^-)\ell^+\ell^-$, which was described in Section 3.3.5 in the context of beam background. In most cases, the two electrons continue down the beam pipe, leaving at most two leptons to interact with the detector. The major backgrounds are other two-photon production processes such as $e^+e^- \to (e^+e^-)\pi^+\pi^-$. In order to suppress these backgrounds, a tag-and-probe approach is used, in which tight selections are made on one lepton, and the other is used to calculate the efficiency. Tag electrons are required to have electronID > 0.95, and tag muons are required to have muonID > 0.95 and $p_{\text{lab}} > 0.7\,\text{GeV}$. The efficiency in data is measured as the ratio of non-background candidates after tag selection and after probe selection; in this approach, the number of background candidates is small and may be estimated from MC. The MC efficiency is obtained directly from truth-level information.
- Finally, the efficiency of high-momentum leptons ($p_{\text{lab}} < 6.5\,\text{GeV}$) is studied through radiative Bhabha events ($e^+e^- \to e^+e^-(\gamma)$) and radiate dimuon events ($e^+e^- \to \mu^+\mu^-(\gamma)$). A similar tag-and-probe approach is also used here, resulting in highly pure control samples.

Kaon and pion ID corrections are derived through the inclusive study of $D^{*+} \to [D^0 \to K^-\pi^+]\pi^+$ decays [121]. In this study, D^0 candidates are reconstructed from oppositely charged tracks with the kaon and pion mass hypotheses. This is combined with a slow pion to form a D^{*+} candidate. The control sample is selected with a momentum cut on the D^{*+} candidates, and on the mass difference between D^{*+} and D^0. Fits are then performed to the D^0 mass distributions for data and MC (binned in polar angle and momentum), from which the efficiency and fake rates are calculated.

[121]: Sandilya et al. (2019), *Study of Kaon and Pion Identification Performances in Phase III data with D^{*+} sample*

5.5.5. Neutral pion efficiency

The signal-side reconstruction of the charged B signal mode contains a π^0, and there are known efficiency differences in π^0 reconstruction in data and MC. To correct for this, we apply the table of correction factors shown in Table 5.10. The same procedure of Gaussian weight variations described in Section 5.5.1 is used here to propagate the uncertainty of these corrections to the final measurement.

The efficiency is calculated through the measurement of $D^0 \to K^-\pi^+$ and $D^0 \to K^-\pi^+\pi^0$ decays [122]. In order to increase the purity of the sample, D^0 candidates are only retained if they may be combined with a slow π^+ to form a D^{*+} candidate which passes cuts on $M_{D^*} - M_D$ and p_{D^*}. The D^0 mass distribution in data and MC are fitted individually for the

[122]: Koga (2020), *Momentum dependent π^0 efficiency measurement with D decays*

two decay channels. The π^0 efficiency correction factors are then derived from a double ratio of the signal yields.

Table 5.10: Efficiency correction factors for π^0. Values taken from Table XIV of Reference [122].

Momentum range (GeV)	$\epsilon_{\text{data}}/\epsilon_{\text{MC}} \pm$ stat. $\pm$ sys.
0.2–0.4	$0.960 \pm 0.012 \pm 0.060$
0.4–0.6	$0.985 \pm 0.009 \pm 0.041$
0.6–0.8	$1.032 \pm 0.009 \pm 0.052$
0.8–1.0	$1.033 \pm 0.010 \pm 0.041$
1.0–1.5	$1.052 \pm 0.008 \pm 0.046$
1.5–2.0	$1.060 \pm 0.011 \pm 0.046$
2.0–3.0	$1.023 \pm 0.011 \pm 0.047$

5.5.6. Slow pion efficiency

Differences between the modelling of slow pions and the actual detector response leads to efficiency differences between MC and data. The values listed in Tables 5.11 and 5.12 were used to correct these efficiency differences and estimate the associated systematic uncertainty. The same procedure of Gaussian weight variations described in Section 5.5.1 is used here to propagate the uncertainty of these corrections. The π_s^0 uncertainties were treated as completely uncorrelated with the π^0 efficiency corrections described in Section 5.5.5.

[123]: Lyu (2022), "Slow π Tracking Efficiency"

The slow charged pion correction factors are calculated through the inclusive reconstruction of $B^0 \to D^{*-}[\to D^0\pi_s^-]\pi^+$ decays [123]. D^0 candidates are reconstructed in modes $K^-\pi^+$, $K^-\pi^+\pi^+\pi^-$, and $K_S^0\pi^+\pi^-$. The number of signal candidates is extracted through a binned maximum likelihood fit to the ΔE distribution. This fit is performed in three bins of slow pion momentum. The slow neutral pion efficiency correction is studied through inclusive reconstruction of $B^+ \to D^{*0}[\to D^0[\to K^-\pi^+]\pi_s^0]\pi^+$ decays [122]. The number of signal candidates is extracted through a fit to the PDF of the mass difference of the D^{*0} and D^0. The number of signal candidates in data and MC may then be used to determine the efficiency correction.

[122]: Koga (2020), *Momentum dependent π^0 efficiency measurement with D decays*

Table 5.11: Efficiency correction factors for π_s^0. Values taken from Table XXV of Reference [122].

Momentum range (GeV)	$\epsilon_{\text{data}}/\epsilon_{\text{MC}} \pm$ stat. $\pm$ sys.
0.05–0.2	$1.020 \pm 0.053 \pm 0.052$

Table 5.12: Efficiency correction factors for π_s^+. Values taken from Reference [123].

Momentum range (GeV)	$\epsilon_{\text{data}}/\epsilon_{\text{MC}} \pm$ stat. $\pm$ sys.
0.05–0.12	$0.909 \pm 0.032 \pm 0.020$
0.12–0.16	$1.033 \pm 0.026 \pm 0.022$
0.16–0.20	$0.972 \pm 0.027 \pm 0.021$

5.5.7. $D^{(*)}$ form factor uncertainties

When generating $B \to D^{(*)}\ell\nu$ decays, there are certain assumptions made about the parametrisation of the form factors. The theoretical uncertainty inherent in these assumptions may be taken into account by reweighting the MC and generating weight variations.

EvtGen uses the BGL parametrisation [124] to simulate $B \to D^*\ell\nu$ decays. The parameters passed to the generator were from version 1 of Reference [125]. We reweighted the generic $B\overline{B}$ MC to the most recent parameters from version 3 of this paper, using the eFFORT package [114].

[124]: Boyd et al. (1995), "Constraints on Form Factors For Exclusive Semileptonic Heavy to Light Meson Decays"
[125]: Ferlewicz et al. (2021), "Revisiting fits to $B^0 \to D^{*-}\ell^+\nu_\ell$ to measure $|V_{cb}|$ with novel methods and preliminary LQCD data at non-zero recoil"
[114]: Prim (2020), *eFFORT*

The differential decay rate of $B \to D^*\ell\nu$ is governed by four variables:

- w: the velocity transfer from the initial to final state, calculated as

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

- θ_ℓ: the angle between the direction of lepton and the direction of the B meson in the virtual W rest frame.
- θ_V: the angle between the direction of the D meson and the direction of the B meson, in the D^* rest frame.
- χ: the angle between the two planes formed by the decays of the W and D meson, defined in the B rest frame.

Given two sets of parameters, eFFORT adjusts each event weight according to its location in the phase space of the above variables. The reweighting is performed such that the total decay rate is held fixed, and only the shape of the differential rate is adjusted. The formula used is shown in Equation (5.6). For each set of parameters, the differential decay rate is calculated using the BGL parametrisation, and the total rate is found by integrating over the above variables.

$$\text{weight} = \frac{\Gamma_{\text{old}}}{\Gamma_{\text{new}}} \frac{\frac{\mathrm{d}^4\Gamma_{\text{new}}}{\mathrm{d}w\,\mathrm{d}\cos\theta_l\,\mathrm{d}\cos\theta_V\,\mathrm{d}\chi}}{\frac{\mathrm{d}^4\Gamma_{\text{old}}}{\mathrm{d}w\,\mathrm{d}\cos\theta_l\,\mathrm{d}\cos\theta_V\,\mathrm{d}\chi}} \tag{5.6}$$

This process is an adjustment of the underlying distributions, not a correction of any reconstruction effect. Accordingly, generator-level information was used for all steps of this reweighting. The four kinematic variables were calculated using the generator-level 4-momenta, and events containing a $B \to D^*\ell\nu$ decay in one or both B mesons were identified using the generator-level PDG ID values.

Equation (5.6) may also be used to produce weight variations in a similar manner to that described in Section 5.5.1. We first generate 200

variations on the BGL parameters by sampling from a multivariate Gaussian, with the parameter covariance matrix calculated as the sum of the statistical and systematic covariance matrices. The i^{th} weight variation for an event is then given by Equation (5.7).

$$\text{weight}_i = \frac{\Gamma_{\text{new}}}{\Gamma_i} \frac{\frac{\mathrm{d}^4\Gamma_i}{\mathrm{d}w\,\mathrm{d}\cos\theta_l\,\mathrm{d}\cos\theta_V\,\mathrm{d}\chi}}{\frac{\mathrm{d}^4\Gamma_{\text{new}}}{\mathrm{d}w\,\mathrm{d}\cos\theta_l\,\mathrm{d}\cos\theta_V\,\mathrm{d}\chi}} \tag{5.7}$$

For $B \to D\ell\nu$ decays, the most recent form factor parameters are those used by EvtGen, so the MC did not need to be reweighted to incorporate new values. However, we did generate 200 variations in order to estimate the systematic uncertainty. A similar expression to Equation (5.7) applies here, except that the differential decay rate is governed only by w. The central values are taken from Reference [126], and the uncertainties and correlation matrix were derived from a fit [127].

[126]: Belle Collaboration (2016), "Measurement of the decay $B \to D\ell\nu_\ell$ in fully reconstructed events and determination of the Cabibbo-Kobayashi-Maskawa matrix element $|V_{cb}|$"

[127]: Bernlochner, *Private Communication*

The central values, uncertainties and correlation matrices of all parameters used in this Section are given in Appendices A.2 and A.3.

5.5.8. Testing for bias in efficiency ratio due to FEI tagging

We chose the normalisation approach in this analysis to enable the use of the uncalibrated SL FEI. Here we explain why the FEI needs calibration, and describe a study performed to validate the normalisation approach.

Differences in FEI performance on data and MC can arise in various ways. These include mismodelling of input variables (including PID variables), mismodelling of candidate multiplicity, and theoretical input assumptions in the generated MC. These discrepancies are further complicated by the structure of the algorithm, as each classifier is an input to later classifiers. As a result, the mismodelling of input variables cannot be addressed by the per-object efficiency correction table approach used earlier in this Section.

The approach taken for the hadronic FEI is to calibrate the global efficiency of the tagging. The calibration procedure is detailed in Reference [128]. The FEI calibration is currently statistically-limited, which restricts the granularity of the correction factors—the present calibrations provide only two numbers, one for B^0 tags and the other for B^+ tags. Regardless, the calibrated hadronic FEI has been successfully used to measure a variety of signal-side decays, including $B \to \pi\ell\nu$, $B \to \rho\ell\nu$, $B \to X_s\gamma$, and $B \to D^*\ell\nu$. The calibration procedure gives rise to a systematic uncertainty contribution on the order of 3%.

[128]: Belle II Collaboration (2020), *A calibration of the Belle II hadronic tag-side reconstruction algorithm with $B \to X\ell\nu$ decays*

The SL FEI has not seen the same success regarding calibration. Reference [96] details a calibration performed using an inclusive semileptonic signal side, $B \to X\ell\nu$. The correction tables are binned by tag mode and across $\mathscr{P}_{\text{tag}}$, since the $\mathscr{P}_{\text{tag}}$ distributions of each classifier are not necessarily related. However, independent calibrations performed with different signal sides have yielded contradictory correction factors. For reasons not yet fully understood by the collaboration, the SL FEI is more sensitive than the hadronic FEI to changes in the signal side. Two quantities of the signal-side which have a large impact are the track multiplicity and the lepton flavour.

[96]: Alina Manthei et al. (2021), *A first calibration of the Belle II semileptonic tag-side reconstruction algorithm using $B \to X\ell\nu$ decays with 62.8 fb^{-1}*

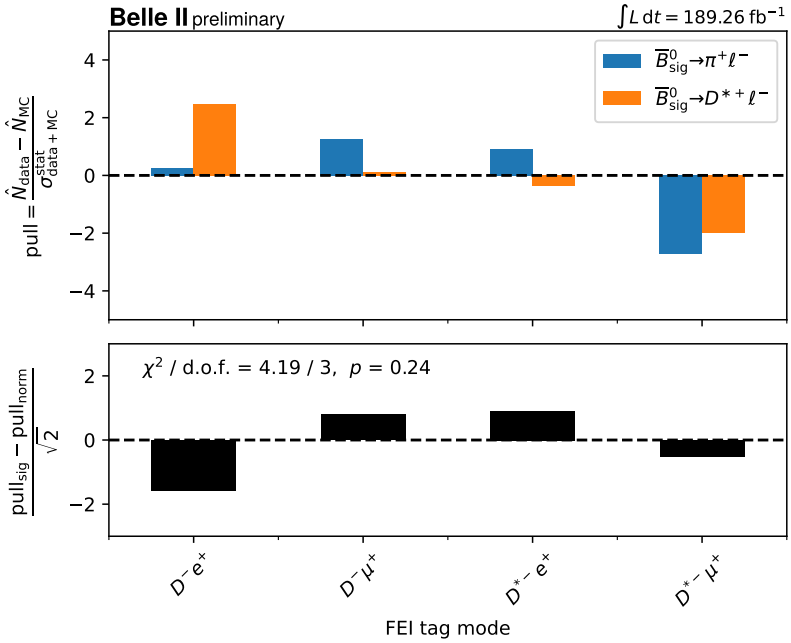
The normalisation approach is a compromise that bypasses the need for calibration. Rather than attempting to provide a calibration that is robust to *all* possible signal sides, we only need to check the compatibility of two signal-side modes. The ratio measurement is not impacted by absolute efficiency errors, as these are shared by the numerator and denominator. However, the measurement may be biased by *differing* efficiency errors in the signal and normalisation modes, provided that these differences are mismodelled. If the differences in MC are the same as the differences in data (up to statistical precision), then we can conclude that the ratio is not significantly biased by the FEI tagging.

To test this, we compared the fraction of events reconstructed in each tag mode, using the strategy detailed in Reference [129]. First we select signal-like events by requiring $\cos^2\Phi_B < 5$ for the signal mode. We normalise the distributions of tag modes, and then record the fraction of events in each mode for data and MC, $\hat{N}_{\text{data}}$ and $\hat{N}_{\text{MC}}$. The per-channel pulls are calculated as the difference between $\hat{N}_{\text{data}}$ and $\hat{N}_{\text{MC}}$ divided by the total statistical uncertainty, $\sigma^{\text{stat}}_{\text{data+MC}} = \sqrt{\left(\sigma^{\text{stat}}_{\text{data}}\right)^2 + \left(\sigma^{\text{stat}}_{\text{MC}}\right)^2}$. This quantifies the difference between data and MC broken down by tag mode for the signal and normalisation modes. These pulls are shown in the upper panels of Figure 5.8.

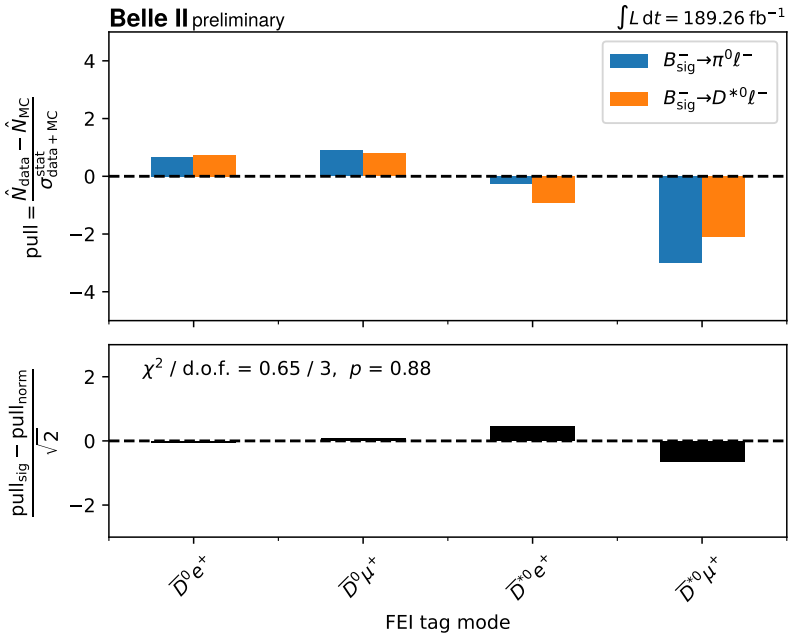
[129]: Bernlochner et al. (2022), *Tests of lepton universality in angular observables of hadronically tagged $B \to D^*\ell\nu$ decays at Belle II*

These first-order pulls being non-zero is not necessarily a problem—a biased ratio may only occur if the per-mode pulls are significantly different between the signal and normalisation modes. The second-order pulls are calculated as the difference of the first-order pulls divided by their combined statistical uncertainty, $\sqrt{2}$. These are shown in the lower panels of Figure 5.8.

The second-order pulls give χ^2 values of 4.19 and 0.65 for B^0 and B^+, respectively. The number of degrees of freedom is taken as $n_{\text{tag modes}} - 1$, giving p-values of 0.24 and 0.88. Because these are outside the threshold of statistical significance, we conclude that any potential error from mismodelling the tag-side efficiency is covered by the statistical uncertainty. Based on this, we do not assign any additional uncertainty to the final result.



(a) B^0 reconstruction.



(b) B^+ reconstruction.

Figure 5.8: Pulls for each tag mode for signal and normalisation modes. The upper panel of each figure shows the difference in the fraction of tags for signal and MC, and the lower panel shows the second-order pulls between signal and normalisation modes.

5.6. Template fit procedure

This Section details the first fitting step of the analysis. The most precise $|V_{ub}|$ measurements take advantage of the q^2 spectrum of $B \to \pi\ell\nu$. Accordingly, we measure the partial branching fractions of $B \to \pi\ell\nu$ in three q^2 bins, and normalise these to the total $B \to D^*\ell\nu$ branching fraction. The q^2 regions are $[0, 8)$, $[8, 16)$, and $[16, 26.4]$—these are referred hereafter to as "low," "mid," and "high." We perform template fits on the $\cos^2\Phi_B$ distribution of the signal mode and the $\cos\theta_{BY}$ distribution of the normalisation mode—a different variable is used for the normalisation mode, as there is very little discrimination between signal and background in $\cos^2\Phi_B$ for normalisation mode events. We extract event yields from the template fit results. We unfold the event yields of the signal mode to account for resolution and reconstruction effects which migrate events between q^2 bins. Finally, we calculate partial branching fraction ratios using the unfolded yields.

5.6.1. Notation for branching fraction ratios

We denote the partial branching fraction ratio in the i^{th} q^2 bin for the neutral and charged B cases as R_i^0 and R_i^-, as shown in Equations (5.8) and (5.9).

$$R_i^0 = \frac{\Delta\mathscr{B}_i(\overline{B}^0 \to \pi^+\ell^-\bar{\nu}_\ell)}{\mathscr{B}(\overline{B}^0 \to D^{*+}\ell^-\bar{\nu}_\ell)} \tag{5.8}$$

$$R_i^- = \frac{\Delta\mathscr{B}_i(B^- \to \pi^0\ell^-\bar{\nu}_\ell)}{\mathscr{B}(B^- \to D^{*0}\ell^-\bar{\nu}_\ell)} \tag{5.9}$$

R_i^0 and R_i^- can be inferred from the event yields and reconstruction efficiencies using Equations (5.10) and (5.11). In these equations, N and N_i denote the unfolded event yields, and ϵ and ϵ_i denote the signal efficiencies (the calculation of which is explained in Section 5.6.2).

$$R_i^0 = \frac{N_i(\pi^+\ell^-\bar{\nu}_\ell)\epsilon(D^{*+}\ell^-\bar{\nu}_\ell)}{N(D^{*+}\ell^-\bar{\nu}_\ell)\epsilon_i(\pi^+\ell^-\bar{\nu}_\ell)} \cdot \mathscr{B}(D^{*+} \to D^0\pi^+) \cdot \mathscr{B}(D^0 \to K^-\pi^+) \tag{5.10}$$

$$R_i^- = \frac{N_i(\pi^0\ell^-\bar{\nu}_\ell)\epsilon(D^{*0}\ell^-\bar{\nu}_\ell)}{N(D^{*0}\ell^-\bar{\nu}_\ell)\epsilon_i(\pi^0\ell^-\bar{\nu}_\ell)} \cdot \mathscr{B}(D^{*0} \to D^0\pi^0) \cdot \mathscr{B}(D^0 \to K^-\pi^+) \tag{5.11}$$

5.6.2. Efficiency calculation

In order to convert the signal yields to branching fraction ratios, we must calculate the reconstruction efficiency of signal events. We define the

signal efficiency as

$$\epsilon_{\text{reco}} = \frac{\text{number of signal events remaining after selections}}{\text{number of signal events generated}} . \tag{5.12}$$

The denominator is Equation (5.12) is calculated as the expected number of generated signal events in the sample, and the numerator is the number of events passing the "signal" criterion defined in Section 5.4.3. The numerator is calculated with all available efficiency corrections applied.

For the $B \to \pi \ell \nu$ reconstruction, a dedicated signal MC sample was used. This contained a cocktail of semileptonic decays involving charmless mesons. Each sample was generated with 50M events. Using the branching fractions in the `.dec` file, we expect 16.51×10^6 $B^0 \to \pi^- \ell^+ \nu$ decays and 9.375×10^6 $B^+ \to \pi^0 \ell^+ \nu$ decays. To determine the efficiency for each q^2 bin, ϵ_i, the expected number of generated events for each (generator-level) q^2 bin is calculated via the normalised partial branching fraction,

$$N_{\text{gen in } [q_1^2, q_2^2] \text{ bin}} = N_{\text{gen}} \times \frac{1}{\Gamma} \int_{q_1^2}^{q_2^2} \frac{\mathrm{d}\Gamma(B \to \pi \ell \nu)}{\mathrm{d}q^2} . \tag{5.13}$$

The numerator for each q^2 bin is the number of events that both satisfy the "signal" criterion and have a generator-level q^2 value in that bin.

For the $B \to D^* \ell \nu$ reconstruction, signal events were identified from the generic $B\overline{B}$ MC, so the numbers of generated signal events were calculated using Equations (5.14) and (5.15). A factor of two is included for the two lepton flavours, and another factor of two is included for the two $\mathscr{B}$ mesons.

$$\begin{aligned} N_{\text{gen}}(\overline{B}^0 \to D^{*+} \ell^- \bar{\nu}_\ell) &= N_{\text{gen}}(B^0 \overline{B}^0) \times 2 \times 2 \times \mathscr{B}(\overline{B}^0 \to D^{*+} \ell^- \bar{\nu}_\ell) \\ &\quad \times \mathscr{B}(D^{*+} \to D^0 \pi^+) \times \mathscr{B}(D^0 \to K^- \pi^+) \qquad (5.14) \\ &= 1.39 \times 10^6 \\ N_{\text{gen}}(B^- \to D^{*0} \ell^- \bar{\nu}_\ell) &= N_{\text{gen}}(B^+ B^-) \times 2 \times 2 \times \mathscr{B}(B^- \to D^{*0} \ell^- \bar{\nu}_\ell) \\ &\quad \times \mathscr{B}(D^{*0} \to D^0 \pi^0) \times \mathscr{B}(D^0 \to K^- \pi^+) \qquad (5.15) \\ &= 1.52 \times 10^6 \end{aligned}$$

$500\,\text{fb}^{-1}$ of generic MC was used, and the number of $B\overline{B}$ pairs per $100\,\text{fb}^{-1}$ is 51M for B^0 $\overline{B}^0$ and 54M for B^+ B^-. From this, we expect 5.027×10^6 $B^0 \to D^{*-} \ell^+ \nu$ decays and 6.899×10^6 $B^+ \to D^{*+} \ell^- \bar{\nu}_\ell$ decays.

Tables 5.13 and 5.14 summarise the number of generated signal events, number of passed events, and resulting efficiencies for each fit region.

	$\overline{B}^0 \to D^{*+}\ell^-$	$\overline{B}^0 \to \pi^+\ell^-$			
		All q^2	Low q^2	Mid q^2	High q^2
$N_{\text{total generated}}$	255M	50M	–	–	–
$N_{\text{signal generated}}$	1.39M	16.51M	6.29M	5.88M	4.35M
$N_{\text{signal reconstructed}}$	624	44503	14199	17955	12349
ϵ_{reco} (%)	0.04	0.27	0.23	0.31	0.28

Table 5.13: Values used to calculate signal efficiency for R^0 measurement.

	$B^- \to D^{*0}\ell^-$	$B^- \to \pi^0\ell^-$			
		All q^2	Low q^2	Mid q^2	High q^2
$N_{\text{total generated}}$	270M	50M	–	–	–
$N_{\text{signal generated}}$	1.52M	9.37M	3.55M	3.33M	2.49M
$N_{\text{signal reconstructed}}$	989	56049	18834	22253	14961
ϵ_{reco} (%)	0.07	0.60	0.53	0.67	0.60

Table 5.14: Values used to calculate signal efficiency for R^- measurement.

5.6.3. Template fit model

To extract the signal yields from the reconstructed detector data, we use a binned maximum likelihood fit using MC templates [130]. This method is a useful alternative to the common approach of finding an appropriate analytic form for the distribution of each event type. Template fits may also be straightforwardly extended to incorporate systematic uncertainties which impact the shape of each template. In this analysis, we use BinFit [131] to perform the template fit.

[130]: Barlow et al. (1993), "Fitting using finite Monte Carlo samples"

[131]: Sutcliffe (2020), *BinFit*

The template fit model is defined by n_c components, binned in n_b bins. The free parameters of the model are:

- ν^j: the yield of j-type events, and
- θ_i^j: a bin parameter for j-type events in bin i.

We define p_i^j as the fraction of events in template j which are reconstructed in bin i. The bin parameters θ_i^j are nuisance parameters included to account for statistical and shape systematic uncertainty in each template. Varying the bin parameters varies the shape of the templates without changing their normalisation. The variations on p_i^j are given by

$$\tilde{p}_i^j = \frac{p_i^j(1+\delta_i^j\theta_i^j)}{\sum_k p_k^j(1+\delta_k^j\theta_k^j)}\,, \tag{5.16}$$

where δ_i^j are the 1σ uncertainties on the bin parameters. The yield of events in bin i is then given by

$$\nu_i^{\text{pred}} = \sum_{j=1}^{n_c} \nu^j \tilde{p}_i^j\,, \tag{5.17}$$

The matrix ρ_θ encodes the correlations between bin parameters arising from systematic shape uncertainties. For each source of systematic uncertainty that impacts the template shape (and not merely the efficiency), a covariance matrix is calculated using event-by-event weight

variations, and these are summed to produce the total covariance matrix ρ_θ.

The fit to data is performed by minimising the negative log likelihood defined in Equation (5.18). The first term arises from the assumption that the bin yields are Poisson-distributed. The second is a Gaussian constraint term on the bin parameters, using the calculated correlation matrix.

$$-\ln \mathscr{L} = \sum_{i=1}^{n_b} \left[\nu_i^{\text{pred}} - \nu_i^{\text{obs}} + \nu_i^{\text{obs}} \ln \frac{\nu_i^{\text{obs}}}{\nu_i^{\text{pred}}} \right] + \boldsymbol{\theta}^{\mathsf{T}} \rho_\theta^{-1} \boldsymbol{\theta} \,. \tag{5.18}$$

We perform a simultaneous fit over four regions for each of the charged and neutral B cases. The fit regions are: all $B \to D^* \ell \nu$ events fitted in $\cos\theta_{BY}$, and $B \to \pi \ell \nu$ events fitted in $\cos^2 \Phi_B$ for each of the q^2 regions. Two components are defined for each fit region—a signal component containing events passing the signal criterion, and a background component containing all other events. The $\cos\theta_{BY}$ binning was chosen as a regular binning from -12 to 2 with bin widths of 1. The $\cos^2 \Phi_B$ binning was chosen as a regular binning from 0 to 20 with bin widths of 1, to capture the shapes of the signal and background. The signal component is peaked at zero and is contained almost entirely within $0 \leq \cos^2 \Phi_B < 5$, whereas the background component is relatively flat and slowly drops off at $\cos^2 \Phi_B \sim 1000$. The upper and lower edges of this range were chosen because $\cos^2 \Phi_B$ cannot be less than zero (by construction), and events with large values of $\cos^2 \Phi_B$ may be considered kinematically extreme.

The simultaneous fit allows us to rigorously deal with shape systematic uncertainties when taking ratios of fitted yields. To explain how the simultaneous fit is achieved in the template fit model, we will take the simpler example of two fitting regions: region A with N_A bins, and region B with N_B bins. We first define an extended variable with $N_A + N_B$ bins. The template for region A is non-zero in the first N_A bins and zero in all other bins, and *vice-versa* for region B. The component yields of all regions are allowed to float independently. The covariance matrix of bin parameters now has dimension $(N_A + N_B) \times (N_A + N_B)$, and includes the correlations between the bin parameters of different fit regions. Consequently, when the fit is performed, the off-diagonal elements of the global covariance matrix lead to slight correlations between the fitted yields. These correlations may then be taken into account when we later calculate the ratio of yields.

5.6.4. Unfolding procedure

We extract the signal mode event yields in three q^2 bins, but q^2 has a finite resolution which causes migration of events between bins. To address

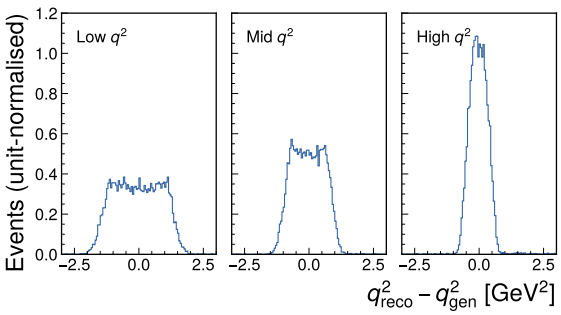


Figure 5.9: Resolution plots in each q^2_{gen} region for signal events of B^0 signal mode.

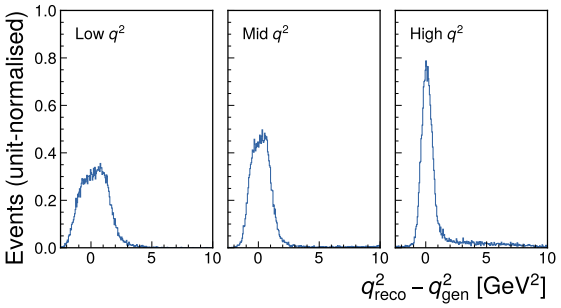


Figure 5.10: Resolution plots in each q^2_{gen} region for signal events of B^+ signal mode.

this, we must unfold the yields from the reconstructed q^2 bins to the generator-level q^2 bins. Here we will explain how q^2 is calculated, and detail the unfolding procedure.

In $B \to \pi\ell\nu$ decays, q^2 is defined as the invariant mass of the W boson,

$$q^2 = (p_B - p_\pi)^2 \,. \tag{5.19}$$

If we were using a hadronic tag, then we could use the 4-momentum of B_{tag} to infer that of B_{sig}. However, since we are using a SL tag, we do not have precise knowledge of the B_{sig} flight direction. We instead make the assumption that the B momentum in the CMS frame is negligible, giving the formula:

$$q^2_{\text{reco}} = (E^*_{\text{beam}} - E^*_\pi)^2 - (p^*_\pi)^2 \tag{5.20}$$

The generator-level quantity q^2_{gen} is calculated using the generator-level 4-momenta of the B and π. For this, we use the exact formula in Equation (5.19).

Figures 5.9 and 5.10 show the difference between the reconstructed and generator-level q^2 values for signal events passing all analysis selections. The plots are broken down by q^2_{gen} region. The q^2 resolutions in each region are calculated as half-width at half-maximum, and are shown in Table 5.15. We find that the assumption in Equation (5.20) leads to a q^2 resolution ranging from 0.42 to 1.45 GeV^2.

Table 5.15: Resolution (GeV^2) of q^2 in each q^2_{gen} bin. Calculated as the half-width at half-maximum of the distributions in Figures 5.9 and 5.10.

Signal mode	Low q^2	Mid q^2	High q^2
$B^0 \to \pi\ell\nu$	1.38	0.90	0.42
$B^+ \to \pi\ell\nu$	1.45	0.92	0.47

We unfold the yields fitted in q^2_{reco} bins into q^2_{gen} bins using a matrix inversion method. In order to do this, we calculate the migration matrix, M. M_{ij} is the conditional probability that an event generated in bin i is reconstructed in bin j. That is,

$$M_{ij} = P(q^2_{reco} \in b_j \mid q^2_{gen} \in b_i)\,. \tag{5.21}$$

The migration matrices for the B^0 and B^+ reconstructions are plotted in Figures 5.11 and 5.12. M relates the yields per q^2_{reco} bin (denoted here as $\boldsymbol{y}$) to the yields per q^2_{gen} bin (denoted here as $\boldsymbol{x}$) via

$$M\boldsymbol{x} = \boldsymbol{y}\,. \tag{5.22}$$

In the matrix inversion method, we unfold the fitted yields and covariance matrix using Equations (5.23) and (5.24).

$$\boldsymbol{x} = M^{-1}\boldsymbol{y} \tag{5.23}$$

$$V_x = M^{-1}V_y(M^{-1})^{\mathsf{T}} \tag{5.24}$$

Equation (5.24) allows us to propagate the uncertainties of the fitted yields to the unfolded yields. Unfolding via the matrix inversion method has the effect of slightly increasing the yield uncertainty, but in this analysis the increase is much smaller than the dominant uncertainty.

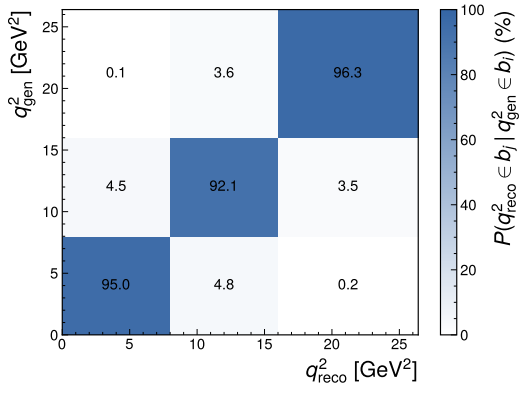


Figure 5.11: Migration matrices for signal events of B^0 signal mode.

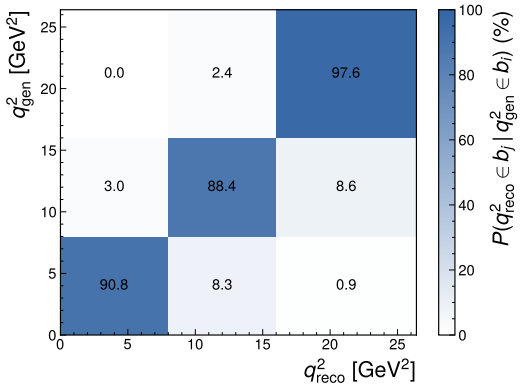


Figure 5.12: Migration matrices for signal events of B^+ signal mode.

5.7. $|V_{ub}|$ extraction procedure

In Section 5.6, we laid out the procedure for calculating the partial branching fraction ratios R_i^0 and R_i^- from the template fit results. Here we will explain how we fit these ratios to the q^2 spectrum of $B \to \pi\ell\nu$ to extract $|V_{ub}|$. This extraction is performed separately for the charged and neutral B reconstructions, as well as a combined fit to both.

Lattice QCD predictions for the q^2 spectrum of $B \to \pi\ell\nu$ decays are restricted by the wide range of possible q^2 values [28]. The statistical and discretisation errors may only be controlled in the upper third of the range[5]. The lattice QCD results may be extrapolated down to $q^2 = 0$, with increasing uncertainties toward the lower end of the range. Conversely, experimental measurements are most precise for low q^2 values, as this is where the differential branching fraction is greatest. One possible approach is to select a middle area which is accessible by both theory and experiment. The approach taken in this analysis is one used by many current analyses: first extract signal yields in separate q^2 bins, and then fit these to the predicted q^2 spectrum, while keeping control of the uncertainty induced by the choice of form factor parametrisation. $|V_{ub}|^2$ is then extracted from the normalisation of the fit result.

[28]: Aoki et al. (2019), "FLAG Review 2019"

5: $B \to D^{(*)}\ell\nu$ decays do not share this same difficulty, as their full kinematic range is reachable by lattice QCD simulations.

The $B \to \pi\ell\nu$ form factors are described by the BCL parametrisation [132]. In this analysis we use results from the Fermilab/MILC lattice collaboration in Reference [30], specifically the eight-parameter fit combining lattice and experimental data. We denote this parameter set as $\boldsymbol{\theta}_{\text{BCL}}$. Their central values, uncertainties, and correlation matrix are given in Appendix A.1. We include these as nuisance parameters in the fit, constraining them to their nominal values according to their covariance matrix, ρ_{BCL}.

[132]: Bourrely et al. (2010), "Model-independent description of $B \to \pi l\nu$ decays and a determination of $|V_{ub}|$"

[30]: Fermilab Lattice Collaboration et al. (2015), "$|V_{ub}|$ from $B \to \pi\ell\nu$ decays and (2+1)-flavor lattice QCD"

We define $\Delta\zeta_i$ in Equation (5.25) as the partial decay rate in the i^{th} q^2 bin divided by $|V_{ub}|^2$. By factoring out $|V_{ub}|^2$, $\Delta\zeta_i$ is determined entirely by theory inputs.

$$\Delta\zeta_i = \int_i \mathrm{d}q^2 \, \frac{1}{|V_{ub}|^2} \frac{\mathrm{d}\Gamma\,(B \to \pi\ell\nu)}{\mathrm{d}q^2} \tag{5.25}$$

[124]: Boyd et al. (1995), "Constraints on Form Factors For Exclusive Semileptonic Heavy to Light Meson Decays"

[133]: Caprini et al. (1998), "Dispersive Bounds on the Shape of $B \to D^{(*)}\ell\nu$ Form Factors"

The $B \to D^*\ell\nu$ decay rate can be described by the BGL [124] or CLN [133] parametrisations. However, because we normalise to the *total* $B \to D^*\ell\nu$ branching fraction, the shape is entirely integrated out and each parameter contributes only to the normalisation. Test fits performed with six BGL parameters were consistently unstable, most likely as a consequence of this collinearity of parameters. Instead, we include the $B \to D^*\ell\nu$ branching fraction as a parameter in the fit, denoting its nominal value and uncertainty as $\mathscr{B}_{D^*\ell\nu}$ and $\sigma_{D^*\ell\nu}$. Using the branching fraction as a parameter also avoids introducing model-dependence in the denominator. We use the current values from Reference [20]:

[20]: Particle Data Group (2022), "Review of Particle Physics"

- $\mathscr{B}(B^+ \to \overline{D}^{*0}\ell^+\nu_\ell) = (5.58 \pm 0.22)\%$
- $\mathscr{B}(B^0 \to \overline{D}^{*-}\ell^+\nu_\ell) = (4.97 \pm 0.12)\%$

The parameters of the fit model are: $|V_{ub}|$, the eight BCL parameters of the $B \to \pi\ell\nu$ form factors, and the $D^*\ell\nu$ branching fraction. Given these parameters, the predicted partial branching fraction ratio in the i^{th} q^2 region may be calculated using Equation (5.26). τ_B in this equation is the B lifetime.

$$R_i^{\text{pred}} = |V_{ub}|^2 \cdot \frac{\tau_B}{\mathscr{B}_{D^*\ell\nu}} \cdot \Delta\zeta_i(\boldsymbol{\theta}_{\text{BCL}}) \tag{5.26}$$

The χ^2 to be minimised is defined in Equation (5.27). $\boldsymbol{R}^{\text{meas}}$ and $\boldsymbol{R}^{\text{pred}}$ are vectors of R_i values, ρ_R is the covariance matrix of R_i, and $\boldsymbol{\theta}$ denotes the full parameter set.

$$\begin{aligned}
&\chi^2(|V_{ub}|, \boldsymbol{\theta}_{\text{BCL}}, \mathscr{B}_{D^*\ell\nu}) \\
&\quad = (\boldsymbol{R}^{\text{meas}} - \boldsymbol{R}^{\text{pred}}(\boldsymbol{\theta}))^{\mathsf{T}} \rho_R^{-1} (\boldsymbol{R}^{\text{meas}} - \boldsymbol{R}^{\text{pred}}(\boldsymbol{\theta})) \\
&\qquad + (\boldsymbol{\theta}_{\text{BCL}} - \boldsymbol{\theta}_{\text{BCL}}^{\text{nominal}})^{\mathsf{T}} \rho_{\text{BCL}}^{-1} (\boldsymbol{\theta}_{\text{BCL}} - \boldsymbol{\theta}_{\text{BCL}}^{\text{nominal}}) \\
&\qquad + (\mathscr{B}_{D^*\ell\nu} - \mathscr{B}_{D^*\ell\nu}^{\text{nominal}})^2 / \sigma_{D^*\ell\nu}^2
\end{aligned} \tag{5.27}$$

The first term of Equation (5.27) is where the measured ratios from the template fit enter this model. The second and third terms are Gaussian constraints for the nuisance parameters, constraining them to their nominal values according to their covariance matrix. $|V_{ub}|$ is left as a free parameter in the fit.

The covariance matrix ρ_R is calculated as the sum of two covariance matrices:

- $\rho_{R,\text{shape}}$ encodes the correlation of shape uncertainties after unfolding. This information comes from the post-fit covariance matrix of the template fit.
- $\rho_{R,\text{efficiency}}$ encodes the correlation of uncertainties in the efficiency. We calculate this matrix using MC weight variations.

We also perform a combined fit to R_i^- and R_i^0, minimising the χ^2 shown in Equation (5.28).

$$\begin{aligned}
&\chi^2(|V_{ub}|, \boldsymbol{\theta}_{\text{BCL}}, \mathscr{B}_{B^-\to D^*\ell\nu}, \mathscr{B}_{B^0\to D^*\ell\nu}) \\
&\quad = (\boldsymbol{R}^{-,\text{meas}} - \boldsymbol{R}^{-,\text{pred}}(\boldsymbol{\theta}))^{\mathsf{T}} \rho_{R^-}^{-1} (\boldsymbol{R}^{-,\text{meas}} - \boldsymbol{R}^{-,\text{pred}}(\boldsymbol{\theta})) \\
&\qquad + (\boldsymbol{R}^{0,\text{meas}} - \boldsymbol{R}^{0,\text{pred}}(\boldsymbol{\theta}))^{\mathsf{T}} \rho_{R^0}^{-1} (\boldsymbol{R}^{0,\text{meas}} - \boldsymbol{R}^{0,\text{pred}}(\boldsymbol{\theta})) \\
&\qquad + (\boldsymbol{\theta}_{\text{BCL}} - \boldsymbol{\theta}_{\text{BCL}}^{\text{nominal}})^{\mathsf{T}} \rho_{\text{BCL}}^{-1} (\boldsymbol{\theta}_{\text{BCL}} - \boldsymbol{\theta}_{\text{BCL}}^{\text{nominal}}) \\
&\qquad + (\mathscr{B}_{B^-\to D^*\ell\nu} - \mathscr{B}_{B^-\to D^*\ell\nu}^{\text{nominal}})^2 / \sigma^2_{B^-\to D^*\ell\nu} \\
&\qquad + (\mathscr{B}_{B^0\to D^*\ell\nu} - \mathscr{B}_{B^0\to D^*\ell\nu}^{\text{nominal}})^2 / \sigma^2_{B^0\to D^*\ell\nu}
\end{aligned} \tag{5.28}$$

5.8. Asimov fit results

Here we show the results of a fit to the *Asimov dataset*—this refers to replacing the detector data with the MC expectation. Asimov data is a useful tool for testing the fit procedure, as the final results should exactly match the generator inputs. We demonstrate the template fit and $|V_{ub}|$ extraction, and find that the extracted value does indeed match that used by the generator.

5.8.1. Template fit

We fitted the Asimov dataset using the simultaneous template fit model described in Section 5.6.3. The resultant yields are shown in Table 5.16. The post-fit covariance matrices of the yields are given in Tables 5.17 and 5.18.

We estimate the significance of the signal mode yields with the following method. For each region, we perform a fit to the Asimov dataset using the signal and background components. We then perform a second fit to the same data, this time using only the background template. We calculate the significance, Σ, via the log likelihood ratio between the signal+background and background-only models, as shown in Equation (5.29).

$$\Sigma = \sqrt{-2 \ln \left(\frac{\mathscr{L}_B}{\mathscr{L}_{S+B}} \right)} \tag{5.29}$$

In this particular instance, we perform individual per-region fits rather than the full simultaneous fit, so that the interpretation of Σ is more straightforward. The resulting significances are shown in Table 5.16.

Table 5.16: Resultant yields and signal significances from simultaneous fit to Asimov dataset. The stated errors are the combined statistical and shape systematic errors, as determined by `MINUIT` varying the parameter values about the minimum of the negative log likelihood.

Fit region	Signal yield	Background yield	Significance
$B^0 \to D^* \ell \nu$	236.2 ± 18.1	82.3 ± 13.2	–
$B^0 \to \pi \ell \nu$ low q^2	48.5 ± 12.2	178.4 ± 16.7	4.2σ
$B^0 \to \pi \ell \nu$ mid q^2	60.1 ± 11.8	238.2 ± 17.8	5.8σ
$B^0 \to \pi \ell \nu$ high q^2	42.3 ± 10.9	332.0 ± 20.2	4.4σ
$B^+ \to D^* \ell \nu$	374.5 ± 23.8	187.2 ± 19.5	–
$B^+ \to \pi \ell \nu$ low q^2	60.9 ± 12.2	143.2 ± 15.2	5.5σ
$B^+ \to \pi \ell \nu$ mid q^2	74.5 ± 16.3	449.6 ± 25.3	4.9σ
$B^+ \to \pi \ell \nu$ high q^2	55.2 ± 17.7	787.7 ± 32.3	3.3σ

Table 5.17: Post-fit covariance matrix of signal yields in simultaneous fit to Asimov dataset (B^0 reconstruction).

	$D^* \ell \nu$	$\pi \ell \nu$ low q^2	$\pi \ell \nu$ mid q^2	$\pi \ell \nu$ high q^2
$D^* \ell \nu$	328.6	1.04×10^{-5}	-1.38×10^{-3}	5.22×10^{-3}
$\pi \ell \nu$ low q^2		148.7	3.48×10^{-3}	-4.07×10^{-3}
$\pi \ell \nu$ mid q^2			139.6	1.91×10^{-3}
$\pi \ell \nu$ high q^2				118.2

Table 5.18: Post-fit covariance matrix of signal yields in simultaneous fit to Asimov dataset (B^+ reconstruction).

	$D^* \ell \nu$	$\pi \ell \nu$ low q^2	$\pi \ell \nu$ mid q^2	$\pi \ell \nu$ high q^2
$D^* \ell \nu$	565.9	-1.17×10^{-2}	-4.16×10^{-2}	1.47×10^{-2}
$\pi \ell \nu$ low q^2		148.8	1.37×10^{-2}	2.13×10^{-2}
$\pi \ell \nu$ mid q^2			266.2	-9.78×10^{-3}
$\pi \ell \nu$ high q^2				312.3

At this point, we unfold these yields and covariance matrices according to the procedure in Section 5.6.4. The unfolded yields are shown in Table 5.19 and the unfolded covariance matrices in Tables 5.20 and 5.21.

Table 5.19: Unfolded signal mode yields. Compare to Table 5.16.

Fit region	Signal yield
$B^0 \to \pi \ell \nu$ low q^2	48.1 ± 12.9
$B^0 \to \pi \ell \nu$ mid q^2	61.2 ± 12.9
$B^0 \to \pi \ell \nu$ high q^2	41.6 ± 11.3
$B^+ \to \pi \ell \nu$ low q^2	64.5 ± 13.5
$B^+ \to \pi \ell \nu$ mid q^2	76.9 ± 18.6
$B^+ \to \pi \ell \nu$ high q^2	49.2 ± 18.2

	$D^*\ell\nu$	$\pi\ell\nu$ low q^2	$\pi\ell\nu$ mid q^2	$\pi\ell\nu$ high q^2
$D^*\ell\nu$	328.6	8.66×10^{-5}	-1.72×10^{-3}	5.48×10^{-3}
$\pi\ell\nu$ low q^2		165.8	−16.3	0.34
$\pi\ell\nu$ mid q^2			166.7	−11.0
$\pi\ell\nu$ high q^2				128.1

Table 5.20: Unfolded post-fit covariance matrix for B^0 signal yields. Compare to Table 5.17.

	$D^*\ell\nu$	$\pi\ell\nu$ low q^2	$\pi\ell\nu$ mid q^2	$\pi\ell\nu$ high q^2
$D^*\ell\nu$	565.9	-1.13×10^{-2}	-4.65×10^{-2}	1.93×10^{-2}
$\pi\ell\nu$ low q^2		182.1	−28.6	1.2
$\pi\ell\nu$ mid q^2			346.6	−39.1
$\pi\ell\nu$ high q^2				331.9

Table 5.21: Unfolded post-fit covariance matrix for B^+ signal yields. Compare to Table 5.18.

We then calculate R_i^0 and R_i^- by inputting these yields and the signal efficiencies in Tables 5.13 and 5.14 into Equations (5.10) and (5.11). The resulting values are shown in Table 5.22.

Ratio	Asimov result
R^0_{low}	1.08×10^{-3}
R^0_{mid}	1.02×10^{-3}
R^0_{high}	7.42×10^{-4}
R^-_{low}	5.42×10^{-4}
R^-_{mid}	5.13×10^{-4}
R^-_{high}	3.65×10^{-4}

Table 5.22: Partial branching fraction ratios calculated from fit to Asimov dataset.

The uncertainties listed in the Table 5.16 include both the statistical and shape systematic uncertainty. These are separated out into individual contributions by first performing the template fit and recording the yield uncertainties, then fixing all bin parameters to their fitted values and rerunning the fit. The resultant (fully uncorrelated) yield uncertainties are purely statistical. We subtract these in quadrature from the original yield uncertainties to derive the template shape uncertainties.

The contributions from each source of systematic uncertainty are given in Table 5.23. The (partially correlated) shape uncertainties are handled by passing the post-fit covariance matrix to the `correlated_values` method of the `uncertainties` package [134], and then dividing the yields. This ensures that correlations between fitted yields are correctly handled in the ratio. For each efficiency correction, the uncertainty contribution was determined by calculating R^0 and R^- for each Gaussian weight variation, and then taking the standard deviation of the resulting values.

[134]: Lebigot (2023), *uncertainties*

Finally, in order to use the fitted values of R^0 and R^- to extract $|V_{ub}|$, we will also need the covariance matrix of the three R_i^0 and R_i^- values. The $\rho_{R,\text{shape}}$ is derived using the unfolded post-fit covariance matrix and the

Table 5.23: Relative systematic uncertainty contributions (%) to Asimov ratio measurements. Note that the values broken down by q^2 bin are the values *prior to unfolding*.

Ratio and q^2 region	R^0			R^-		
	Low	Mid	High	Low	Mid	High
Statistical $B \to \pi\ell\nu$	23.5	18.8	24.5	19.0	20.3	29.0
Statistical $B \to D^*\ell\nu$	7.4	7.4	7.4	6.1	6.1	6.1
Template shape	9.3	5.9	8.2	6.5	8.5	13.5
Lepton ID	0.04	0.1	0.4	0.2	0.3	0.6
Pion ID	0.2	0.09	0.3	0.09	0.09	0.09
Kaon ID	0.3	0.3	0.3	0.2	0.2	0.2
Tracking	0.6	0.6	0.6	0.6	0.6	0.6
Slow π^+ efficiency	2.6	2.6	2.6	–	–	–
Slow π^0 efficiency	–	–	–	6.1	6.1	6.1
π^0 efficiency	–	–	–	4.8	4.8	4.7
D FF	0.07	0.06	0.07	0.07	0.06	0.07
D^* FF	0.3	0.3	0.3	0.3	0.3	0.3
Total	26.4	21.3	27.0	22.4	24.1	33.5

function `uncertainties.correlated_values`. Its elements are shown in Tables 5.24 and 5.25. $\rho_{R,\text{efficiency}}$ is derived by calculating R_i^0 and R_i^- for all Gaussian variations (binned using q^2_{gen}), and calculating the covariance of the results. Its elements are shown in Tables 5.26 and 5.27. The total covariance matrix, ρ_R, is the sum of these two.

Table 5.24: Covariance matrix of statistical and shape systematics, $\rho_{R,\text{shape}}$, for R_i^0 fit.

	R^0_{low}	R^0_{mid}	R^0_{high}
R^0_{low}	9.03×10^{-8}	3.64×10^{-10}	4.85×10^{-9}
R^0_{mid}		5.20×10^{-8}	1.18×10^{-9}
R^0_{high}			4.41×10^{-8}

Table 5.25: Covariance matrix of statistical and shape systematics, $\rho_{R,\text{shape}}$, for R_i^- fit.

	R^-_{low}	R^-_{mid}	R^-_{high}
R^-_{low}	1.40×10^{-8}	-4.83×10^{-10}	8.71×10^{-10}
R^-_{mid}		1.65×10^{-8}	-1.18×10^{-9}
R^-_{high}			1.88×10^{-8}

Table 5.26: Covariance matrix for efficiency systematics, $\rho_{R,\text{efficiency}}$, for R_i^0 fit.

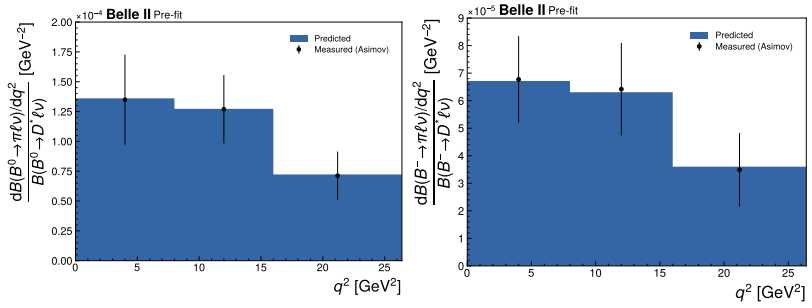
	R^0_{low}	R^0_{mid}	R^0_{high}
R^0_{low}	8.32×10^{-10}	7.76×10^{-10}	5.67×10^{-10}
R^0_{mid}		7.32×10^{-10}	5.36×10^{-10}
R^0_{high}			4.07×10^{-10}

Table 5.27: Covariance matrix for efficiency systematics, $\rho_{R,\text{efficiency}}$, for R_i^- fit.

	R^-_{low}	R^-_{mid}	R^-_{high}
R^-_{low}	1.77×10^{-9}	1.67×10^{-9}	1.18×10^{-9}
R^-_{mid}		1.59×10^{-9}	1.12×10^{-9}
R^-_{high}			7.98×10^{-10}

5.8.2. $|V_{ub}|$ extraction

Using the results of the Asimov template fit in Section 5.8.1, we performed a fit with the model described in Section 5.7. The pre-fit plots of the partial branching fraction ratios are shown in Figure 5.13.



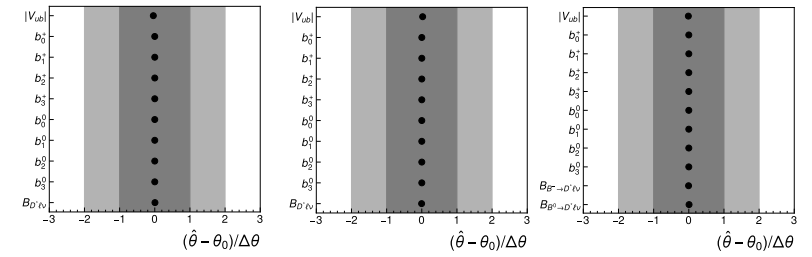
(a) R_i^0 fit. **(b)** R_i^- fit.

Figure 5.13: Pre-fit plots of the fits to partial branching fraction ratios.

The resulting $|V_{ub}|$ values are listed in Table 5.28, and the parameter pulls are shown in Figure 5.14. Figure 5.15 shows the post-fit differential decay rate distribution with resultant uncertainty bands. The value of $|V_{ub}|$ used by EvtGen is 3.72×10^{-3}. The values extracted from the Asimov fit match this within 0.4%. The relative uncertainty in $|V_{ub}|$ is ~9%.

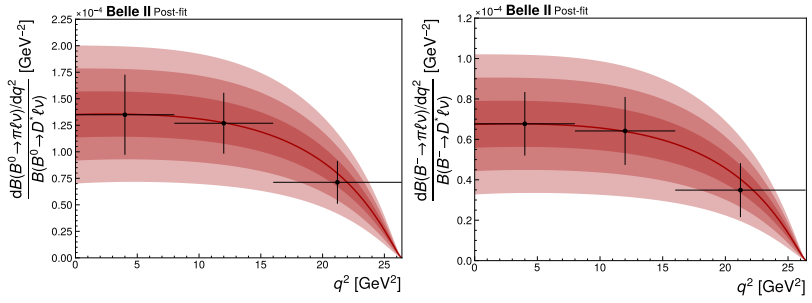
Fit	$\lvert V_{ub}\rvert$ measurement	$\chi^2/d.o.f.$
R_i^0	$(3.707 \pm 0.322) \times 10^{-3}$	0.001
R_i^-	$(3.729 \pm 0.351) \times 10^{-3}$	0.006
Combined	$(3.717 \pm 0.257) \times 10^{-3}$	0.008

Table 5.28: Extracted values of $|V_{ub}|$ from the two fits.



(a) R_i^0 fit. **(b)** R_i^- fit. **(c)** Combined fit.

Figure 5.14: Post-fit pull plots of parameters of fits to partial branching fraction ratios.



(a) R_i^0 fit. **(b)** R_i^- fit.

Figure 5.15: Post-fit distribution of differential branching fraction ratio. Shaded regions indicate 1σ, 2σ and 3σ uncertainty bands.

5.8.3. Summary

We have now demonstrated the entire fit procedure using the Asimov dataset. The procedure works as expected, giving us back the input value of $|V_{ub}|$, with all pulls in Figure 5.14 being zero. Finally, we have produced an estimate for the expected uncertainty on $|V_{ub}|$.

5.9. Unblinded pre-fit plots

We have now constructed the full analysis pipeline, from reconstruction to template fitting to $|V_{ub}|$ extraction. We have applied all relevant MC corrections, and validated the data-MC agreement in the sideband region. We now unblind and look at the detector data in the signal region, in preparation for fitting. Figures 5.16 to 5.18 show the variables to be used for signal extraction. The signal mode plots are shown for each q^2 region individually and for all q^2. All MC efficiency corrections are included in these plots.

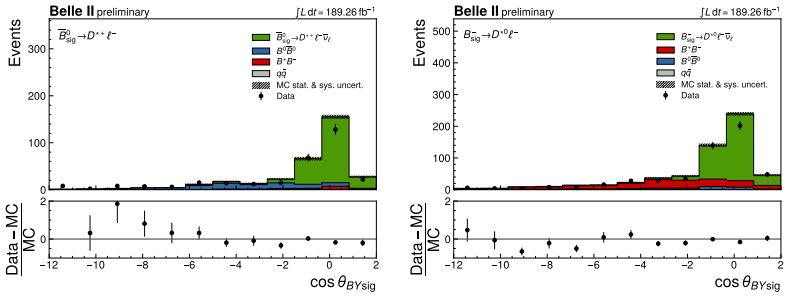


Figure 5.16: Pre-fit plots of signal extraction variable for B^0 and B^+ normalisation mode.

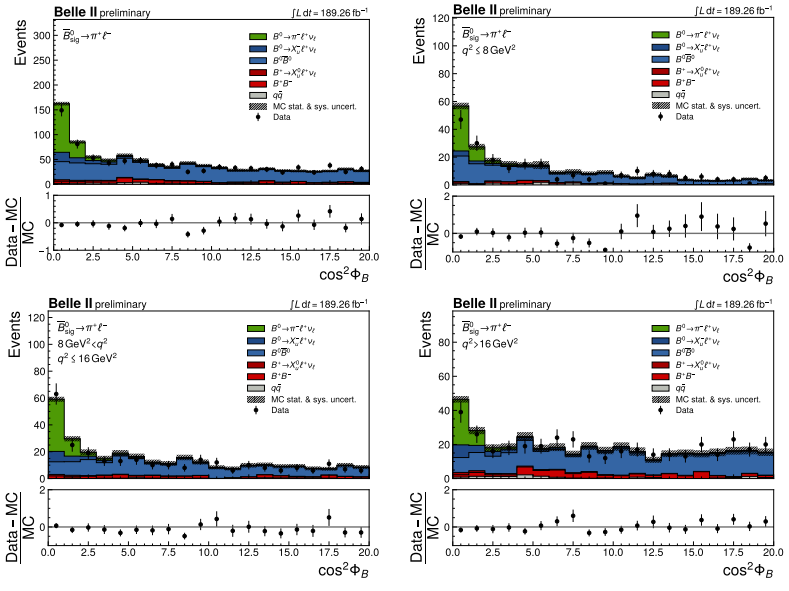


Figure 5.17: Pre-fit plots of signal extraction variable for B^0 signal mode. Distribution is shown for all q^2 and for the three q^2 regions.

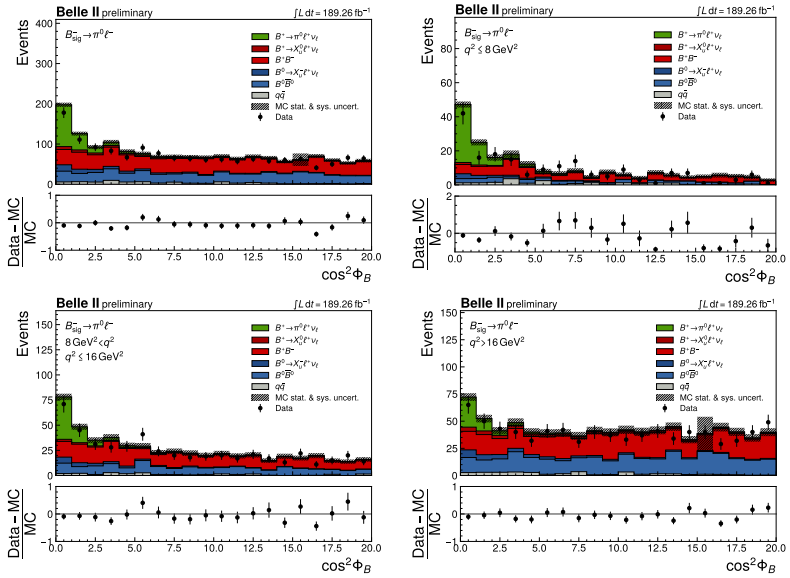


Figure 5.18: Pre-fit plots of signal extraction variable for B^+ signal mode. Distribution is shown for all q^2 and for the three q^2 regions.

5.10. Results of fit to detector data

5.10.1. Template fit

We performed a simultaneous template fit to the detector data using the model described in Section 5.6.3. The resultant yields are shown in Table 5.29, along with the signal yield significances. The post-fit covariance matrices of the yields are given in Tables 5.30 and 5.31.

Fit region	Signal yield	Background yield	Significance
$B^0 \to D^*\ell\nu$	200.7 ± 16.9	95.3 ± 13.4	–
$B^0 \to \pi\ell\nu$ low q^2	38.4 ± 11.7	172.6 ± 16.5	3.4σ
$B^0 \to \pi\ell\nu$ mid q^2	67.2 ± 11.6	201.8 ± 16.4	6.8σ
$B^0 \to \pi\ell\nu$ high q^2	28.1 ± 10.1	348.9 ± 20.6	3.0σ
$B^+ \to D^*\ell\nu$	360.7 ± 23.0	154.3 ± 17.9	–
$B^+ \to \pi\ell\nu$ low q^2	47.7 ± 11.3	133.3 ± 14.6	4.6σ
$B^+ \to \pi\ell\nu$ mid q^2	62.7 ± 15.6	433.3 ± 24.8	4.3σ
$B^+ \to \pi\ell\nu$ high q^2	50.3 ± 17.2	745.7 ± 31.5	3.1σ

Table 5.29: Resultant yields and signal significances from simultaneous fit to data. The stated errors are the combined statistical and shape systematic errors, as determined by `MINUIT` varying the parameter values about the minimum of the negative log likelihood.

	$D^*\ell\nu$	$\pi\ell\nu$ low q^2	$\pi\ell\nu$ mid q^2	$\pi\ell\nu$ high q^2
$D^*\ell\nu$	285.1	1.26×10^{-2}	-7.09×10^{-3}	-4.71×10^{-3}
$\pi\ell\nu$ low q^2		136.7	1.04×10^{-2}	-1.98×10^{-2}
$\pi\ell\nu$ mid q^2			135.3	-4.99×10^{-3}
$\pi\ell\nu$ high q^2				102.4

Table 5.30: Post-fit covariance matrix of signal yields in simultaneous fit to data (B^0 reconstruction).

Table 5.31: Post-fit covariance matrix of signal yields in simultaneous fit to data (B^+ reconstruction).

	$D^*\ell\nu$	$\pi\ell\nu$ low q^2	$\pi\ell\nu$ mid q^2	$\pi\ell\nu$ high q^2
$D^*\ell\nu$	527.4	-7.82×10^{-3}	-4.30×10^{-2}	2.02×10^{-2}
$\pi\ell\nu$ low q^2		126.8	3.42×10^{-2}	-3.03×10^{-2}
$\pi\ell\nu$ mid q^2			243.8	-2.75×10^{-2}
$\pi\ell\nu$ high q^2				294.9

The pre-fit and post-fit distributions are shown in Figures 5.19 and 5.20. We check the goodness-of-fit by calculating the χ^2, taking the number of degrees of freedom as $n_{\text{bins}} - n_{\text{yields}} = 72 - 8$. This gives us values of $\chi^2/d.o.f.$ of $43.53/64 = 0.68$ and $45.87/64 = 0.72$ for the B^0 and B^+ fits, respectively.

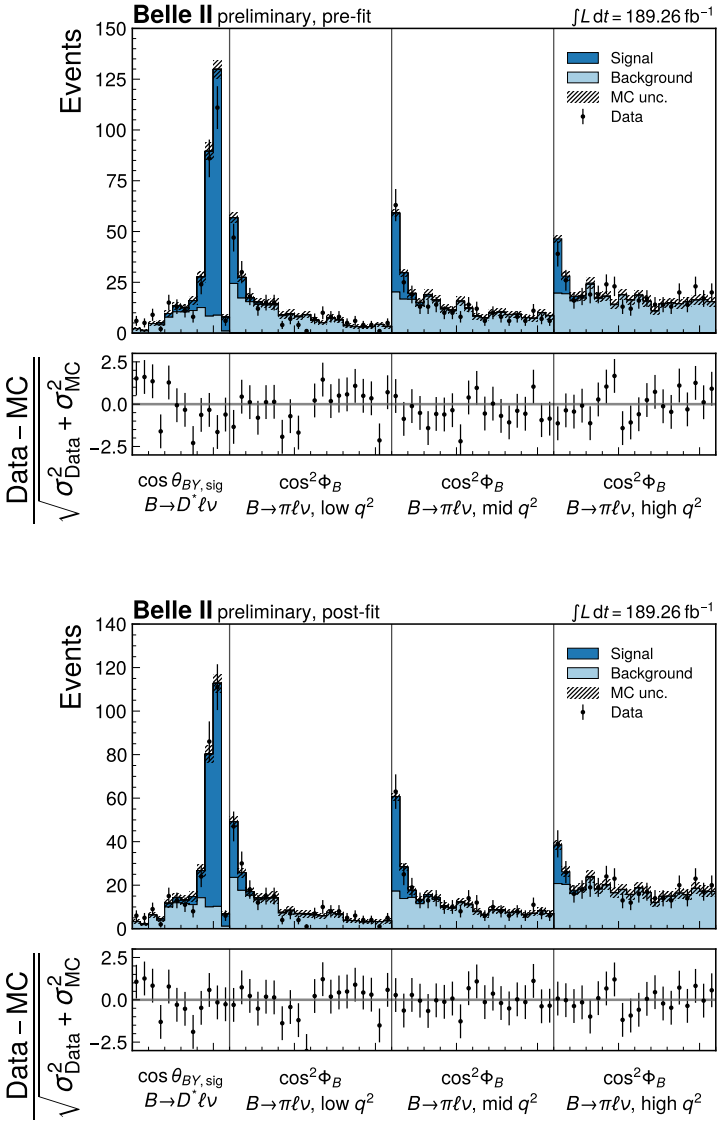


Figure 5.19: Pre-fit (top) and post-fit (bottom) plots of all regions in B^0 template fit. $\cos\theta_{BY,\text{sig}}$ is binned from -12 to 2 with bin widths of 1, and $\cos^2\Phi_B$ is binned from 0 to 20 with bin widths of 1; this is the same binning used in Figures 5.16 and 5.17.

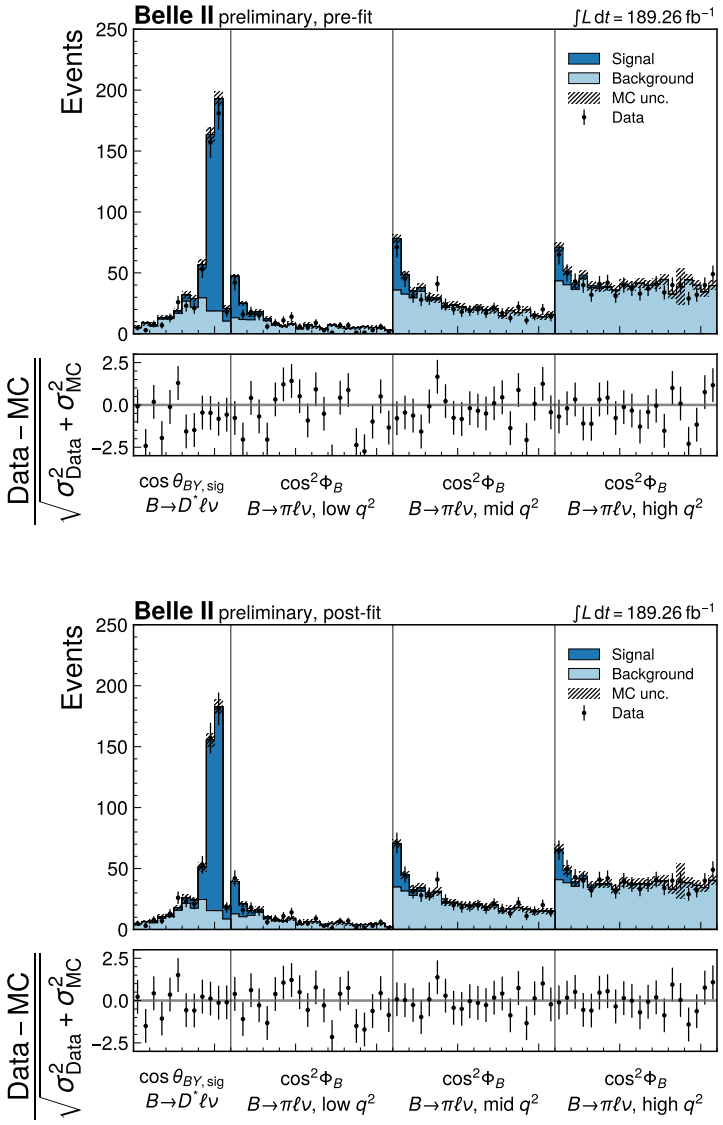


Figure 5.20: Pre-fit (top) and post-fit (bottom) plots of all regions in B^+ template fit. $\cos\theta_{BY,\text{sig}}$ is binned from -12 to 2 with bin widths of 1, and $\cos^2\Phi_B$ is binned from 0 to 20 with bin widths of 1; this is the same binning used in Figures 5.16 and 5.18.

We now unfold these yields and covariance matrices according to the procedure in Section 5.6.4. The unfolded yields are shown in Table 5.32 and the unfolded covariance matrices are shown in Tables 5.33 and 5.34.

Fit region	Signal yield
$B^0 \to \pi\ell\nu$ low q^2	37.1 ± 12.3
$B^0 \to \pi\ell\nu$ mid q^2	70.1 ± 12.7
$B^0 \to \pi\ell\nu$ high q^2	26.6 ± 10.5
$B^+ \to \pi\ell\nu$ low q^2	50.4 ± 12.5
$B^+ \to \pi\ell\nu$ mid q^2	65.0 ± 17.8
$B^+ \to \pi\ell\nu$ high q^2	45.3 ± 17.7

Table 5.32: Unfolded signal mode yields. Compare to Table 5.29.

Table 5.33: Unfolded post-fit covariance matrix for B^0 signal yields. Compare to Table 5.30.

	$D^*\ell\nu$	$\pi\ell\nu$ low q^2	$\pi\ell\nu$ mid q^2	$\pi\ell\nu$ high q^2
$D^*\ell\nu$	285.1	1.36×10^{-2}	-8.22×10^{-3}	-4.62×10^{-3}
$\pi\ell\nu$ low q^2		152.5	−15.4	0.30
$\pi\ell\nu$ mid q^2			161.5	−10.1
$\pi\ell\nu$ high q^2				110.9

Table 5.34: Unfolded post-fit covariance matrix for B^+ signal yields. Compare to Table 5.31.

	$D^*\ell\nu$	$\pi\ell\nu$ low q^2	$\pi\ell\nu$ mid q^2	$\pi\ell\nu$ high q^2
$D^*\ell\nu$	527.4	-6.99×10^{-3}	-4.87×10^{-2}	2.50×10^{-2}
$\pi\ell\nu$ low q^2		155.3	−25.1	1.0
$\pi\ell\nu$ mid q^2			317.3	−36.1
$\pi\ell\nu$ high q^2				313.3

We then calculate R_i^0 and R_i^- by inputting these yields and the signal efficiencies in Tables 5.13 and 5.14 into Equations (5.10) and (5.11). The resulting values are shown in Table 5.35.

Table 5.35: Partial branching fraction ratios calculated from fit to data.

Ratio	Data result
R^0_{low}	9.80×10^{-4}
R^0_{mid}	1.37×10^{-3}
R^0_{high}	5.58×10^{-4}
R^-_{low}	4.39×10^{-4}
R^-_{mid}	4.50×10^{-4}
R^-_{high}	3.49×10^{-4}

The contributions from each source of systematic uncertainty to the ratios are given in Table 5.36. This table was calculated using the same procedure as Table 5.23, except using the results of the fit to detector data rather than to Asimov data.

Table 5.36: Relative systematic uncertainty contributions (%) to data ratio measurements. Note that the values broken down by q^2 bin are the values *prior to unfolding*.

Ratio and q^2 region	R^0			R^-		
	Low	Mid	High	Low	Mid	High
Statistical $B \to \pi\ell\nu$	28.2	16.8	33.9	22.2	22.9	30.9
Statistical $B \to D^*\ell\nu$	8.1	8.1	8.1	6.1	6.1	6.1
Template shape	11.7	4.8	12.5	8.1	10.0	14.6
Lepton ID	0.04	0.1	0.4	0.2	0.3	0.6
Pion ID	0.2	0.09	0.3	0.09	0.09	0.09
Kaon ID	0.3	0.3	0.3	0.2	0.2	0.2
Tracking	0.6	0.6	0.6	0.6	0.6	0.6
Slow π^+ efficiency	2.6	2.6	2.6	–	–	–
Slow π^0 efficiency	–	–	–	6.1	6.1	6.1
π^0 efficiency	–	–	–	4.8	4.8	4.7
D FF	0.07	0.06	0.07	0.07	0.06	0.07
D^* FF	0.3	0.3	0.3	0.3	0.3	0.3
Total	31.7	19.4	37.1	25.7	26.9	35.6

The covariance matrices $\rho_{R,\text{efficiency}}$ and $\rho_{R,\text{shape}}$ were calculated using the same method detailed in Section 5.8.1. $\rho_{R,\text{shape}}$ for the R_i^0 and R_i^- fits are shown in Tables 5.37 and 5.38, and $\rho_{R,\text{efficiency}}$ for the R_i^0 and R_i^- fits are shown in Tables 5.39 and 5.40.

	R^0_{low}	R^0_{mid}	R^0_{high}
R^0_{low}	1.13×10^{-7}	1.54×10^{-9}	4.04×10^{-9}
R^0_{mid}		7.49×10^{-8}	1.25×10^{-9}
R^0_{high}			5.12×10^{-8}

Table 5.37: Covariance matrix of statistical and shape systematics, $\rho_{R,\text{shape}}$, for R_i^0 fit.

	R^-_{low}	R^-_{mid}	R^-_{high}
R^-_{low}	1.26×10^{-8}	-7.16×10^{-10}	6.91×10^{-10}
R^-_{mid}		1.61×10^{-8}	-1.29×10^{-9}
R^-_{high}			1.91×10^{-8}

Table 5.38: Covariance matrix of statistical and shape systematics, $\rho_{R,\text{shape}}$, for R_i^- fit.

	R^0_{low}	R^0_{mid}	R^0_{high}
R^0_{low}	6.86×10^{-10}	9.50×10^{-10}	3.87×10^{-10}
R^0_{mid}		1.33×10^{-9}	5.43×10^{-10}
R^0_{high}			2.30×10^{-10}

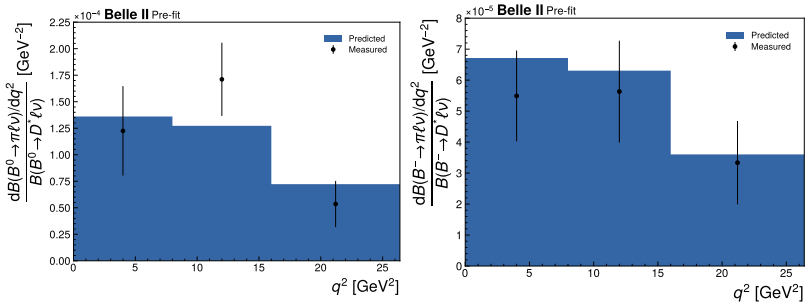
Table 5.39: Covariance matrix for efficiency systematics, $\rho_{r,\text{efficiency}}$, for R_i^0 fit.

	R^-_{low}	R^-_{mid}	R^-_{high}
R^-_{low}	1.16×10^{-9}	1.19×10^{-9}	9.14×10^{-10}
R^-_{mid}		1.22×10^{-9}	9.41×10^{-10}
R^-_{high}			7.29×10^{-10}

Table 5.40: Covariance matrix for efficiency systematics, $\rho_{R,\text{efficiency}}$, for R_i^- fit.

5.10.2. $|V_{ub}|$ extraction

Using the results of the template fit to data in Section 5.10.1, we performed a fit with the model described in Section 5.7. The pre-fit plots of the partial branching fraction ratios are shown in Figure 5.21.



(a) R_i^0 fit. **(b)** R_i^- fit.

Figure 5.21: Pre-fit plots of the fits to partial branching fraction ratios.

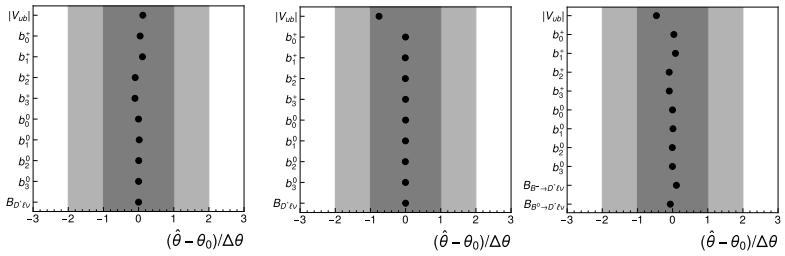
The resulting $|V_{ub}|$ values are listed in Table 5.41, and the parameter pulls are shown in Figure 5.22. Figure 5.23 shows the post-fit differential

decay rate distribution with resultant uncertainty bands. The three determinations of $|V_{ub}|$ are all consistent with the current average for exclusive measurements, which is $(3.70 \pm 0.10 \pm 0.12) \times 10^{-3}$ [20].

[20]: Particle Data Group (2022), "Review of Particle Physics"

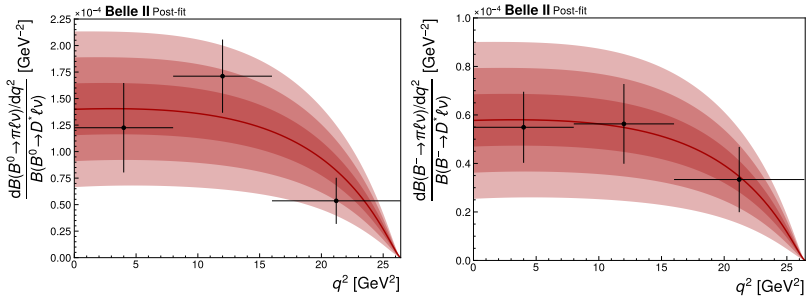
Table 5.41: Extracted values of $|V_{ub}|$ from the two fits.

Fit	$\lvert V_{ub}\rvert$ measurement	$\chi^2/d.o.f.$
R_i^0	$(3.763 \pm 0.350) \times 10^{-3}$	1.23
R_i^-	$(3.455 \pm 0.351) \times 10^{-3}$	0.05
Combined	$(3.598 \pm 0.266) \times 10^{-3}$	1.52



(a) R_i^0 fit. **(b)** R_i^- fit. **(c)** Combined fit.

Figure 5.22: Post-fit pull plots of parameters of fits to partial branching fraction ratios.



(a) R_i^0 fit. **(b)** R_i^- fit.

Figure 5.23: Post-fit distribution of differential branching fraction ratio. Shaded regions indicate 1σ, 2σ and 3σ uncertainty bands.

5.11. Summary

In this Chapter, we detailed our extraction of $|V_{ub}|$ from exclusive semileptonic decays using SL FEI tagging. The combined fit result for $|V_{ub}|$ is $(3.598 \pm 0.266) \times 10^{-3}$, which is consistent with the current average for exclusive determinations.

This analysis is one of the first Belle II analyses to use the SL FEI. There are under-studied effects in the SL FEI that have led to difficulties for constructing a general-purpose calibration. However, this analysis circumvented this issue by instead using a normalisation mode approach. We validated the approach prior to unblinding, and the consistency of the unblinded results confirms its validity.

There are two main areas for future development of this analysis. Firstly, the most recent version of the FEI has addressed the sculpting

of the tag-side $\cos\theta_{BY}$ distribution for background events. Since $\cos\theta_{BY}$ is an input to the fitting variable, the new FEI training could provide even stronger discrimination between signal and background events. Secondly, the dominant uncertainties in this analysis are statistical, and thus are a function of the amount of data Belle II has collected. The dataset used in this analysis was $189.9\,\text{fb}^{-1}$, compared to Belle II's lifetime target of $50\,\text{ab}^{-1}$. There is considerable future scope for this analysis to be improved with more data.

Global fit of leptoquark models with flavour constraints

6.

In the last decade, there has been renewed interest in leptoquarks, largely due to their ability to explain the flavour anomalies. There are two complementary approaches to testing LQ models: precision measurements of flavour observables, and direct searches for LQ production. By combining information from both of these approaches, we can arrive at a complete picture of the present constraints on LQ models. This is best achieved within the framework of *global fitting*.

In this Chapter, we investigate two models using global fits, each of which is a combination of two scalar LQs. We first explain the mathematical basis of global fits, and introduce GAMBIT, a tool for performing global fits to a wide range of BSM models. We then describe the implementation of the LQ models in GAMBIT and the code written to calculate new likelihoods and observables. Finally, we present preliminary results of global fits to the models, and discuss the viability of observing these models at the LHC.

6.1. Global fits

BSM models typically introduce a large number of free parameters. Consequently, testing a particular BSM model against experimental observations is a challenging task. A global fit is the modern approach to this task, in which we scan over the parameter space and test many parameter combinations.

The central object in a global fit is the *likelihood*, which is a function of the model parameters. Suppose we have a model M with parameters $\boldsymbol{\theta}$, and have some set of experimental observations $\boldsymbol{D}$. The likelihood is defined as the conditional probability of observing $\boldsymbol{D}$ given M and $\boldsymbol{\theta}$,

$$\mathscr{L}(\boldsymbol{\theta}) = P(\boldsymbol{D}|\boldsymbol{\theta}, M)\,. \tag{6.1}$$

If we have multiple sets of observations, then the combined likelihood is calculated as the product over the individual likelihoods. Doing so allows the global fit to combine measurements from different experiments or even different fields of physics. Higher values of the combined likelihood indicate that that parameter combination is a better fit to the experimental observations. Thus, the goal of a global fit is to find the parameters which maximise the combined likelihood[1].

1: In practice, global fits typically minimise the negative log-likelihood, as this is more numerically stable.

In this Chapter, we use differential evolution to efficiently find the maximum likelihood. This has been shown to be a very efficient global optimiser [135], and has become the standard within the GAMBIT community for the presentation of results in a frequentist setting.

[135]: The DarkMachines High Dimensional Sampling Group et al. (2021), "A Comparison of Optimisation Algorithms for High-Dimensional Particle and Astrophysics Applications"

In order to visualise the results of a scan, we employ a procedure known as *profiling*. Suppose that we wish to visualise one or two parameters out of the full set $\boldsymbol{\theta}$—we will denote these parameters $\boldsymbol{\phi}$ and all other parameters as $\boldsymbol{\nu}$. The profile likelihood is defined as

$$\mathscr{L}(\boldsymbol{\phi}, \hat{\hat{\boldsymbol{\nu}}}) \,, \tag{6.2}$$

where $\hat{\hat{\boldsymbol{\nu}}}$ is the value of $\boldsymbol{\nu}$ which maximises the likelihood for a given value of $\boldsymbol{\phi}$. The profile likelihood ratio is defined by dividing Equation (6.2) by the likelihood at the best fit point, $\mathscr{L}(\hat{\boldsymbol{\phi}}, \hat{\boldsymbol{\nu}})$,

$$\lambda(\boldsymbol{\phi}) = \frac{\mathscr{L}(\boldsymbol{\phi}, \hat{\hat{\boldsymbol{\nu}}})}{\mathscr{L}(\hat{\boldsymbol{\phi}}, \hat{\boldsymbol{\nu}})} \,. \tag{6.3}$$

Finally, we define a test statistic $t = -2 \ln \lambda(\boldsymbol{\theta})$. According to Wilks' theorem [136], t is χ^2-distributed in the asymptotic limit. We may then identify values of $\boldsymbol{\phi}$ as "not excluded" if t is greater than the value of the quantile function of the χ^2 distribution for a chosen significance level α,

[136]: Cowan et al. (2011), "Asymptotic Formulae for Likelihood-Based Tests of New Physics"

$$-2 \ln(\lambda(\boldsymbol{\phi})) \leq \chi^2_{n,1-\alpha} \,. \tag{6.4}$$

By performing this type of inference using global fit results, we can determine which regions of the parameter space are compatible with experimental observations.

6.2. GAMBIT

The Global and Modular BSM Inference Tool (GAMBIT) is a framework for performing global fits of BSM models [137, 138]. If any particular BSM model is true, then its signals could appear in astroparticle physics, cosmology, dark matter direct detection experiments, precision measurements, and collider experiments. GAMBIT was created to take measurements from all these fields of physics, and combine them to provide a complete picture of the constraints on BSM physics.

[137]: GAMBIT Collaboration (2017), "GAMBIT"
[138]: GAMBIT Collaboration (2020), "GAMBIT and Its Application in the Search for Physics Beyond the Standard Model"

The structure of GAMBIT is illustrated in Figure 6.1, and this section will detail

The framework is organised into modules which are given the suffix `Bit`. Some modules handle a particular domain of observations, such as ColliderBit for HEP collider experiments [139], FlavBit for flavour

[139]: Balázs et al. (2017), "ColliderBit"

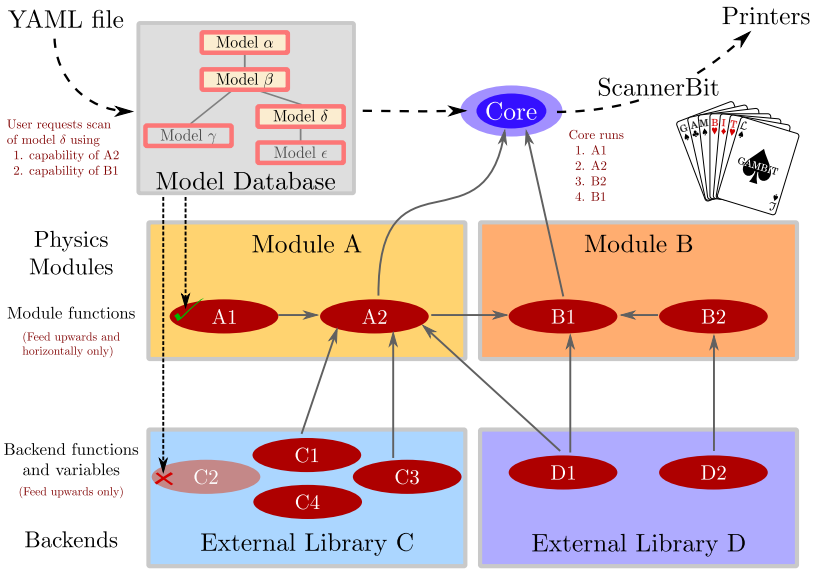


Figure 6.1: High-level overview of the GAMBIT framework [137]. The entire GAMBIT run is initiated with a user-inputted YAML steering file. Physics modules provide capabilities for a particular domain of physics, which may be required by other physics modules. Backends provide additional capabilities, with a flexible dependency structure. ScannerBit controls the sampling and likelihoods, and printer modules implement output to various formats.

physics measurements [140], and DarkBit for dark matter direct detection results. SpecBit and DecayBit handle the phenomenology of BSM models, namely their mass spectra and decays [141]. Finally, ScannerBit handles the sampling of high-dimensional spaces and the statistical treatment of the combined likelihoods [142].

[140]: Bernlochner et al. (2017), "FlavBit"

[141]: The GAMBIT Models Workgroup: et al. (2018), "SpecBit, DecayBit and PrecisionBit"

[142]: The GAMBIT Scanner Workgroup et al. (2017), "Comparison of Statistical Sampling Methods with ScannerBit, the GAMBIT Scanning Module"

GAMBIT's modular design allows for new observations to be straightforwardly included. A *capability* is a quantity with a specified type that is calculated by some function. Capabilities are chained together via *dependencies*, in which a function is declared to require another capability. If multiple functions provide a capability, any of them may be chosen as-needed at runtime, as long as all of its dependencies are fulfilled. When a new observation is to be included in a global fit, it is simply coded up as a capability. Furthermore, many capabilities that are written for one global fit can be easily reused in others.

GAMBIT has a highly flexible backend system, allowing outside code to interface with GAMBIT at runtime. The backend system is an extension of the capability concept; multiple backends can provide a given capability, and backends can be swapped out as-needed. This allows GAMBIT to make use of open source packages which perform specific calculations, such as Pythia for simulating high energy interactions [143], SuperIso for flavour observables [144], and DDCalc for dark matter direct detection observables [145].

[143]: Sjöstrand et al. (2015), "An Introduction to PYTHIA 8.2"

[144]: Mahmoudi (2008), "SuperIso"

[145]: Bringmann et al. (2017), "DarkBit"

GAMBIT is also designed to allow for addition of new models and extension of existing ones. A model in GAMBIT is defined by its free parameters. GAMBIT contains a model hierarchy, where models are declared as inheriting from or being friends of other models. Functions that are declared as model-specific can be reused by other models, according to their relation to the original model. GAMBIT allows multiple models

to be scanned over simultaneously—in particular, the parameters of the SM can be included in a BSM global fit as nuisance parameters.

6.3. Leptoquark global fit

[146]: Angelescu et al. (2021), "On the single leptoquark solutions to the B-physics anomalies"
[147]: Fajfer et al. (2016), "Vector leptoquark resolution of R_K and $R_{D^{(*)}}$ puzzles"
[37]: LHCb Collaboration (2022), "Measurement of lepton universality parameters in $B^+ \to K^+\ell^+\ell^-$ and $B^0 \to K^{*0}\ell^+\ell^-$ decays"

Much of the interest in leptoquark models in the last decade has been motivated by the anomalous measurements of $R(K^{(*)})$ and $R(D^{(*)})$ [146, 147]. Until recently, the $R(K^{(*)})$ measurement stood at a $\sim 3\sigma$ tension with the SM prediction. However, an updated $R(K^{(*)})$ measurement by LHCb in December 2022 identified a background component that was missing in previous analyses [37]. The inclusion of this component brought the measurement into agreement with the SM prediction within the uncertainties. In light of this development, it is interesting to revisit the flavour constraints on LQ models, since the new LHCb result can be expected to shrink the viable parameter space. This Section details a global fit with GAMBIT to two LQ models using a selection of flavour constraints.

Having performed the fit to flavour observables, we then turn our attention to the prospects for discovering LQs at ATLAS and CMS in the near future. This can be accomplished approximately by taking points identified by the flavour fit and testing them against the signal regions of LHC analyses designed to detect direct LQ production. In this analysis, we consider three of the most relevant ATLAS analyses; the inclusion of CMS searches is left as future work.

We chose to analyse two models, which are each combinations of two scalar LQs: S_3 & R_2, and S_1 & S_3. This choice of models is motivated by their ability to explain $R(D^{(*)})$ and $R(K^{(*)})$. The vector LQ U_1 is able to explain the flavour anomalies on its own; however, predictions with vector LQs are more sensitive to the choice of ultraviolet completion. Because of this, only scalar LQs were considered in this analysis.

The Lagrangians governing all possible interactions of R_2, S_1 and S_3 with SM fermions are shown in Equations (2.13) to (2.15). In this global fit, we consider only quark-lepton-LQ interactions, and fix all Yukawa couplings involving first-generation fermions to zero. This leaves us with five Yukawa matrices, each with four free elements: $y^{RL}_{R_2}$, $y^{LR}_{R_2}$, $y^{LL}_{S_1}$, $y^{RR}_{S_1}$, and $y^{LL}_{S_3}$. We only consider real Yukawa couplings, not for any theoretical reason, but due to software limitations. The Lagrangians of the two models are given in Equations (6.5) and (6.6) with weak isospin and generational indices suppressed.

$$\mathscr{L}_{R_2+S_3} = -y^{RL}_{R_2}\bar{u}_R R_2 i\tau_2 L_L + y^{LR}_{R_2}\bar{e}_R R_2^* Q_L + y^{LL}_{S_3}\bar{Q}^C_L i\tau_2(\tau^k S_3^k)L_L + \text{h.c.} \quad (6.5)$$

$$\mathscr{L}_{S_1+S_3} = y^{LL}_{S_1}\bar{Q}^C_L S_1 i\tau_2 L_L + y^{RR}_{S_1}\bar{u}^C_R S_1 e_R + y^{LL}_{S_3}\bar{Q}^C_L i\tau_2(\tau^k S_3^k)L_L + \text{h.c.} \quad (6.6)$$

6.3.1. Implementing models in GAMBIT

The GAMBIT Universal Model Machine (GUM) is a recent addition to GAMBIT, and is a powerful tool for automating the implementation of new BSM models [148]. GUM reads Lagrangian-level inputs and generates the code required for each GAMBIT module. This includes the model definition, mass spectrum code, decay tables, and collider simulation code.

[148]: Bloor et al. (2021), "The GAMBIT Universal Model Machine"

To use GUM, the Lagrangian must first be written symbolically in a Mathematica model file [149], which is then processed by FeynRules [150] or SARAH [151] to produce Universal FeynRules Output (UFO) files. The UFO files define the particles and interactions of the model, and these can be passed as input to many software packages. In particular, GUM passes the UFO files to MadGraph [152] to generate matrix elements for Pythia, which may be used for collider simulation and calculation of cross-sections. GUM creates a GAMBIT backend for a patched version of Pythia containing these matrix elements. GUM also uses CalcHEP [153] to produce code which calculates decay widths.

[149]: Wolfram Research, Inc. (2022), *Mathematica*
[150]: Alloul et al. (2014), "FeynRules 2.0 - A Complete Toolbox for Tree-Level Phenomenology"
[151]: Staub (2012), *Sarah*
[152]: Alwall et al. (2011), "MadGraph 5"
[153]: Belyaev et al. (2013), "CalcHEP 3.4 for Collider Physics within and beyond the Standard Model"

The Lagrangians in Equations (6.5) and (6.6) were written up as FeynRules models by Sanjay Bloor. FeynRules was used rather than SARAH, as only the former is designed to work with non-renormalisable Lagrangians. Running GUM on these models produced code for the ColliderBit, DecayBit and SpecBit modules.

The second necessary step in implementing the LQ models is to incorporate them into FlavBit. The theoretical predictions for observables in FlavBit are almost exclusively calculated by SuperIso. The observables relevant to this analysis are calculated by SuperIso using the Wilson coefficients for $b \to s\mu\mu$ and $b \to s\tau\tau$ processes. The FlavBit interface to SuperIso fills a data object with these values. I wrote model-specific capabilities for each of the Wilson coefficients, using the expressions in Equation (55) and Table 4 of Reference [51]. I extended the FlavBit interface to pass these to SuperIso, enabling the use of the already-available observables.

[51]: Doršner et al. (2016), "Physics of leptoquarks in precision experiments and at particle colliders"

6.3.2. Likelihoods and observables

The likelihoods used in the scan are experimental likelihoods of flavour observables, which are implemented in HEPLike [154]. For each point, the corresponding observable is calculated by SuperIso and passed to HEPLike.

[154]: Bhom et al. (2020), "HEPLike"

We include the following flavour observables in the combined likelihood.

- $\mathcal{B}(B_s^0 \to \phi\mu\mu)$ [155]

[155]: LHCb Collaboration (2015), "Angular analysis and differential branching fraction of the decay $B_s^0 \to \phi\mu^+\mu^-$"

- $\mathscr{B}(B_s^0 \to \mu\mu)$ and $\mathscr{B}(B^0 \to \mu\mu)$ [156–158]
- $R(K^{(*)})$ [37]

The $\mathscr{B}(B_{(s)}^0 \to \mu\mu)$ measurement is consistent with SM predictions, while the $\mathscr{B}(B_s^0 \to \phi\mu\mu)$ measurement is 3σ lower than the SM prediction. The $R(K^{(*)})$ measurement is the recent LHCb result, which removes a previous anomaly.

The publications for $\mathscr{B}(B_s^0 \to \phi\mu\mu)$, $\mathscr{B}(B_s^0 \to \mu\mu)$, and $\mathscr{B}(B^0 \to \mu\mu)$ contain the n-dimensional likelihoods of each observable, in the form of a ROOT histogram. HEPLike interpolates the histograms to provide likelihood functions for these observables. The published $R(K)$ and $R(K^*)$ measurements were each performed in the q^2 regions $[0.1\,\text{GeV}^2, 1.1\,\text{GeV}^2]$ and $[1.1\,\text{GeV}^2, 6.0\,\text{GeV}^2]$. The four-dimensional likelihood of these observables was not made public, so we instead construct a Gaussian likelihood using the central values, symmetrised uncertainties, and correlation matrix of the four values.

Having performed the flavour fit, we are interested in evaluating the prospects for direct detection of LQs at the LHC. To this end, we are interested in calculating the total production cross-section of LQ processes. Using this, we can obtain an estimate of the expected number of events by multiplying the cross-section by the integrated luminosity. We calculate the cross-section using the LQ-specific patched Pythia version. For each point, we create an SLHA file with the model parameters, and pass this to a standalone executable which generates 1000 events at $\sqrt{s} = 13\,\text{TeV}$ and returns the total cross-section. We can then determine whether the cross-section of the non-excluded points is sufficiently high to be observed at the LHC.

The final stage of this analysis is to test these points against the event yields in signal regions of LHC analyses targeting LQ production. These constraints from direct searches may be included in a scan using ColliderBit. ColliderBit contains approximately sixty of the 8 TeV and 13 TeV LHC analyses from ATLAS and CMS. The cuts in each of these analyses are implemented in a `ColliderBit::Analysis` object, as well as the measured yields for each signal region. To test a point in the parameter space against the implemented analyses, ColliderBit generates MC events using Pythia. The events are then passed through BuckFast, which performs a fast detector simulation to incorporate the systematics of the ATLAS and CMS detectors into the generated events[2]. Finally, the analysis cuts are applied to the events, and the signal region yields are recorded. The fraction of events passing the selections is scaled according to the production cross-section, and the likelihood contribution from one signal region of one analysis is calculated as

$$\mathscr{L}_i(s_i(\boldsymbol{\theta})) = \text{Poisson}(n_i | s_i(\boldsymbol{\theta}) + b_i)\,, \tag{6.7}$$

[156]: LHCb Collaboration (2017), "Measurement of the $B_s^0 \to \mu^+\mu^-$ branching fraction and effective lifetime and search for $B^0 \to \mu^+\mu^-$ decays"
[157]: CMS Collaboration (2020), "Measurement of properties of $B_s^0 \to \mu^+\mu^-$ decays and search for $B^0 \to \mu^+\mu^-$ with the CMS experiment"
[158]: ATLAS Collaboration (2019), "Study of the rare decays of B_s^0 and B^0 mesons into muon pairs using data collected during 2015 and 2016 with the ATLAS detector"
[37]: LHCb Collaboration (2022), "Measurement of lepton universality parameters in $B^+ \to K^+\ell^+\ell^-$ and $B^0 \to K^{*0}\ell^+\ell^-$ decays"

2: BuckFast is an internal package in GAMBIT.

where n_i, b_i and s_i respectively refer to the number of measured events, number of predicted background events, and number of predicted signal events for the point $\boldsymbol{\theta}$. The log-likelihood contribution from each analysis is given by Equation (6.7) normlised background-only likelihood ($s_i = 0$),

$$\Delta \ln \mathscr{L}_i(s_i(\boldsymbol{\theta})) = \ln \mathscr{L}_i(s_i(\boldsymbol{\theta})) - \ln \mathscr{L}_i(s_i = 0)\,. \tag{6.8}$$

The combined likelihood is the sum of the individual log-likelihoods marginalised over the systematic uncertainties.

Since the LHC likelihoods are calculated via Equation (6.8), it is possible for the combined likelihood to attain values greater than one for particular parameter combinations. This indicates that those models provide a better fit to the data than the SM does, which can occur if a signal region has a small excess. We take a conservative approach in assuming that any excesses are due to statistical fluctuations. In doing so, we cap the combined likelihood at the SM expectation, thereby preventing points from being favoured on account of fitting to statistically insignificant excesses. The capped likelihood is defined as

$$\ln\left(\mathscr{L}_{\text{LHC, capped}}\right) = \min[\ln(\mathscr{L}_{\text{LHC}}), 0]\,. \tag{6.9}$$

I implemented in ColliderBit the ATLAS searches detailed in References [58–60]. These analyses are direct searches for final states with a third-generation quark combined with a τ lepton or missing transverse energy, and thus they are able to probe leptoquark production through the processes shown in Figure 6.2. I coded up the selections listed in each paper, namely, selections on analysis objects (*e.g.* p_{T} of muons), event topology (*e.g.* number of b-tagged tracks), and kinematic variables (*e.g.* M_{T2} and missing transverse energy). The signal regions also include selections on specially-trained multivariate classifiers; the trained weights and code for applying them are not made publicly available, so these could not be implemented in ColliderBit. However, these are designed for background rejection, and so typically have a very high efficiency on signal events. Finally, any selections based on jet substructure (*e.g.* number of tracks in a jet) were not implemented, due to limitations of the data structures in ColliderBit at the time of writing.

I validated these analyses by implementing signal MC cutflows and comparing these to the cutflows published as supplementary data. In order to match the benchmark points of the published cutflows, I only activated a subset of the available Pythia processes and fixed the LQ masses. We aim for signal yield agreement within around 50%—better agreement is often not possible due to differences in simulation, and not strictly necessary for the purposes of ColliderBit.

[58]: ATLAS Collaboration (2021), *Search for pair production of third-generation scalar leptoquarks decaying into a top quark and a τ-lepton in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*

[59]: ATLAS Collaboration (2021), "Search for new phenomena in pp collisions in final states with tau leptons, b-jets, and missing transverse momentum with the ATLAS detector"

[60]: ATLAS Collaboration (2021), *Search for new phenomena in final states with b-jets and missing transverse momentum in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector*

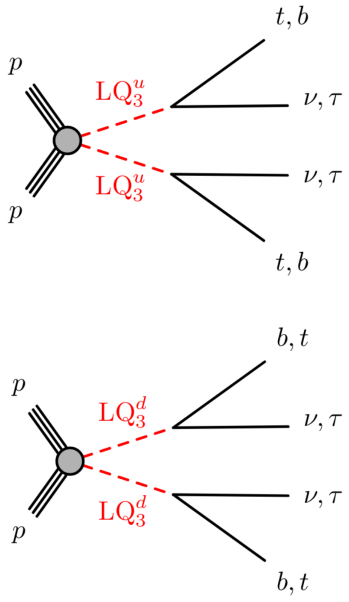


Figure 6.2: Feynman diagrams for LQ production processes with final states containing a third generation quark and either a τ lepton or missing transverse energy. The labels LQ_3^u and LQ_3^u are for the simplified LQ models used in the ATLAS analyses, the details of which are described in References [58–60].

The collider likelihoods are computationally expensive to run in a global fit, and are beyond the scope of this analysis. Instead, as a proof-of-concept, we select points within the 1σ contour of the scan of flavour constraints, and run the analyses on them. The resulting log-likelihoods indicate whether these points have been excluded by direct searches.

6.4. Results

I ran a GAMBIT scan of the two LQ models using the scanner Diver—this is an implementation the λjDE variant of differential evolution [142, 159], which is based on the rand/1/bin algorithm, but dynamically adapts the crossover rate, mutation scale factor, and mutation parameter λ during the run. I chose a population size of 19,200 and a convergence threshold of 10^{-7}. The LQ masses were allowed to float between 100 GeV and 3 TeV, and the non-zero Yukawa couplings between -0.5 and 0.5. Figures 6.3 and 6.4 show the profile likelihoods of the masses, and Figures 6.5 to 6.18 and E.1 to E.10 show the profile of the masses and couplings.

[142]: The GAMBIT Scanner Workgroup et al. (2017), "Comparison of Statistical Sampling Methods with ScannerBit, the GAMBIT Scanning Module"
[159]: Brest et al. (2006), "Self-Adapting Control Parameters in Differential Evolution"

It is important to note that these plots are highly preliminary. Each of the final datasets contained 10M points, after continuing the scans multiple times with lower and lower convergence thresholds. Despite this, the profile likelihood plots showed patchy boundaries, suggesting poor convergence. For these plots, a binning of 35 bins per parameter was chosen in order to give as smooth a boundary as possible while leaving the structure of the constraints visible. One underlying issue is that these are 14-dimensional scans in which several parameters have very little constraint. Future scans will include additional flavour likelihoods which will constrain these parameters, potentially aiding convergence.

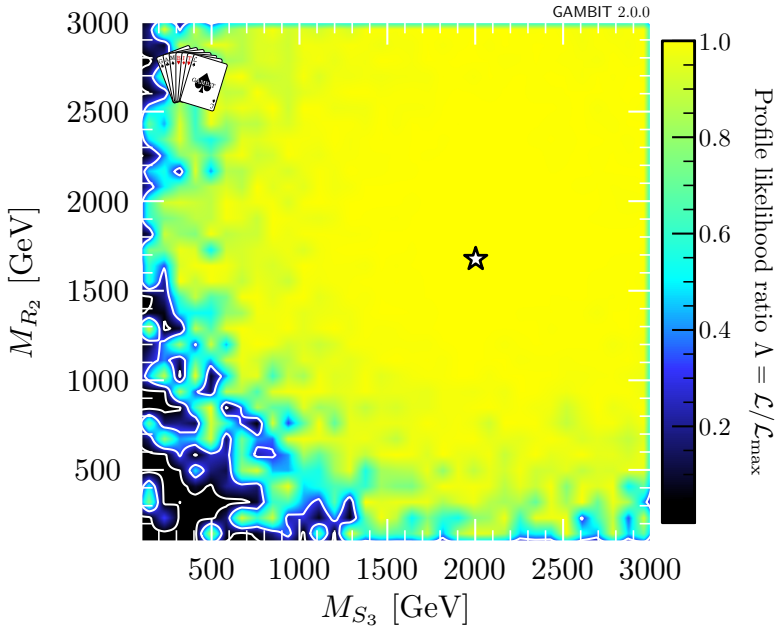


Figure 6.3: Profile likelihood plot of LQ masses in $R_2 + S_3$ scan.

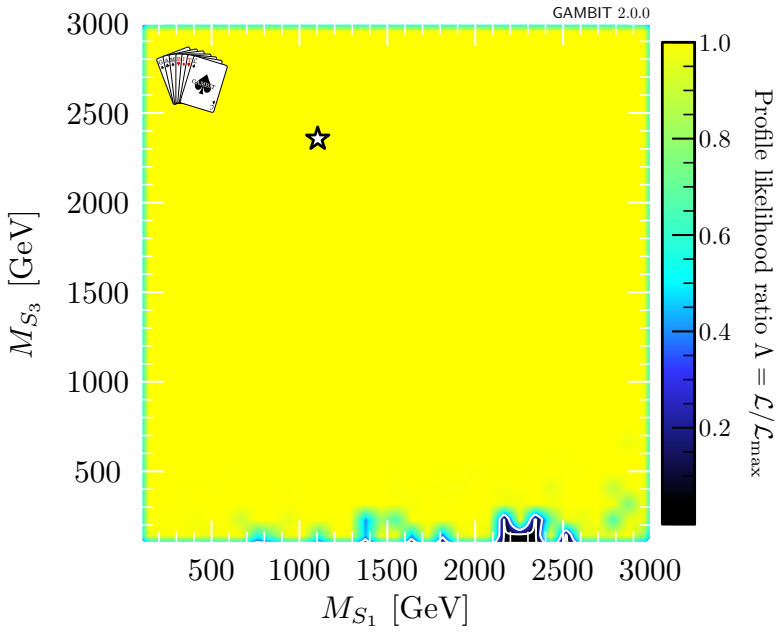


Figure 6.4: Profile likelihood plot of LQ masses in $S_1 + S_3$ scan.

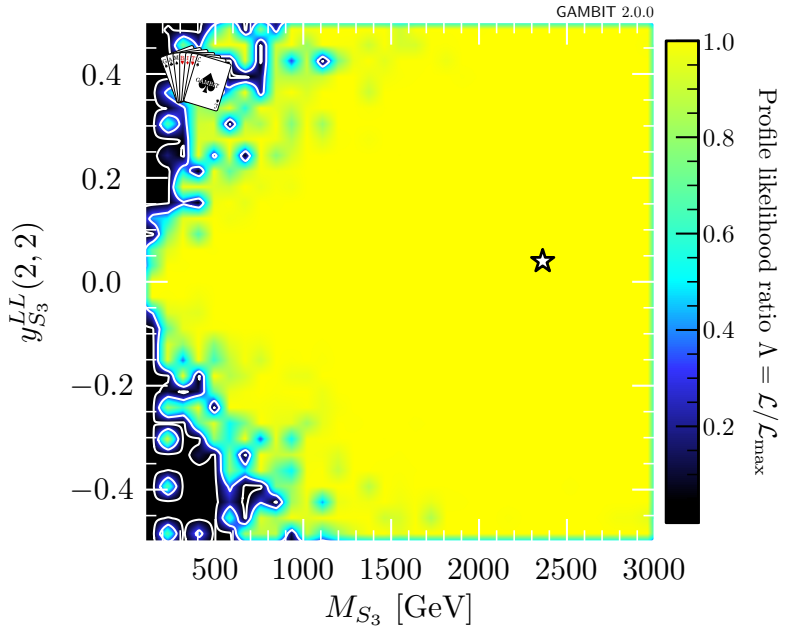


Figure 6.5: Profile likelihood plot of M_{S_3} and $(2,2)$ element of $y^{LL}_{S_3}$ in $S_1 + S_3$ scan.

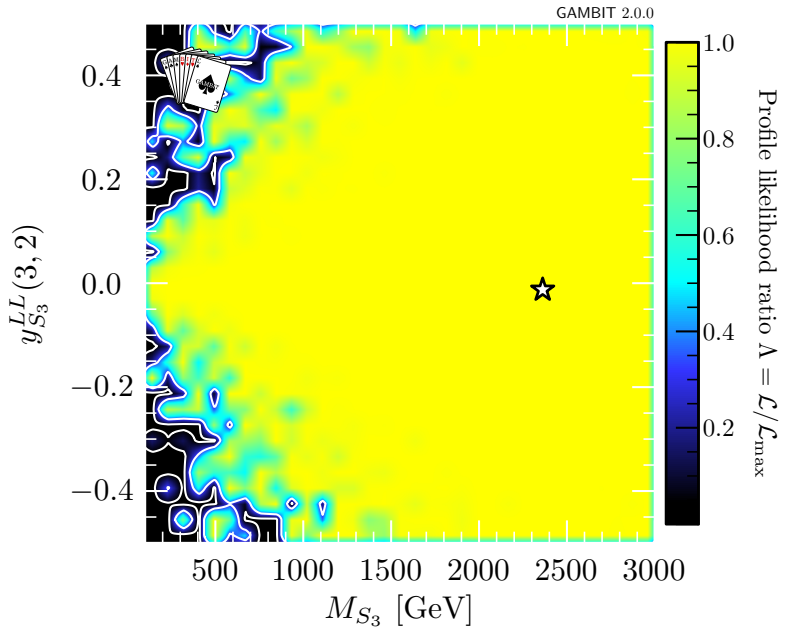


Figure 6.6: Profile likelihood plot of M_{S_3} and $(3,2)$ element of $y^{LL}_{S_3}$ in $S_1 + S_3$ scan.

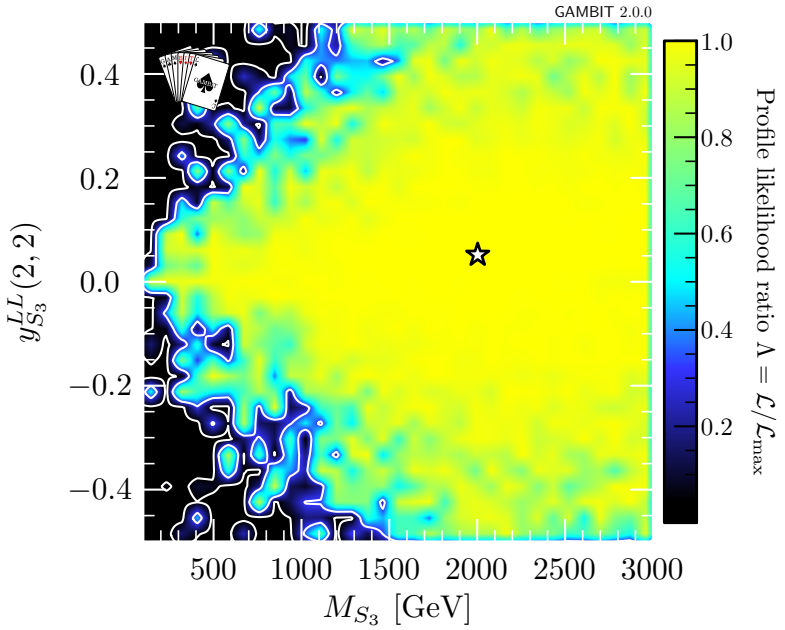


Figure 6.7: Profile likelihood plot of M_{S_3} and $(2,2)$ element of $y^{LL}_{S_3}$ in $R_2 + S_3$ scan.

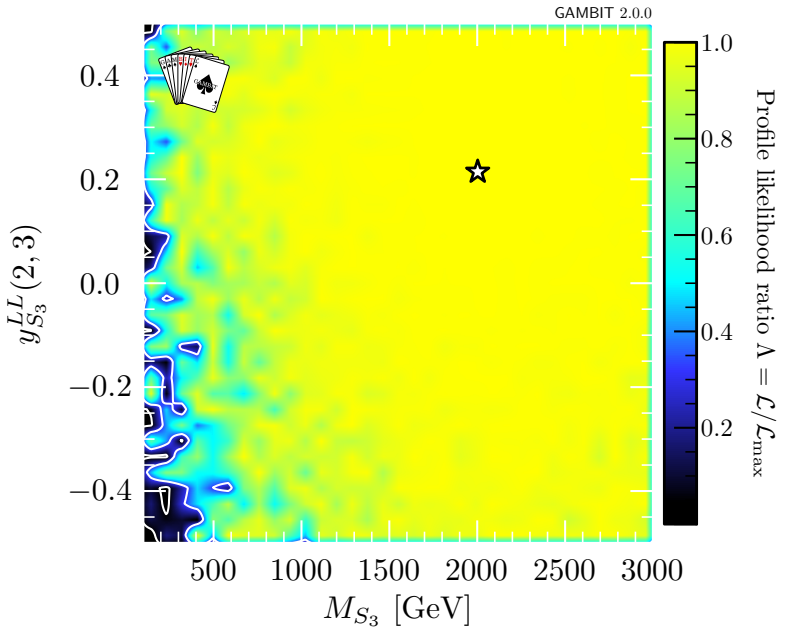


Figure 6.8: Profile likelihood plot of M_{S_3} and $(2,3)$ element of $y^{LL}_{S_3}$ in $R_2 + S_3$ scan.

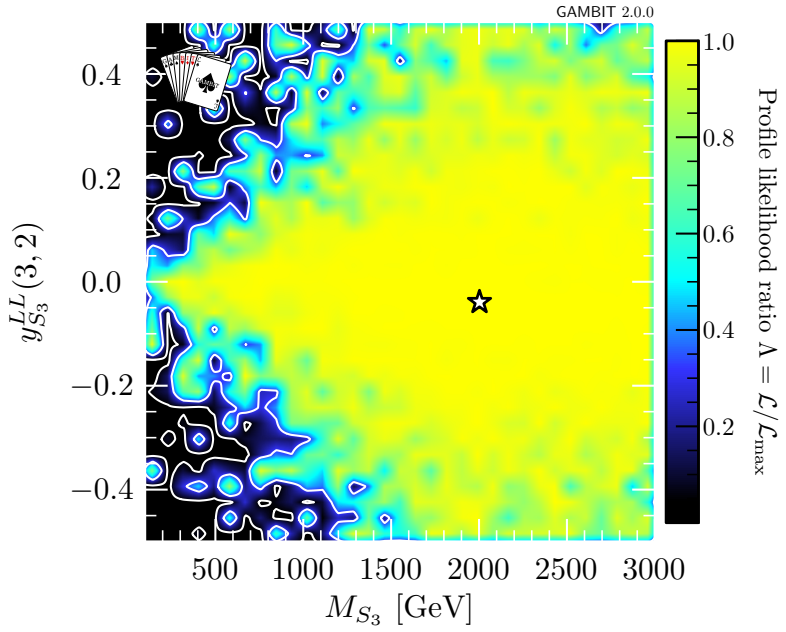


Figure 6.9: Profile likelihood plot of M_{S_3} and $(3,2)$ element of $y^{LL}_{S_3}$ in $R_2 + S_3$ scan.

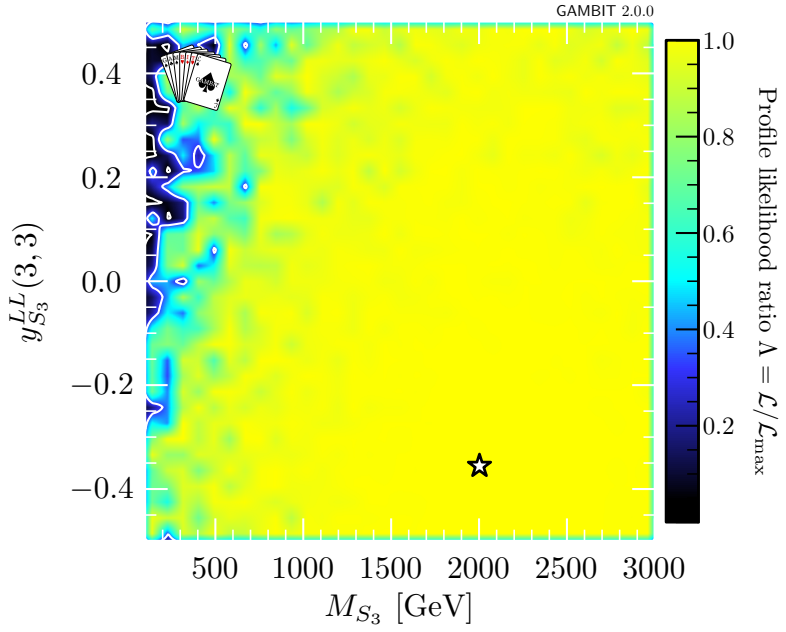


Figure 6.10: Profile likelihood plot of M_{S_3} and $(3,3)$ element of $y^{LL}_{S_3}$ in $R_2 + S_3$ scan.

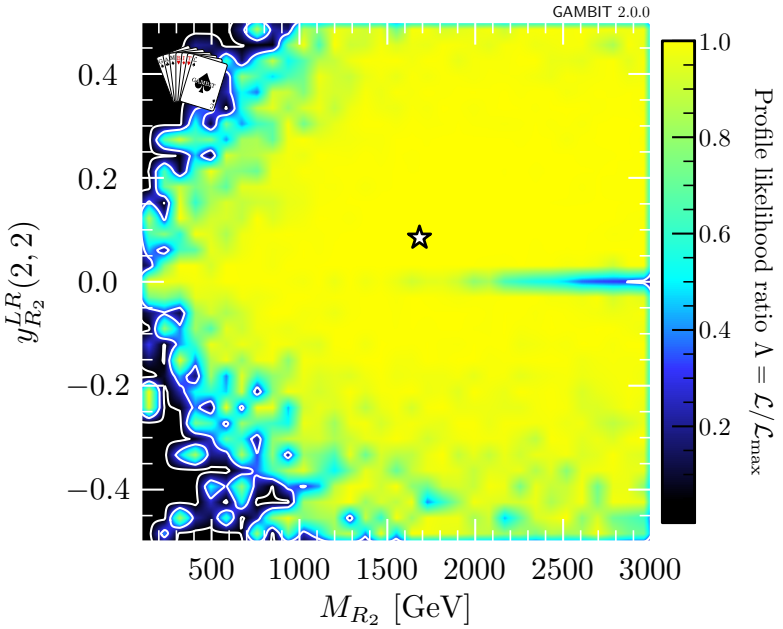


Figure 6.11: Profile likelihood plot of M_{R_2} and $(2,2)$ element of $y^{LR}_{R_2}$ in $R_2 + S_3$ scan.

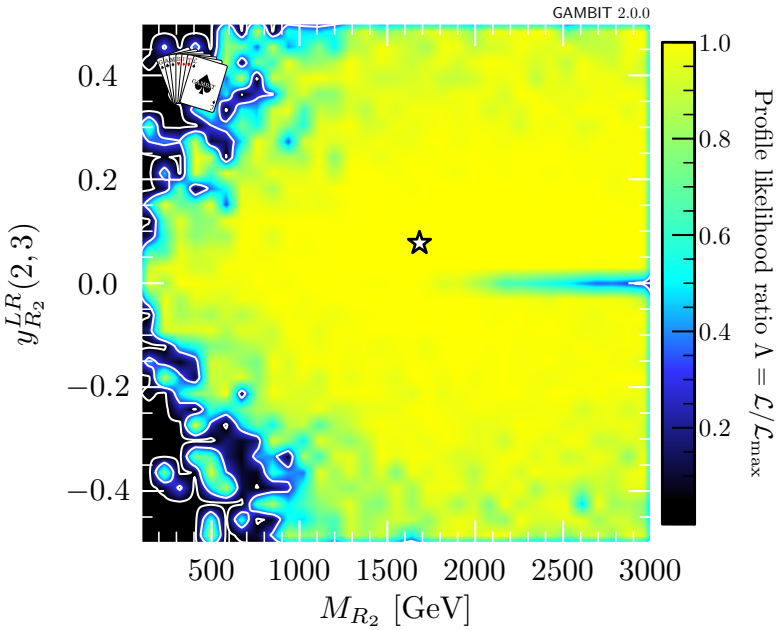


Figure 6.12: Profile likelihood plot of M_{R_2} and $(2,3)$ element of $y^{LR}_{R_2}$ in $R_2 + S_3$ scan.

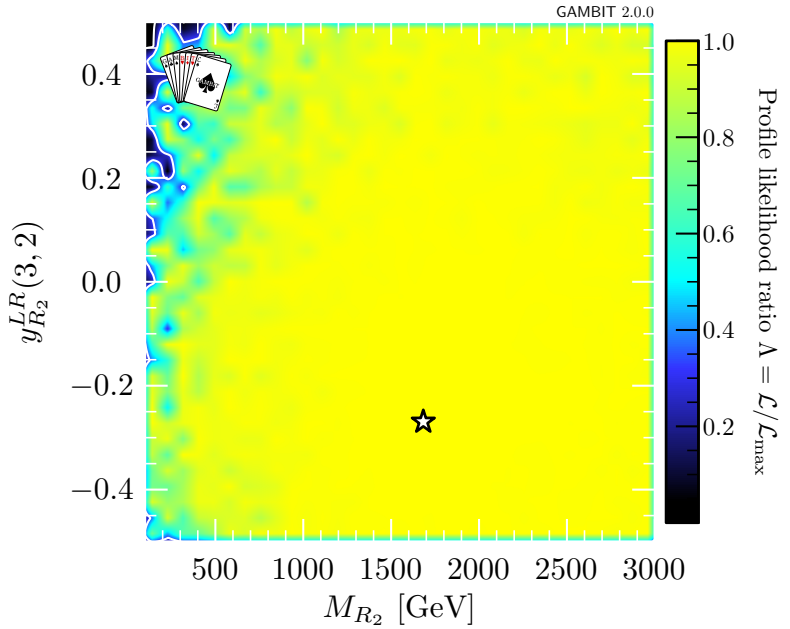


Figure 6.13: Profile likelihood plot of M_{R_2} and $(3,2)$ element of $y^{LR}_{R_2}$ in $R_2 + S_3$ scan.

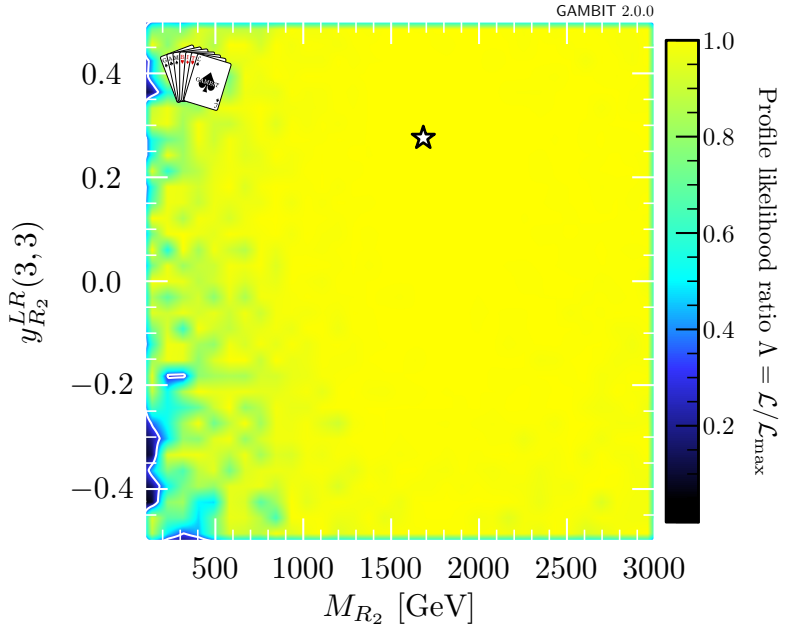


Figure 6.14: Profile likelihood plot of M_{R_2} and $(3,3)$ element of $y^{LR}_{R_2}$ in $R_2 + S_3$ scan.

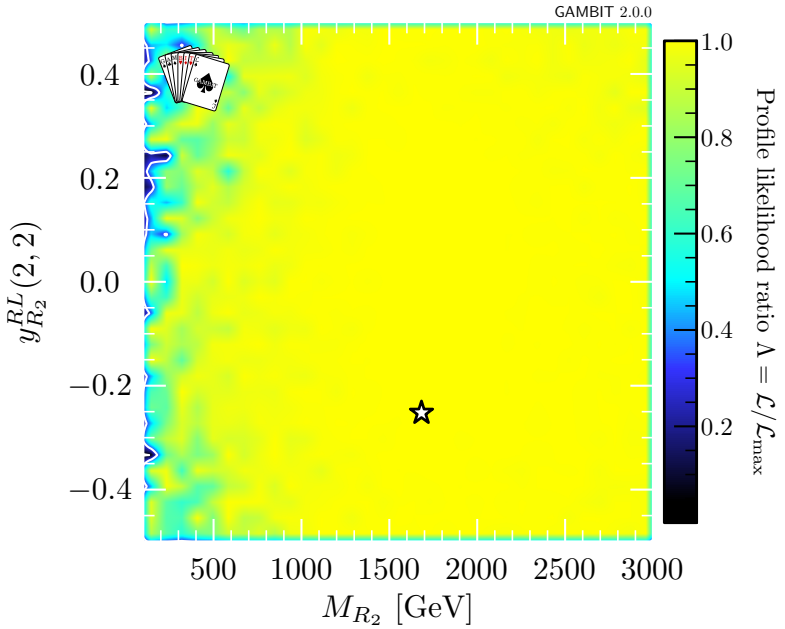


Figure 6.15: Profile likelihood plot of M_{R_2} and $(2,2)$ element of $y^{RL}_{R_2}$ in $R_2 + S_3$ scan.

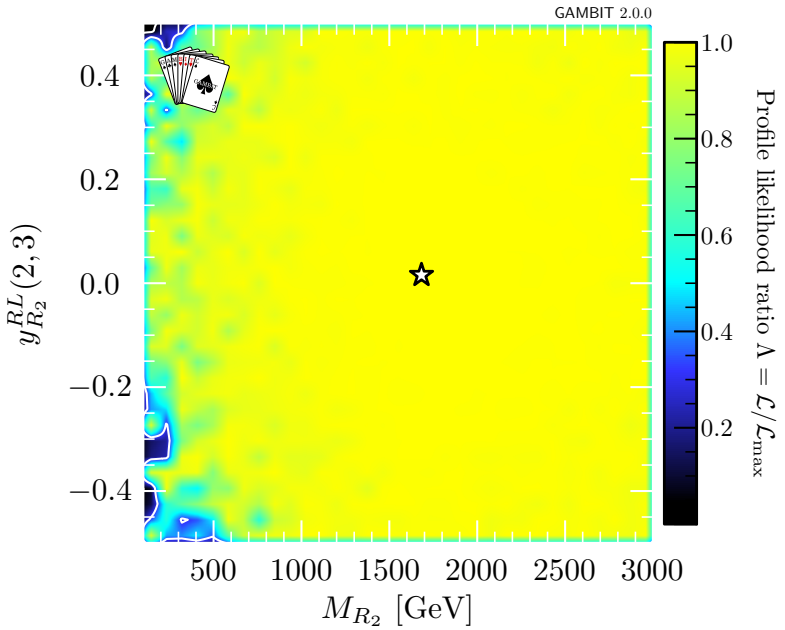


Figure 6.16: Profile likelihood plot of M_{R_2} and $(2,3)$ element of $y^{RL}_{R_2}$ in $R_2 + S_3$ scan.

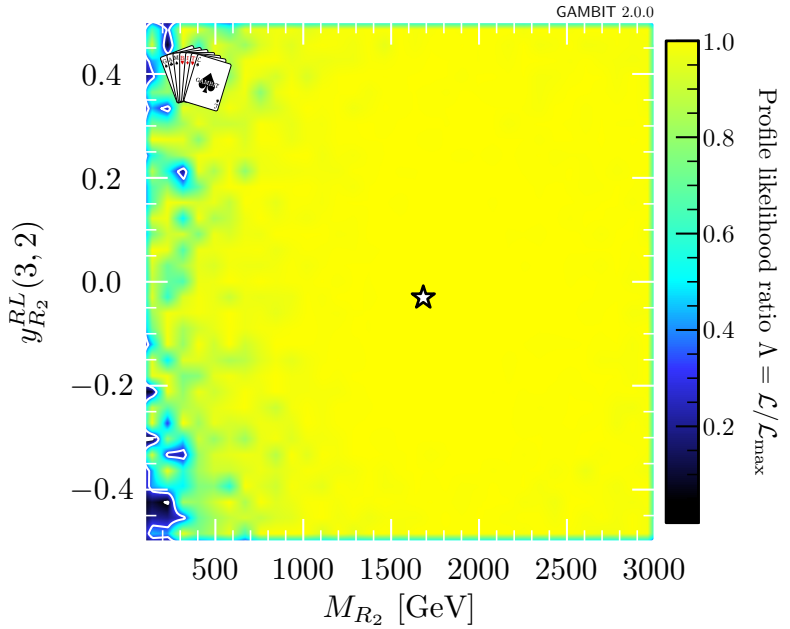


Figure 6.17: Profile likelihood plot of M_{R_2} and $(3,2)$ element of $y^{RL}_{R_2}$ in $R_2 + S_3$ scan.

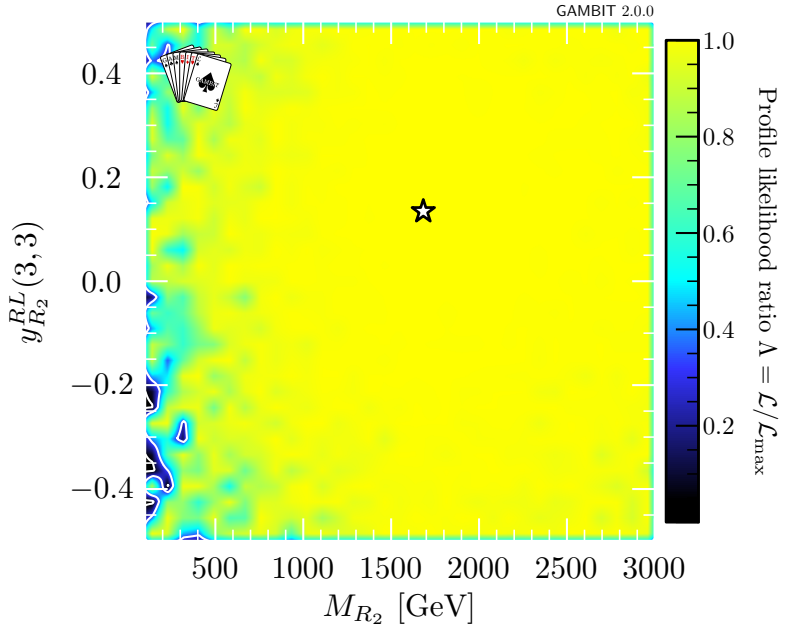


Figure 6.18: Profile likelihood plot of M_{R_2} and $(3,3)$ element of $y^{RL}_{R_2}$ in $R_2 + S_3$ scan.

We make the following observations about these results:

- Figure 6.3 shows a clear exclusion limit for $R_2 + S_3$ combinations where one or both of them has low mass.
- Figure 6.4 also shows that the inclusion of S_1 loosens the exclusion limits on a low mass S_3, as both LQs can give competing contributions to C_9 and C_{10}.
- There are almost no constraints on the masses and couplings of S_1 in this fit. These plots show very little structure, and so are consigned to the appendix (see Figures E.3 to E.10). These parameters would be more constrained by the inclusion of $R(D^{(*)})$ in a future scan.
- Figures 6.5, 6.7, 6.8 and E.1 show that a low mass S_3 is allowable if the $(2,2)$ and $(2,3)$ elements of $y^{LL}_{S_3}$ are small. The $(2,3)$ and $(3,3)$ elements are largely unconstrained by the likelihoods included in the fit, as seen in Figures 6.6, 6.9, 6.10 and E.2.
- Likewise, a low mass R_2 is allowable if the $(2,2)$ and $(3,2)$ elements of $y^{LR}_{R_2}$ are small, as seen in Figures 6.11 and 6.12.
- There is very little constraint on $y^{RL}_{R_2}$ in Figures 6.15 to 6.18.

For each point in the scan, we calculate the total cross-section using Pythia, as described in Section 6.3.2. Each cross-section was calculated using 1000 hard scattering Pythia events, resulting in uncertainties between one and two per cent. In $137\,\text{fb}^{-1}$ of data[3], we expect to observe a single event if the total cross-section is 7.3 ab. This value serves as a threshold for models which are unobservable with the current dataset. Figures 6.19 and 6.20 show the one-dimensional profile likelihood of cross-sections. Values of σ_{total} for which the profile likelihood ratio is above a chosen threshold may be considered not-excluded. As can be seen in these plots, there are points within the 1σ contour which have cross-sections well above the threshold, and as such are accessible at the LHC. The broad range of allowed cross-sections for the $S_1 + S_3$ model is largely due to lack of constraint on the mass and couplings of S_3.

3: This is the current size of the LHC dataset.

Finally, we will briefly discuss the constraints from LHC searches detailed in Section 6.3.2. We first select points within the 1σ contour using Equation (6.4). Of these points, we then select the hundred highest cross-sections for each model, and run the collider simulation on these points. As explained earlier, the LHC log-likelihood is capped at zero, on the assumption that any signal region excesses are due to statistical fluctuations. Using Equation (6.4) with the LHC likelihoods, we find that 96% of the selected $R_2 + S_3$ models and 72% of the selected $S_1 + S_3$ models were not excluded by the LHC searches. These points indicate that there are viable regions of parameter space which are not yet being targeted in LHC searches. A full scan would include the LHC likelihoods in the

total likelihood, rather than calculating only a few points; such a scan is left as future work.

Tables 6.1 and 6.2 show the parameters of the four points with the highest cross-section points which are not excluded by flavour or LHC constraints. The combined LHC likelihoods, total LQ production cross-sections, and expected numbers of events are also shown in the tables. The large cross-sections shown in Table 6.2 are due to the lack of constraints on the parameters of S_1, which allows for small values of M_{S_1}.

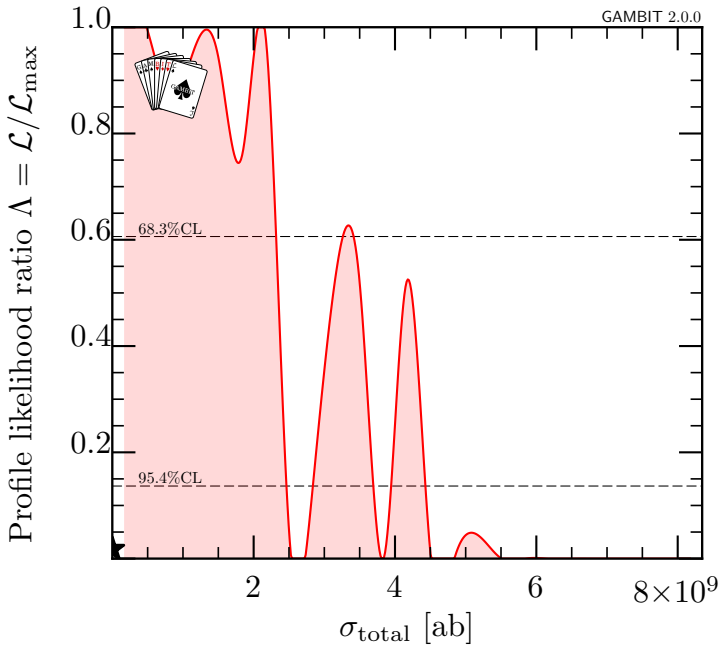


Figure 6.19: Profile likelihood of total cross-section of LQ processes at $\sqrt{s}$ = 13 TeV for $R_2 + S_3$ scan. Values of σ_{total} for which Λ is above a chosen threshold may be considered non-excluded.

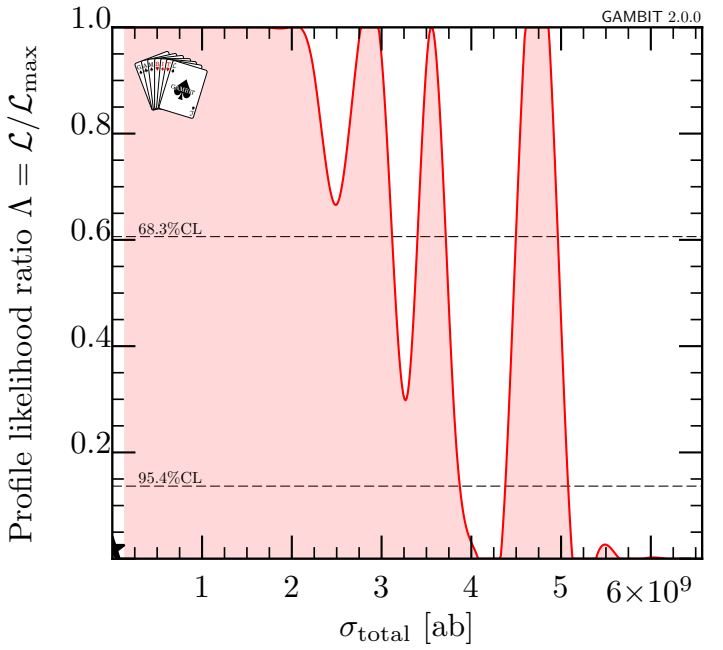


Figure 6.20: Profile likelihood of total cross-section of LQ processes at $\sqrt{s}$ = 13 TeV for $S_1 + S_3$ scan. Values of σ_{total} for which Λ is above a chosen threshold may be considered non-excluded.

Param/observable	Point 1	Point 2	Point 3	Point 4
$\ln\left(\mathscr{L}_{\text{LHC, capped}}\right)$	0	0	0	−1.555
σ_{total} [ab]	1.09×10^4	8.50×10^3	7.54×10^3	7.53×10^3
$n_{\text{expected events}}$	1486	1165	1033	1032
M_{R_2} [GeV]	1770.8	1121.7	1320.9	1158.5
M_{S_3} [GeV]	1132.8	1637.0	2798.5	1487.8
$y_{S_3}^{LL}(2,2)$	0.035	0.072	−0.127	−0.018
$y_{S_3}^{LL}(2,3)$	0.256	0.237	0.207	0.154
$y_{S_3}^{LL}(3,2)$	−0.014	−0.009	0.08	0.143
$y_{S_3}^{LL}(3,3)$	−0.073	−0.274	−0.332	0.284
$y_{R_2}^{LR}(2,2)$	0.073	0.023	0.434	−0.027
$y_{R_2}^{LR}(2,3)$	0.013	0.104	0.019	−0.168
$y_{R_2}^{LR}(3,2)$	−0.215	−0.097	0.042	−0.268
$y_{R_2}^{LR}(3,3)$	0.262	−0.005	0.172	0.157
$y_{R_2}^{RL}(2,2)$	−0.142	−0.209	−0.488	−0.196
$y_{R_2}^{RL}(2,3)$	0.102	0.0	0.198	0.156
$y_{R_2}^{RL}(3,2)$	0.424	0.068	0.026	0.021
$y_{R_2}^{RL}(3,3)$	−0.014	0.06	0.179	0.295

Table 6.1: Parameters and LHC likelihoods of the four highest cross-section points in the $R_2 + S_3$ scan which are not excluded by the implemented flavour constraints or LHC searches. The likelihood from LHC searches is capped at the SM expectation.

Param/observable	Point 1	Point 2	Point 3	Point 4
$\ln\left(\mathscr{L}_{\text{LHC, capped}}\right)$	0.0	−1.575	−0.498	−3.547
σ_{total} [ab]	6.66×10^8	2.62×10^8	2.04×10^8	1.50×10^8
$n_{\text{expected events}}$	9.12×10^7	3.59×10^7	2.80×10^7	2.05×10^7
M_{S_1} [GeV]	126.8	156.8	164.9	176.3
M_{S_3} [GeV]	1954.5	2680.1	2245.1	2898.6
$y_{S_1}^{LL}(2,2)$	0.203	−0.178	0.063	−0.071
$y_{S_1}^{LL}(2,3)$	−0.466	−0.013	0.247	−0.134
$y_{S_1}^{LL}(3,2)$	0.142	−0.137	0.088	0.464
$y_{S_1}^{LL}(3,3)$	−0.098	0.294	0.053	−0.043
$y_{S_1}^{RR}(2,2)$	−0.157	−0.296	−0.12	0.097
$y_{S_1}^{RR}(2,3)$	0.411	−0.124	−0.469	−0.476
$y_{S_1}^{RR}(3,2)$	0.356	−0.065	0.238	−0.384
$y_{S_1}^{RR}(3,3)$	−0.039	−0.457	−0.244	−0.343
$y_{S_3}^{LL}(2,2)$	0.422	0.034	−0.027	0.204
$y_{S_3}^{LL}(2,3)$	−0.21	−0.131	0.491	−0.095
$y_{S_3}^{LL}(3,2)$	−0.002	−0.013	0.02	−0.011
$y_{S_3}^{LL}(3,3)$	−0.048	0.214	−0.104	−0.448

Table 6.2: Parameters and LHC likelihoods of the four highest cross-section points in the S_1+S_3 scan which are not excluded by the implemented flavour constraints or LHC searches. The likelihood from LHC searches is capped at the SM expectation.

6.5. Conclusion

In this Chapter, we performed a global fit to two LQ models. The likelihoods used were the experimental likelihoods for the $B^0_{(s)} \to \mu\mu$ and $B^0_s \to \phi\mu\mu$ branching fractions, and the recent $R(K^{(*)})$ LHCb measurement. We identified regions of the parameter space which are consistent with these measurements. With the results of the fit, we performed a brief proof-of-concept study of a full scan with collider constraints. We calculated the total LQ production cross-section in order to determine which models are observable in the current LHC dataset. Of the non-excluded models, we selected the highest cross-section points and tested them against three ATLAS analyses I implemented in ColliderBit, showing which of these are excluded by direct searches.

This is a preliminary study that lays the groundwork for a more complete analysis. The future directions for this analysis involve the addition of new observables in GAMBIT and the inclusion of existing ones in a larger scan. On the flavour side, FlavBit already includes angular observables for $B \to K^{(*)}\mu\mu$ decays, but these were not included due to time constraints. Future scans should also include $R(D^{(*)})$, which is not currently calculated by SuperIso or FlavBit. On the collider side, this analysis can be extended by including additional direct searches by ATLAS and CMS, and by running a global fit with these likelihoods activated.

Part III.

Conclusion

7. Conclusion

7. Conclusion

Flavour physics offers the most tantalising hints of physics beyond the Standard Model, at a time when direct evidence of any particular BSM model is not forthcoming. The flavour anomalies are sensitive to contributions from high-energy physics, while being accessible at low energies. This thesis details work on two aspects of the ongoing search for BSM physics: precision measurements of SM parameters, and global fits with flavour constraints to make conclusions about specific BSM models.

We measured the branching fraction of $B \to \pi \ell \nu$ relative to $B \to D^* \ell \nu$ using SL FEI tagging. The measurement was performed using $189.9\,\text{fb}^{-1}$ of e^+e^- collisions recorded by the Belle II detector. From a combined fit to the B^+ and B^0 modes, we obtained a result of $|V_{ub}| = (3.598 \pm 0.266) \times 10^{-3}$, which is consistent with the world average. The uncertainty is largely statistical, and so will be greatly reduced in the future by larger datasets. Furthermore, there is ongoing work in improving the performance of the FEI and reducing data-MC discrepancies, which may further improve the precision of this measurement. It is anticipated that measurements of $|V_{ub}|$ at Belle II will lead to better understanding of the discrepancy between inclusive and exclusive determinations.

This is one of the first Belle II analyses to use the SL FEI. The SL FEI has been underutilised compared to the hadronic FEI, due to ongoing challenges in developing a calibration procedure. The SL FEI offers a larger tagging efficiency than the hadronic FEI, so there is considerable interest within the collaboration for finding ways to use it. This analysis demonstrates the viability of the normalisation mode approach, which allows us to obtain meaningful physics results in the absence of calibration. The main implication is that other ratio measurements such as $R(D^{(*)})$ could take advantage of SL FEI tagging.

Belle II is currently in the early days of its data collection. Its dataset size will rapidly increase after the long shutdown, with the goal of $50\,\text{ab}^{-1}$ by 2030. This thesis details the software I developed for the `skim` package of basf2. My work greatly reduced the amount of boilerplate code in the package, and created a more straightforward framework for defining skims. This work helps to prepare Belle II for the massive datasets of the coming years, when analysts will be encouraged to exclusively analyse skimmed data.

Finally, we used GAMBIT to perform a global fit of the LQ models $R_2 + S_3$ and $S_1 + S_3$. I implemented the models in GAMBIT, wrote additional code to calculate flavour observables, and implemented three direct searches from ATLAS in ColliderBit. The flavour constraints used in the fit were the $B^0_{(s)} \to \mu\mu$ and $B^0_s \to \phi\mu\mu$ branching fractions, and the

recent $R(K^{(*)})$ result from LHCb which removes a previous long-standing anomaly. We identified regions of the LQ parameter space which are compatible with these measurements, and have a sufficiently high cross-section to be observed at the LHC. This is a preliminary study, and can be straightforwardly extended in the future with additional flavour constraints and a full scan of the implemented collider searches.

Appendices

A. Form factors parameters

A.1. B→πℓν

$B \to \pi \ell \nu$ form factors are described by the BCL parametrisation, which is defined in Reference [132]. Table XIX of Reference [30] shows the results of a fit combining lattice and experimental data. The nominal values are shown in Table A.1, and their correlation matrix is shown in Table A.2. In Chapter 5, we use these values to constrain the fit for extracting $|V_{ub}|$ from partial branching fraction ratios.

[132]: Bourrely et al. (2010), "Model-independent description of $B \to \pi l \nu$ decays and a determination of $|V_{ub}|$"

[30]: Fermilab Lattice Collaboration et al. (2015), "$|V_{ub}|$ from $B \to \pi \ell \nu$ decays and (2+1)-flavor lattice QCD"

Table A.1: Nominal values and uncertainties of BCL parameters.

Parameter	Nominal value
b_0^+	0.419 ± 0.013
b_1^+	-0.495 ± 0.054
b_2^+	-0.43 ± 0.13
b_3^+	0.22 ± 0.31
b_0^0	0.510 ± 0.019
b_1^0	-1.700 ± 0.082
b_2^0	1.53 ± 0.19
b_3^0	4.52 ± 0.83

Table A.2: Correlation matrix for BCL parameter uncertainties.

	b_0^+	b_1^+	b_2^+	b_3^+	b_0^0	b_1^0	b_2^0	b_3^0
b_0^+	1	0.140	−0.455	−0.342	0.224	0.174	0.047	−0.033
b_1^+		1	−0.789	−0.874	−0.068	0.142	0.025	−0.007
b_2^+			1	0.879	−0.051	−0.253	0.098	0.234
b_3^+				1	0.076	0.038	0.018	−0.200
b_0^0					1	−0.043	−0.604	−0.388
b_1^0						1	−0.408	−0.758
b_2^0							1	0.457
b_3^0								1

A.2. B→D*ℓν

$B \to D^* \ell \nu$ decays are described by the CLN [133] and BGL [124] parametrisations. The MC samples used in this analysis were generated with the BGL parametrisation. The parameters used by the generator were from version 1 of Reference [125]; in Section 5.5.7, we update the MC to the parameters from version 3. Additionally, we estimate the systematic uncertainty due to form factor shape by reweighting the MC using a set of Gaussian variations. The original and updated central values are show

[133]: Caprini et al. (1998), "Dispersive Bounds on the Shape of $B \to D^{(*)} \ell \nu$ Form Factors"

[124]: Boyd et al. (1995), "Constraints on Form Factors For Exclusive Semileptonic Heavy to Light Meson Decays"

[125]: Ferlewicz et al. (2021), "Revisiting fits to $B^0 \to D^{*-} \ell^+ \nu_\ell$ to measure $|V_{cb}|$ with novel methods and preliminary LQCD data at non-zero recoil"

in Table A.3, and the correlation matrices of statistical and systematic uncertainties are shown in Tables A.4 and A.5.

Table A.3: Fitted parameters from BGL$(1,1,2)$ configuration. Statistical and systematic uncertainties are included for updated values used in reweighting.

Parameter	EvtGen values ($\times 10^3$)	Updated values ($\times 10^3$)
$\tilde{a}_0^g$	1.00	$1.00 \pm 0.02 \pm 0.02$
$\tilde{a}_1^g$	-2.33	$-2.35 \pm 0.61 \pm 0.66$
$\tilde{a}_0^f$	0.505	$0.511 \pm 0.004 \pm 0.013$
$\tilde{a}_1^f$	0.66	$0.67 \pm 0.17 \pm 0.30$
$\tilde{a}_0^{\mathcal{F}_1}$	0.29	$0.30 \pm 0.06 \pm 0.08$
$\tilde{a}_1^{\mathcal{F}_1}$	-3.6	$-3.68 \pm 1.26 \pm 1.20$
$\mathcal{F}(1)\eta_{\text{EW}}\lvert V_{cb}\rvert$	34.9	$35.28 \pm 0.24 \pm 0.87$

Table A.4: Correlation matrix for statistical uncertainties of updated parameters in Table A.3.

	$\tilde{a}_0^g$	$\tilde{a}_1^g$	$\tilde{a}_0^f$	$\tilde{a}_1^f$	$\tilde{a}_0^{\mathcal{F}_1}$	$\tilde{a}_1^{\mathcal{F}_1}$
$\tilde{a}_0^g$	1	-0.937	-0.128	0.069	-0.081	0.161
$\tilde{a}_1^g$		1	0.127	-0.222	0.110	-0.192
$\tilde{a}_0^f$			1	-0.800	-0.751	0.624
$\tilde{a}_1^f$				1	0.443	-0.354
$\tilde{a}_0^{\mathcal{F}_1}$					1	-0.978
$\tilde{a}_1^{\mathcal{F}_1}$						1

Table A.5: Correlation matrix for systematic uncertainties of updated parameters in Table A.3.

	$\tilde{a}_0^g$	$\tilde{a}_1^g$	$\tilde{a}_0^f$	$\tilde{a}_1^f$	$\tilde{a}_0^{\mathcal{F}_1}$	$\tilde{a}_1^{\mathcal{F}_1}$
$\tilde{a}_0^g$	1	-0.940	-0.132	0.085	-0.077	0.158
$\tilde{a}_1^g$		1	0.129	-0.228	0.107	-0.189
$\tilde{a}_0^f$			1	-0.806	-0.755	0.629
$\tilde{a}_1^f$				1	0.452	-0.362
$\tilde{a}_0^{\mathcal{F}_1}$					1	-0.977
$\tilde{a}_1^{\mathcal{F}_1}$						1

Using $\mathcal{F}(1) = 0.906$ and $\eta_{\text{EW}} = 1.0066$, the value of $|V_{cb}|$ in the updated parameters is 38.27×10^{-3}.

A.3. B→Dℓν

[126]: Belle Collaboration (2016), "Measurement of the decay $B \to D\ell\nu_\ell$ in fully reconstructed events and determination of the Cabibbo-Kobayashi-Maskawa matrix element $|V_{cb}|$"

$B \to D\ell\nu$ decays are described by the CLN [133] and BGL [124] parametrisations. The MC samples used in this analysis were generated with the BGL parametrisation. We use the central values from Reference [126], and the uncertainties and correlation matrix from a private fit by Florian Bernlochner. These are given in Table A.6 and Table A.7. In Section 5.5.7 we estimate the systematic uncertainty due to form factor shape by reweighting the MC using a set of Gaussian variations.

BGL parameter	Value
V_{cb}	$(41.18 \pm 1.15) \times 10^{-3}$
$a_{+,0}$	$1.26147 \times 10^{-2} \pm 9.78879 \times 10^{-5}$
$a_{+,1}$	$-9.62084 \times 10^{-2} \pm 3.34148 \times 10^{-3}$
$a_{+,2}$	$4.13884 \times 10^{-1} \pm 9.44471 \times 10^{-2}$
$a_{+,3}$	$-1.73699 \times 10^{-1} \pm 8.91809 \times 10^{-1}$

Table A.6: BGL parameters with uncertainties for $B \to D\ell\nu$ form factor.

	V_{cb}	$a_{+,0}$	$a_{+,1}$	$a_{+,2}$	$a_{+,3}$
V_{cb}	1	−0.402	−0.238	−0.110	0.047
$a_{+,0}$		1	0.245	−0.161	0.020
$a_{+,1}$			1	−0.654	0.272
$a_{+,2}$				1	−0.770
$a_{+,3}$					1

Table A.7: Correlation matrix for uncertainties of parameters in Table A.6.

B. Plots of charged track distributions

The following plots display the distributions of signal-side charged tracks, broken down by MC truth information. These tracks are identified using the ROE of the highest signal probability B_{tag} candidate. All plots are generated using a small subset of the full MC sample. The distributions are normalised, and are shown with all selections made *except* for the one being plotted.

The cuts on each particle are summarised in Table 5.3.

B.1. Electron

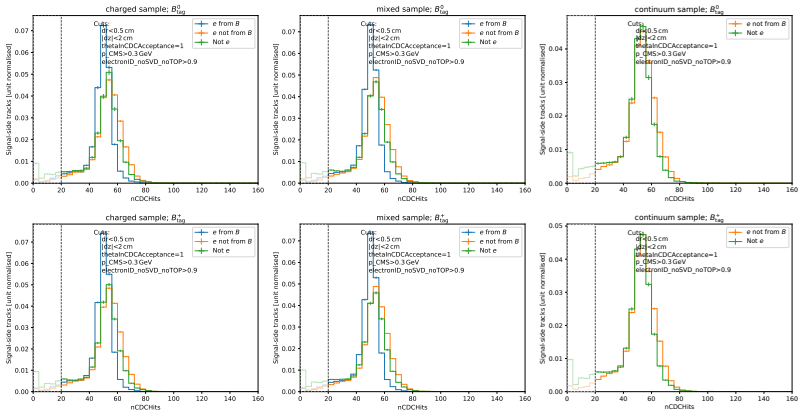


Figure B.1: `nCDCHits` distribution for signal-side charged tracks, separated by whether they are MC truth-matched to e^+.

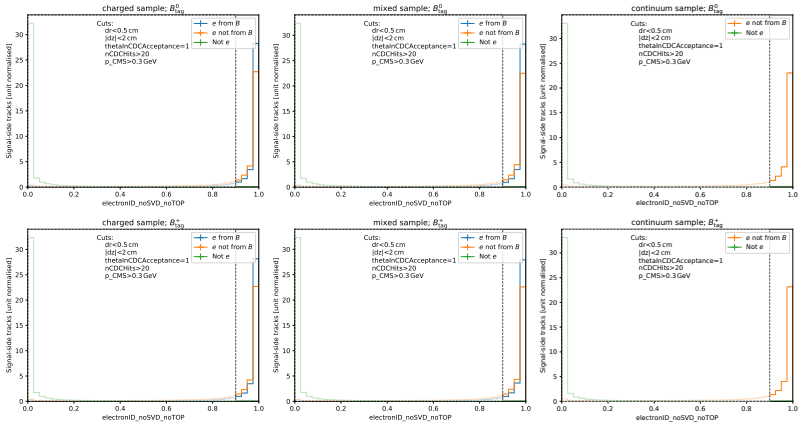


Figure B.2: `electronID` distribution for signal-side charged tracks, separated by whether they are MC truth-matched to e^+.

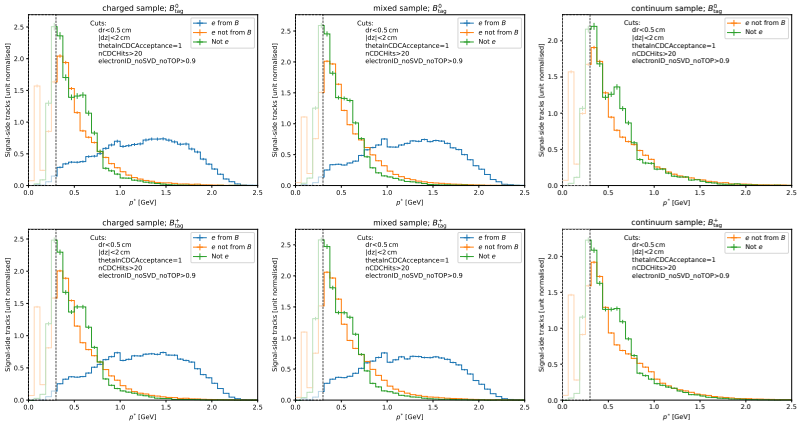


Figure B.3: p^* distribution for signal-side charged tracks, separated by whether they are MC truth-matched to e^+.

B.2. Muon

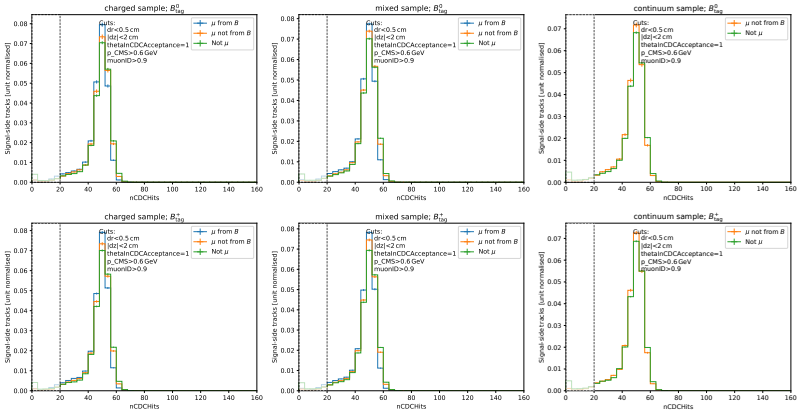


Figure B.4: `nCDCHits` distribution for signal-side charged tracks, separated by whether they are MC truth-matched to μ^+.

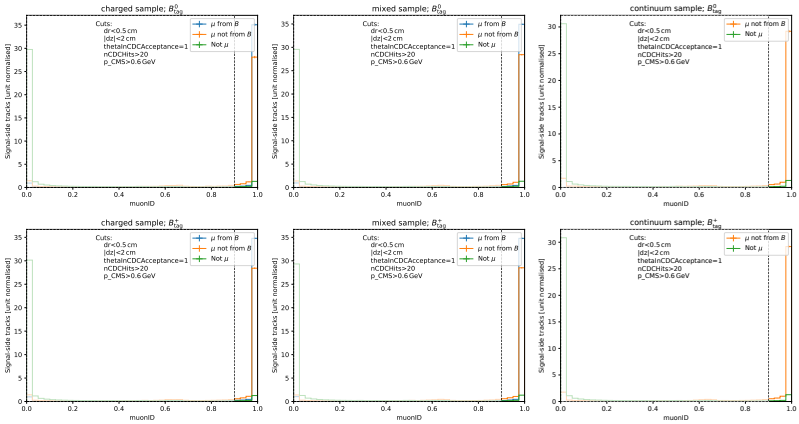


Figure B.5: `muonID` distribution for signal-side charged tracks, separated by whether they are MC truth-matched to μ^+.

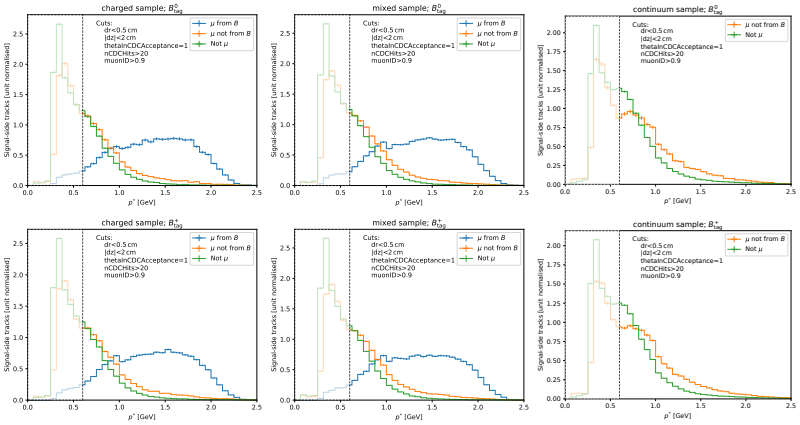


Figure B.6: p^* distribution for signal-side charged tracks, separated by whether they are MC truth-matched to μ^+.

B.3. Pion originating from B^+ or B^0

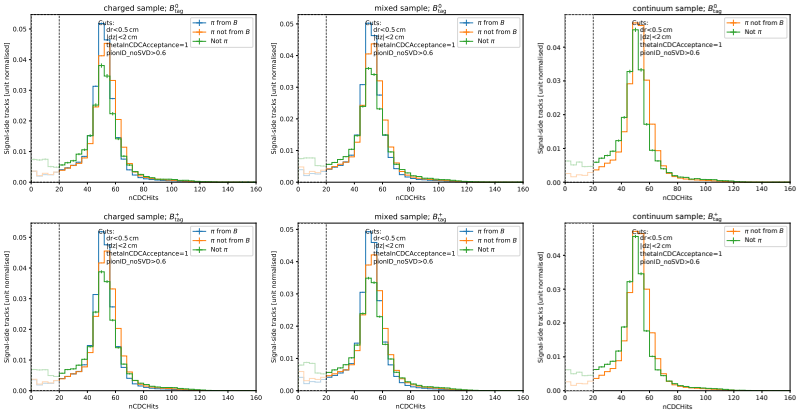


Figure B.7: `nCDCHits` distribution for signal-side charged tracks, separated by whether they are MC truth-matched to a π^+ originating from B^+ or B^0.

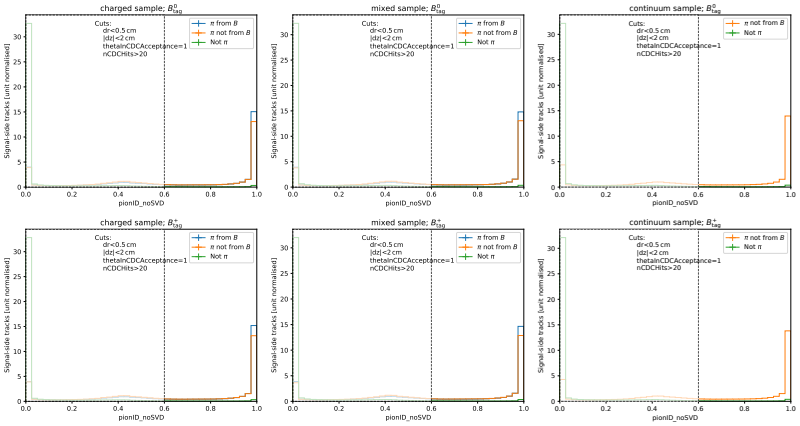


Figure B.8: `pionID` distribution for signal-side charged tracks, separated by whether they are MC truth-matched to a π^+ originating from B^+ or B^0.

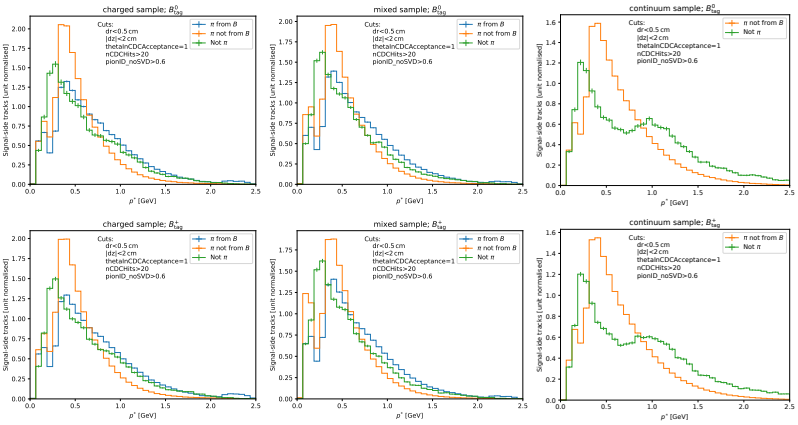


Figure B.9: p^* distribution for signal-side charged tracks, separated by whether they are MC truth-matched to a π^+ originating from B^+ or B^0. No analysis selection was chosen on this variable for this particle hypothesis.

B.4. Pion originating from D^0

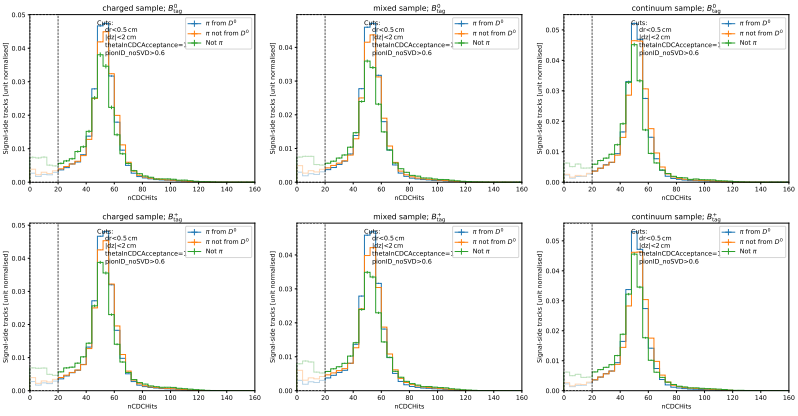


Figure B.10: `nCDCHits` distribution for signal-side charged tracks, separated by whether they are MC truth-matched to a π^+ originating from D^0.

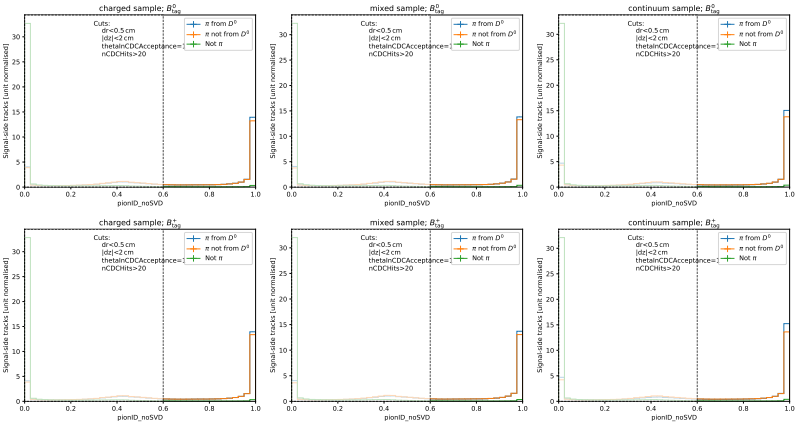


Figure B.11: `pionID` distribution for signal-side charged tracks, separated by whether they are MC truth-matched to a π^+ originating from D^0.

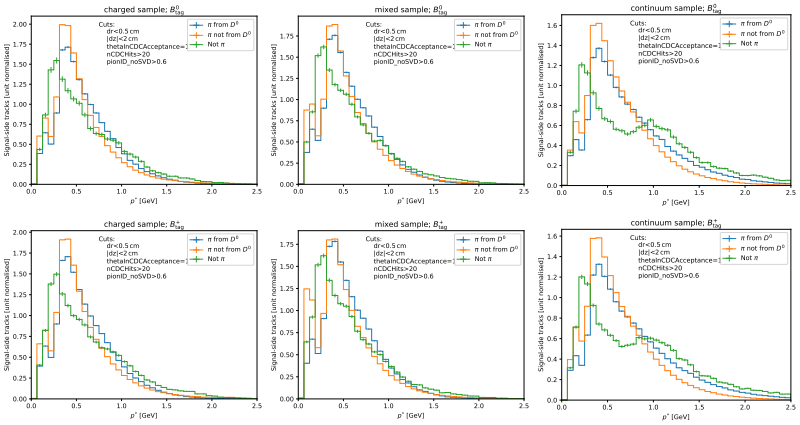


Figure B.12: p^* distribution for signal-side charged tracks, separated by whether they are MC truth-matched to a π^+ originating from D^0. No analysis selection was chosen on this variable for this particle hypothesis.

B.5. Kaon

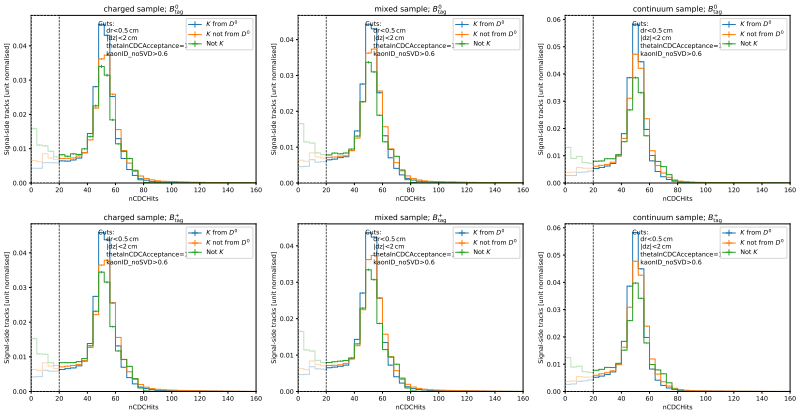


Figure B.13: `nCDCHits` distribution for signal-side charged tracks, separated by whether they are MC truth-matched to K^+.

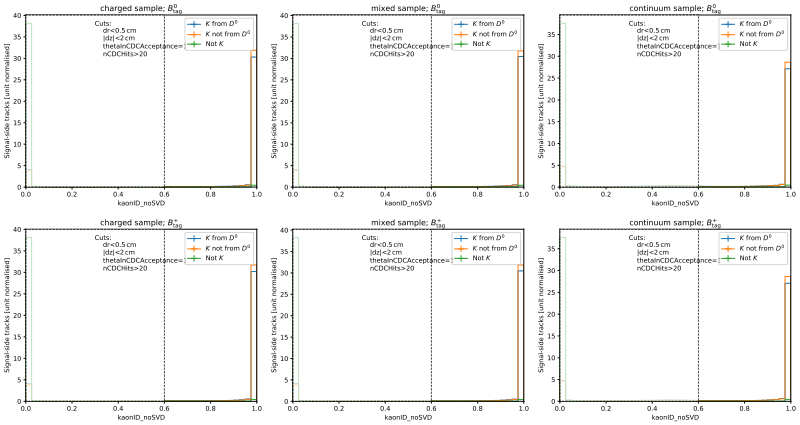


Figure B.14: `kaonID` distribution for signal-side charged tracks, separated by whether they are MC truth-matched to K^+.

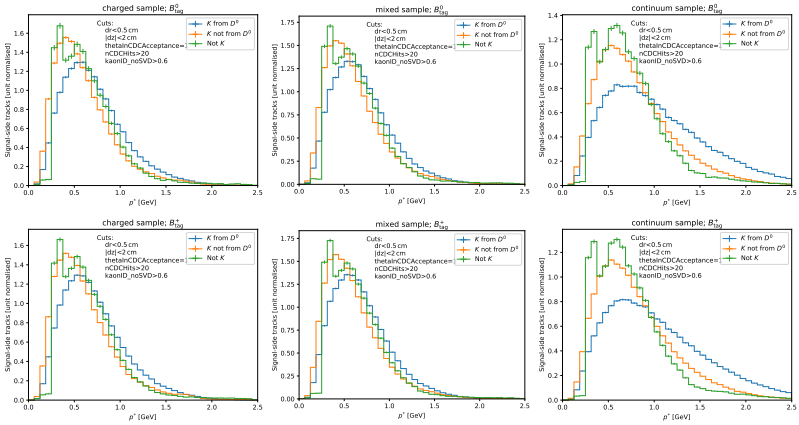


Figure B.15: p^* distribution for signal-side charged tracks, separated by whether they are MC truth-matched to K^+. No analysis selection was chosen on this variable for this particle hypothesis.

C. Plots of ROE photons

The following plots display the distributions of photons in the ROE of the $\Upsilon(4S)$ reconstruction, broken down by MC truth information. These photons are selected after applying the full online $\Upsilon(4S)$ reconstruction and cuts and randomly selecting one $\Upsilon(4S)$ candidate. All plots are generated using a small subset of the full MC sample. The distributions are normalised, and are shown with all selections made *except* for the one being plotted.

The selections on ROE objects are summarised in Table 5.8.

C.1. Forward region

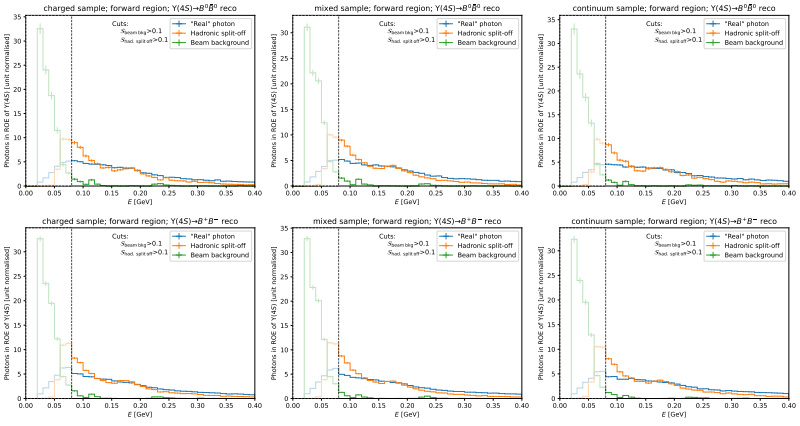


Figure C.1: E distribution for ROE photons in the forward region, separated by whether their origin in the generator.

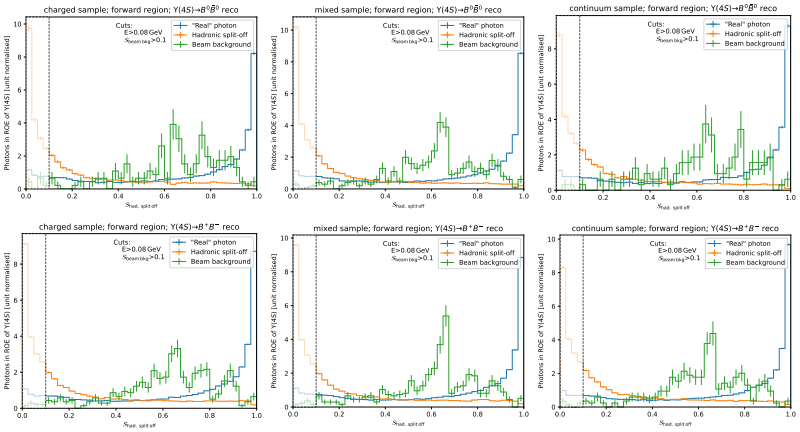


Figure C.2: `hadronicSplitOffSuppression` distribution for ROE photons in the forward region, separated by whether their origin in the generator.

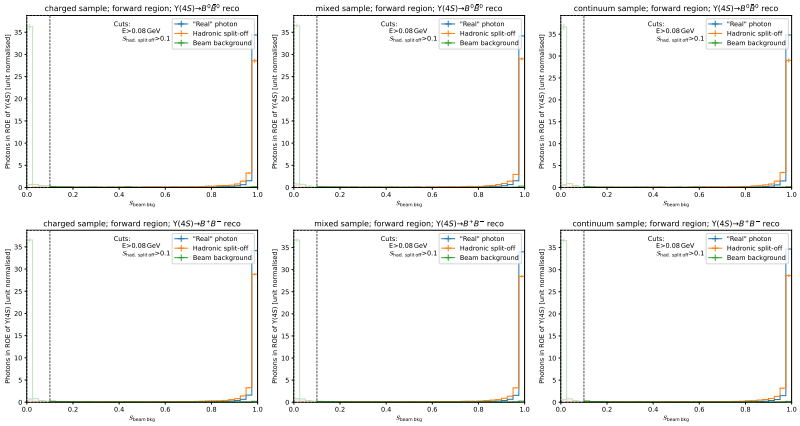


Figure C.3: `beamBackgroundSuppression` distribution for ROE photons in the forward region, separated by whether their origin in the generator.

C.2. Barrel region

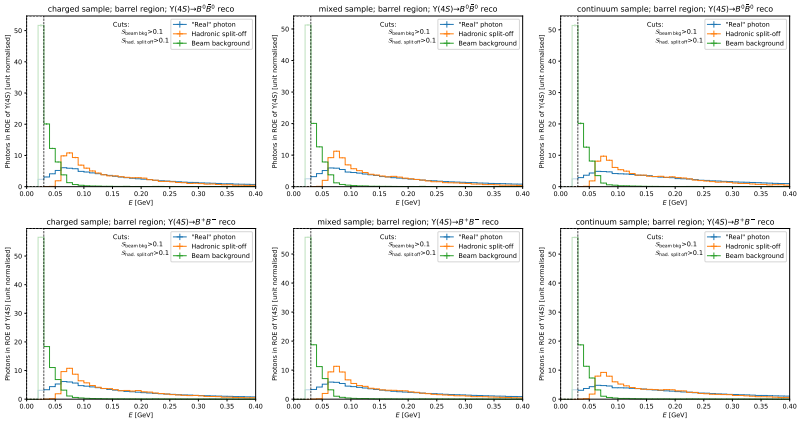


Figure C.4: E distribution for ROE photons in the barrel region, separated by whether their origin in the generator.

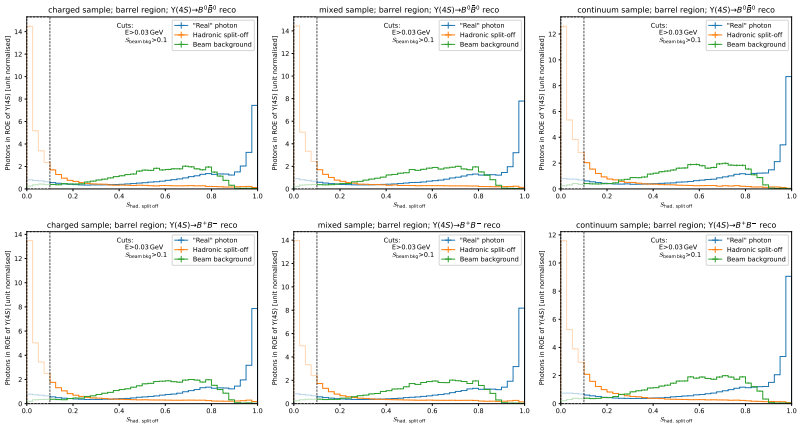


Figure C.5: `hadronicSplitOffSuppression` distribution for ROE photons in the barrel region, separated by whether their origin in the generator.

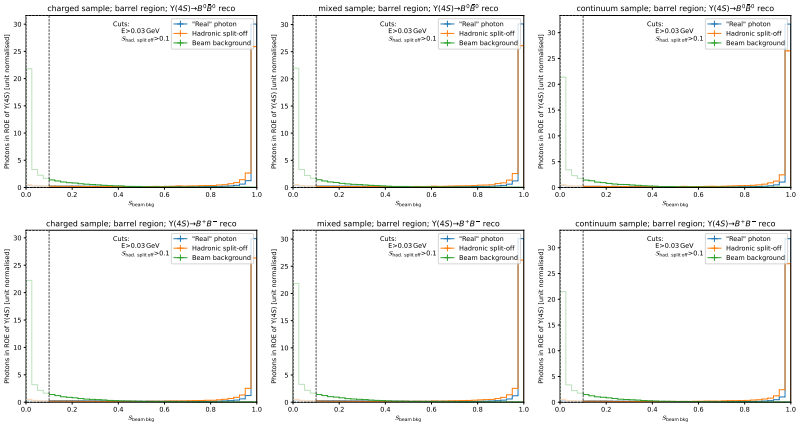


Figure C.6: `beamBackgroundSuppression` distribution for ROE photons in the barrel region, separated by whether their origin in the generator.

C.3. Backward region

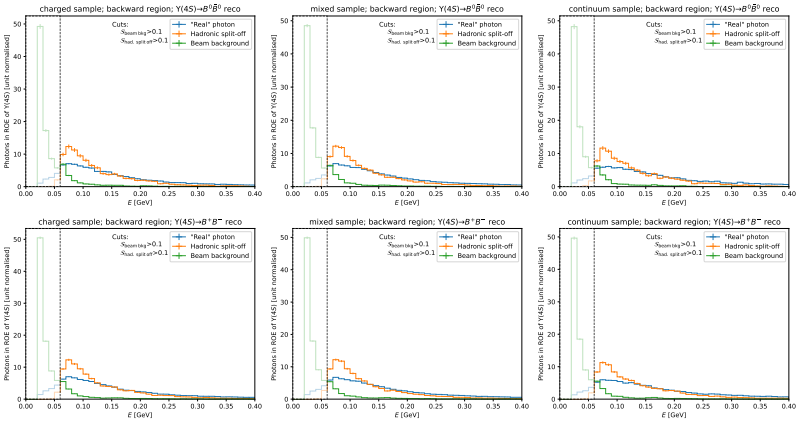


Figure C.7: E distribution for ROE photons in the backward region, separated by whether their origin in the generator.

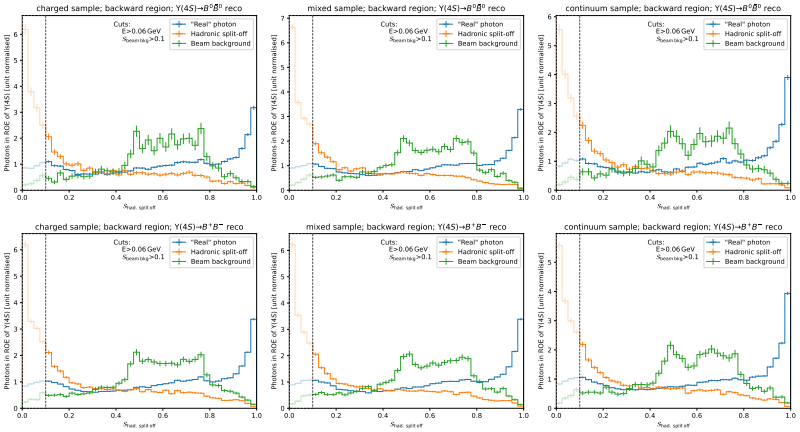


Figure C.8: `hadronicSplitOffSuppression` distribution for ROE photons in the backward region, separated by whether their origin in the generator.

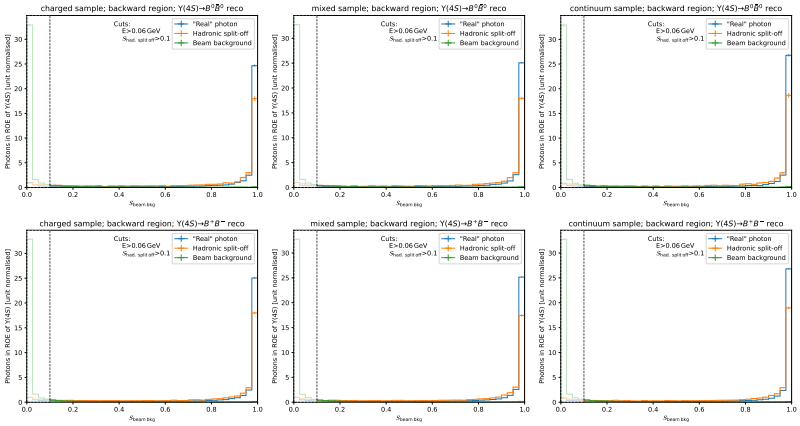


Figure C.9: `beamBackgroundSuppression` distribution for ROE photons in the backward region, separated by whether their origin in the generator.

Data-MC comparisons of kinematic variables

D.

This Section contains plots for checking data-MC agreement. The plots are shown are broken down by the flavour of the signal-side lepton. The signal modes are plotted in the side-band for each q^2 region. All efficiency corrections are included in these plots.

D.1. Signal mode

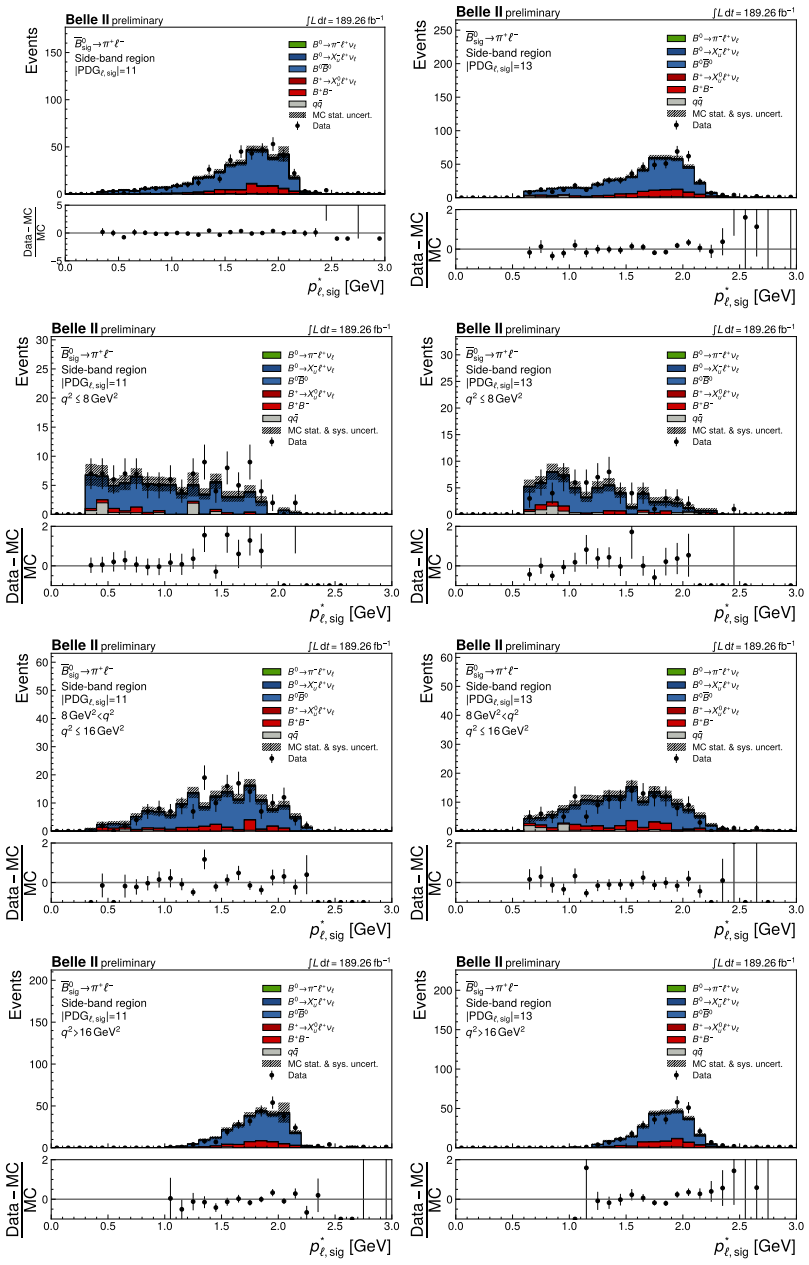


Figure D.1: Signal-side lepton momentum for B^0 signal mode, broken down by signal-side lepton flavour (electron in left column, muon in right column).

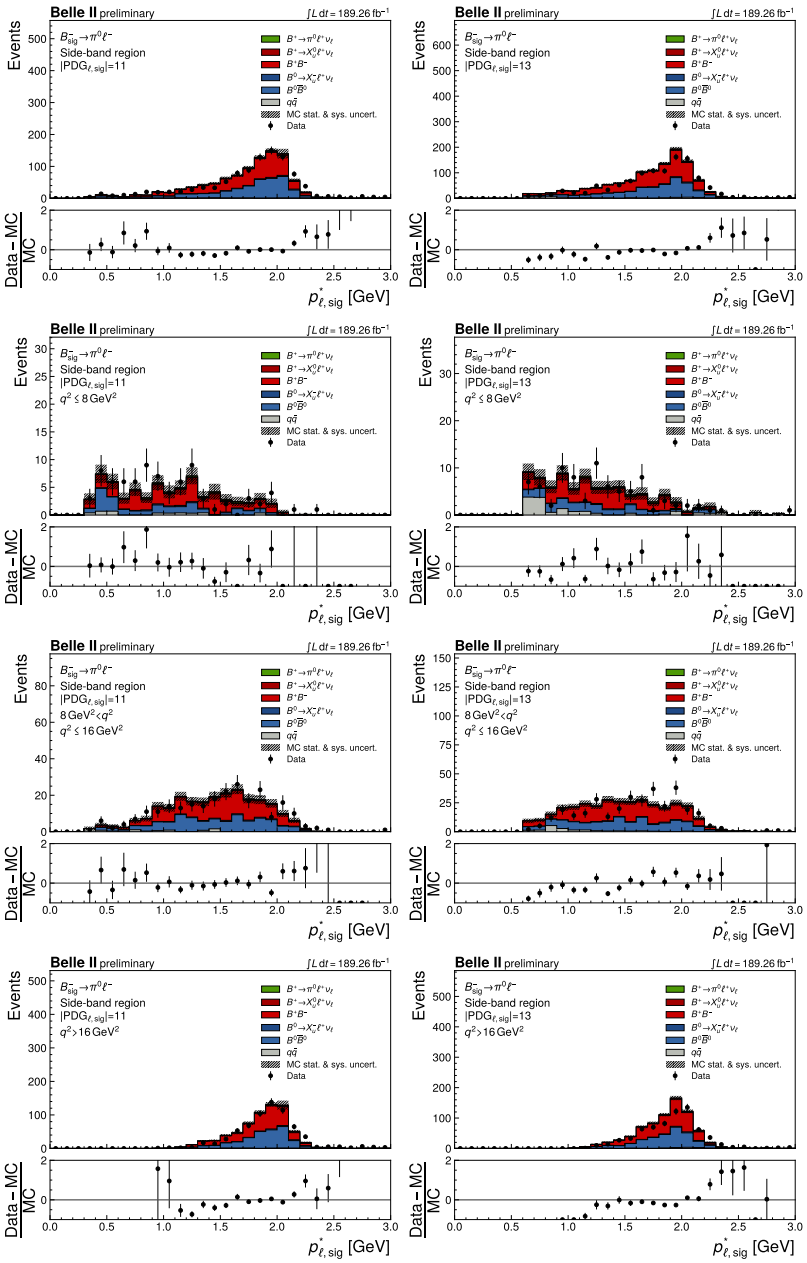


Figure D.2: Signal-side lepton momentum for B^+ signal mode, broken down by signal-side lepton flavour (electron in left column, muon in right column).

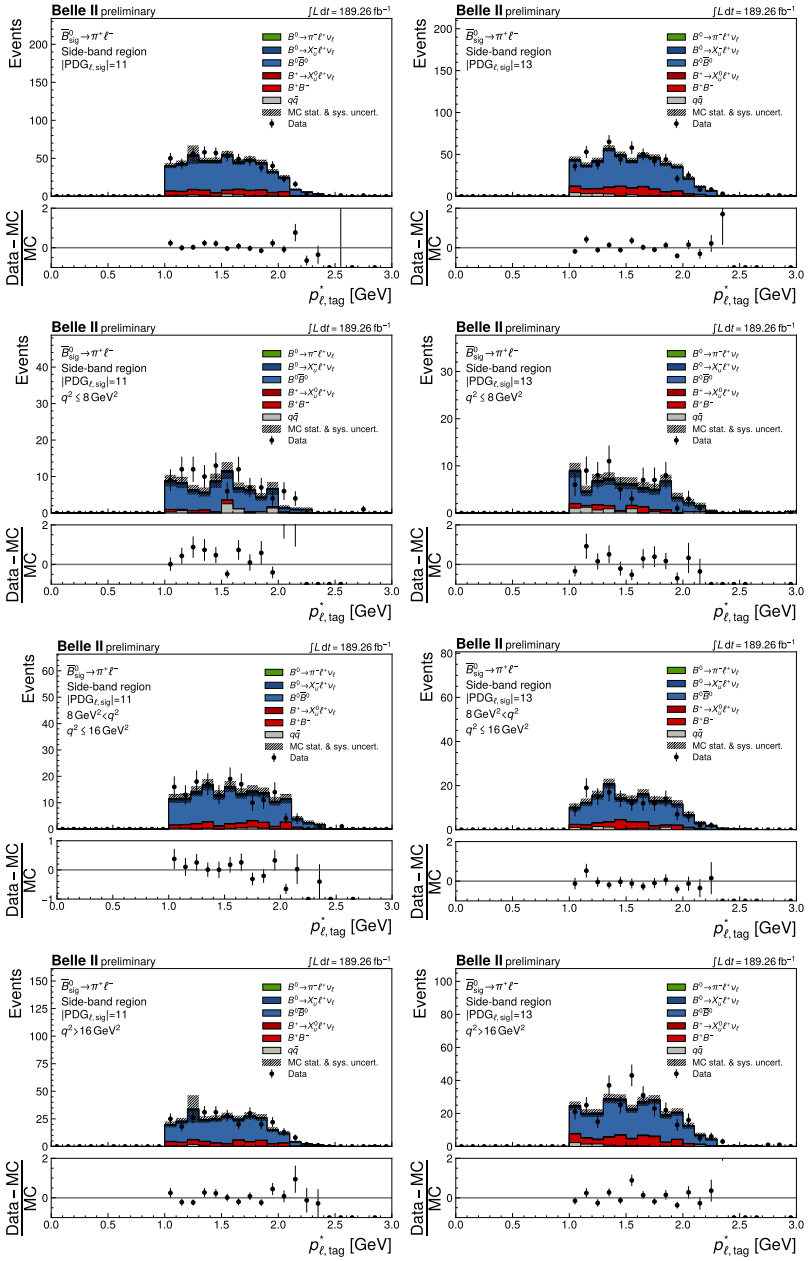


Figure D.3: Tag-side lepton momentum for B^0 signal mode, broken down by signal-side lepton flavour (electron in left column, muon in right column).

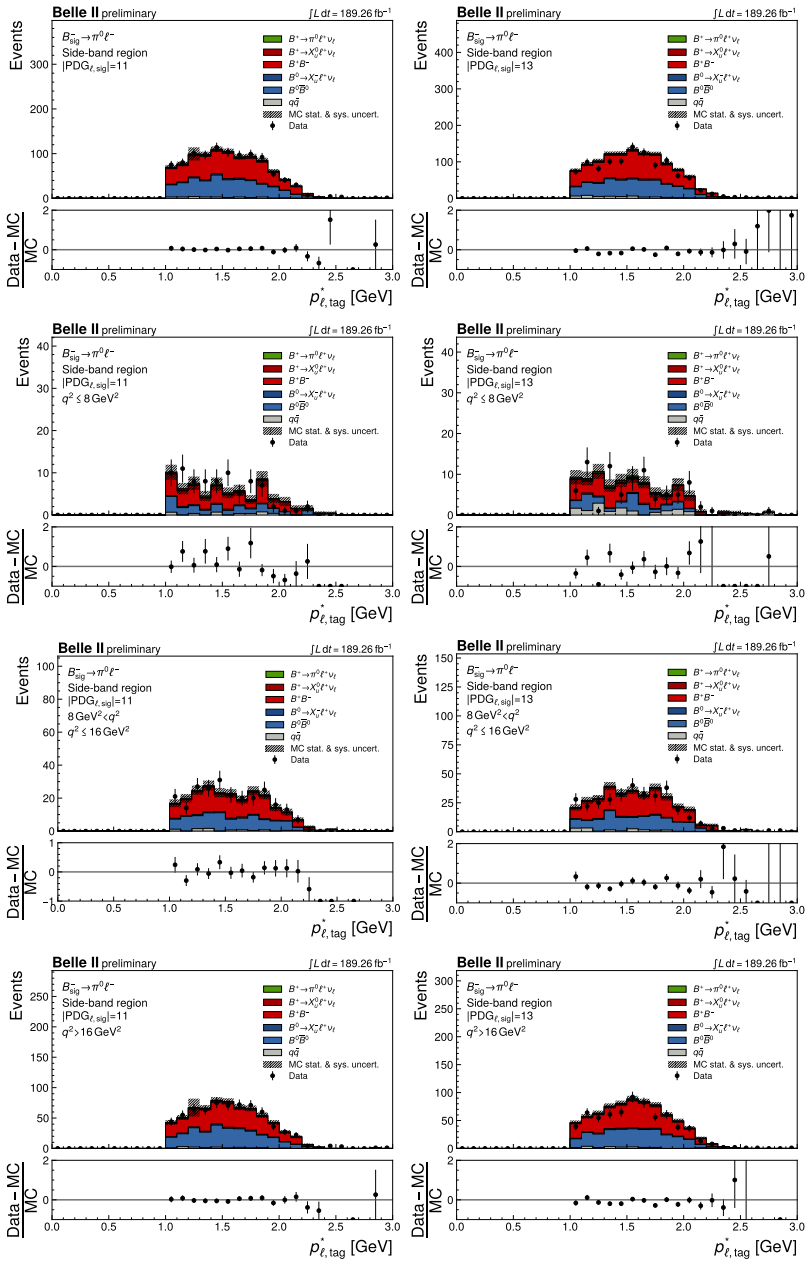


Figure D.4: Tag-side lepton momentum for B^+ signal mode, broken down by signal-side lepton flavour (electron in left column, muon in right column).

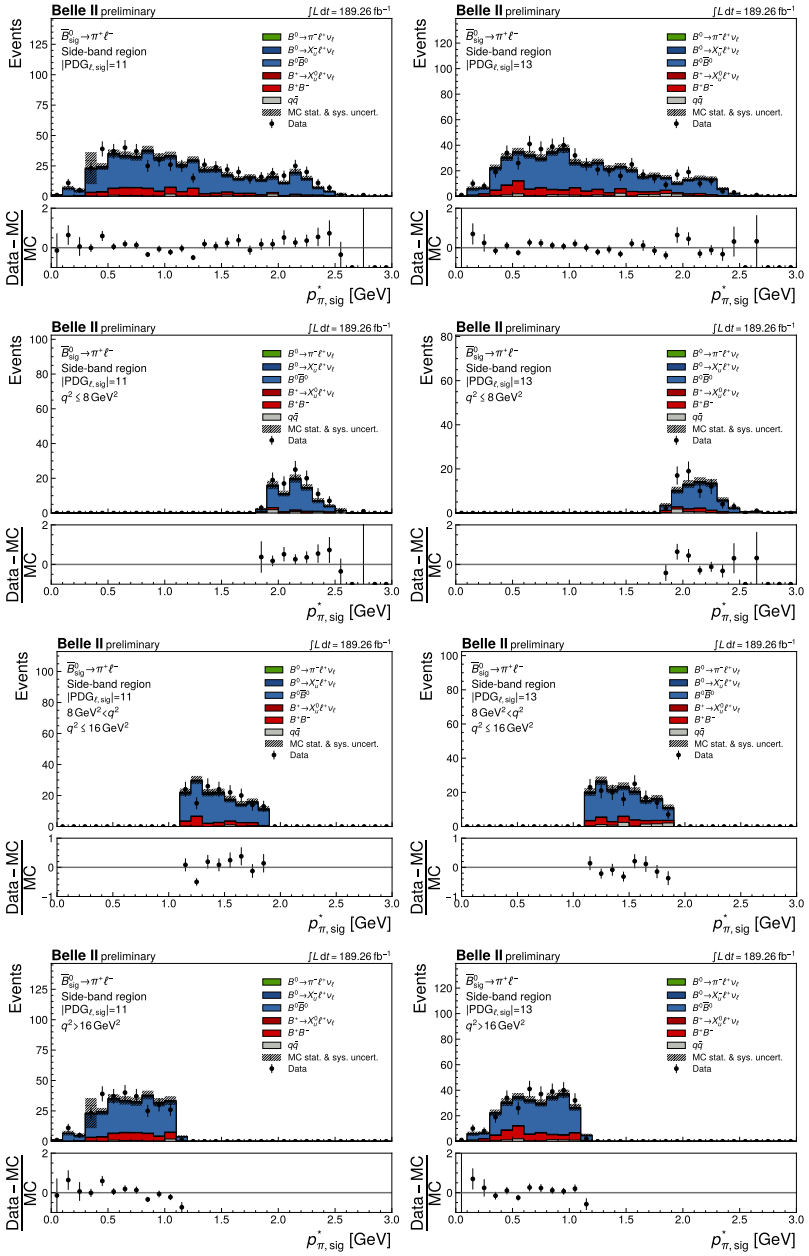


Figure D.5: Signal-side pion momentum for B^0 signal mode, broken down by signal-side lepton flavour (electron in left column, muon in right column).

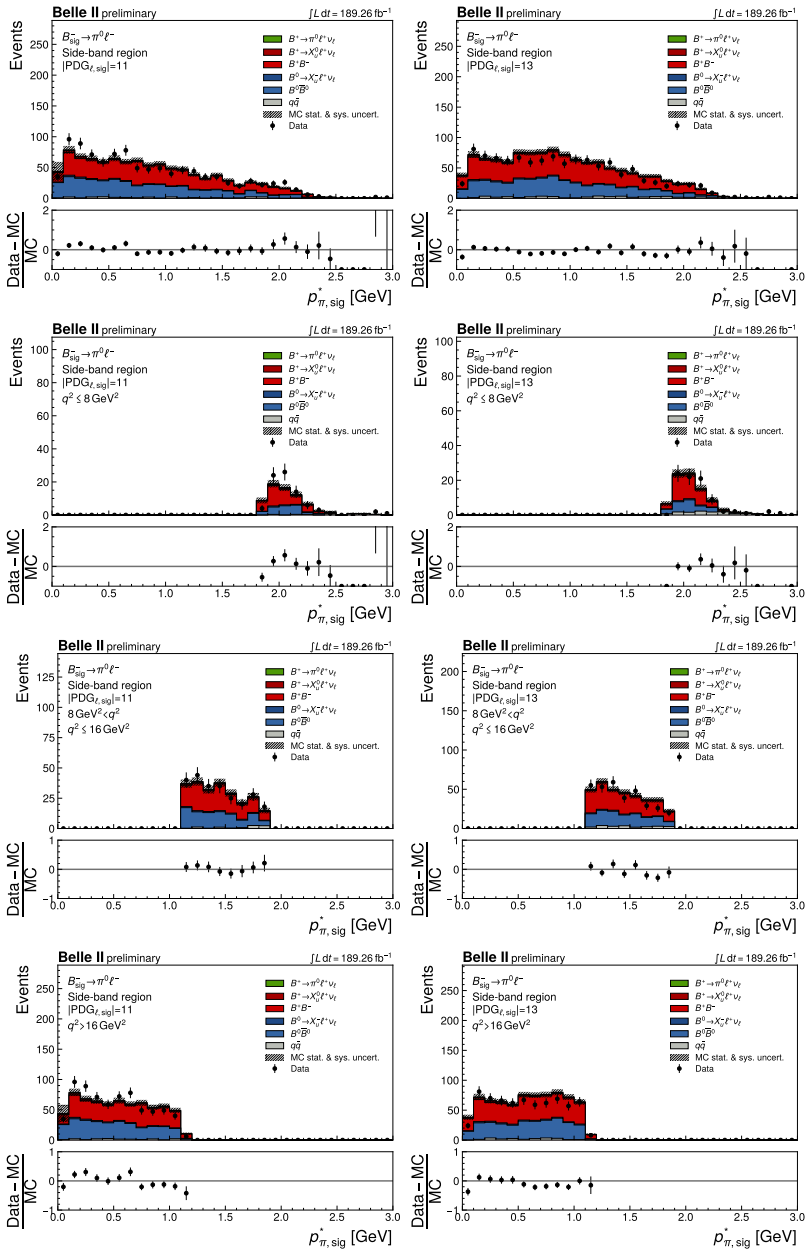


Figure D.6: Signal-side pion momentum for B^+ signal mode, broken down by signal-side lepton flavour (electron in left column, muon in right column).

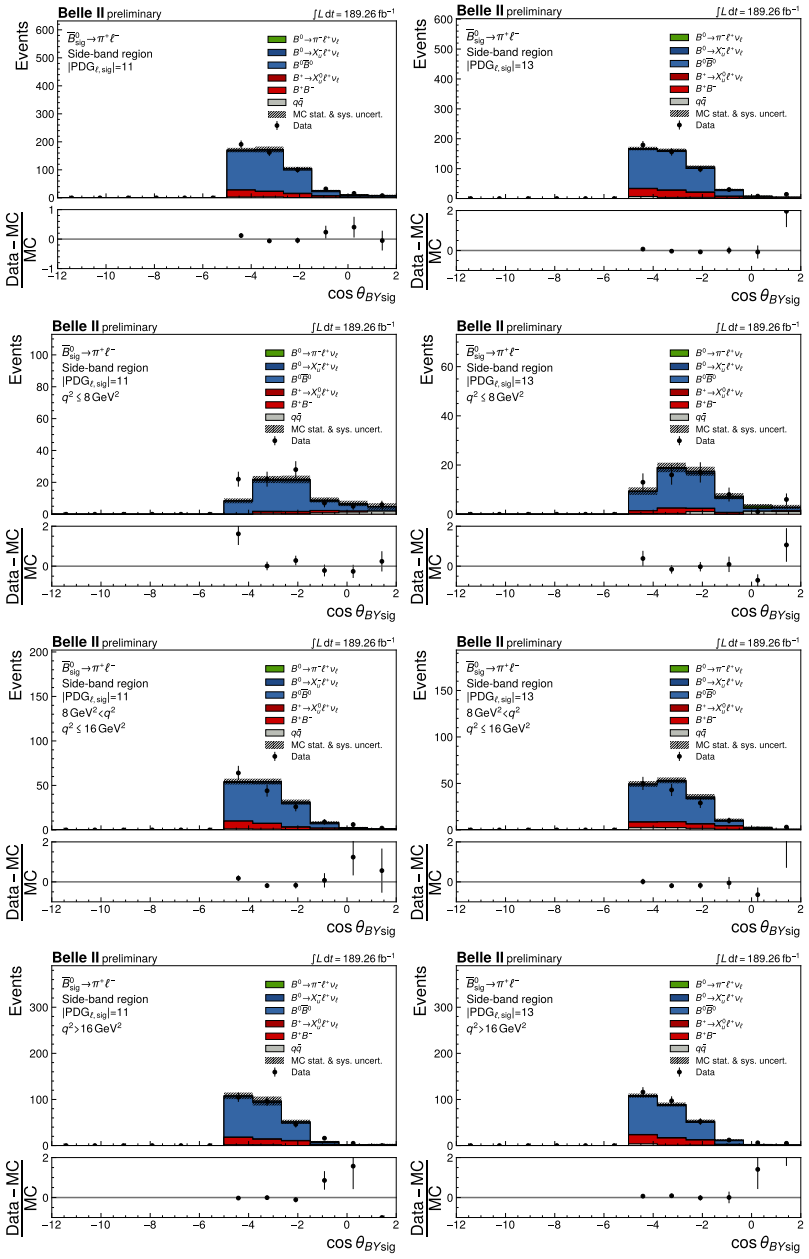


Figure D.7: Signal-side $\cos\theta_{BY}$ for B^0 signal mode, broken down by signal-side lepton flavour (electron in left column, muon in right column).

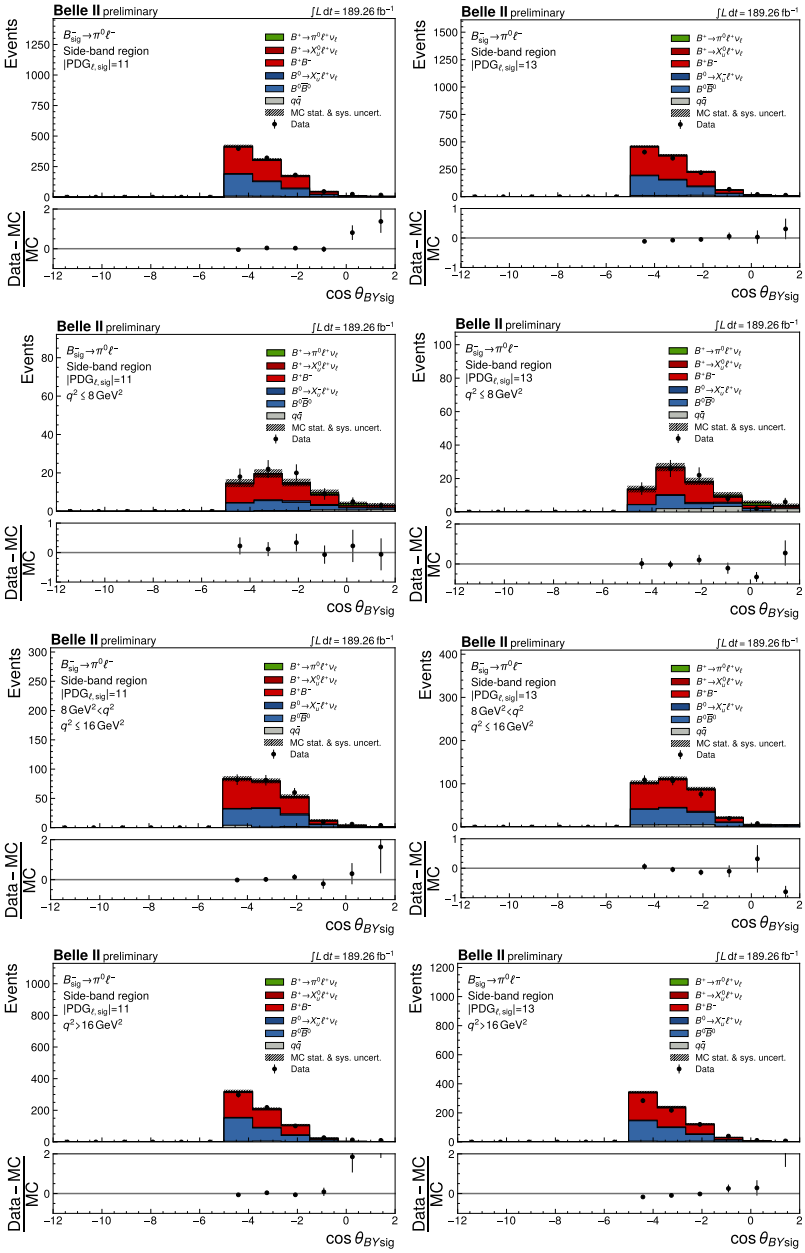


Figure D.8: Signal-side $\cos\theta_{BY}$ for B^+ signal mode, broken down by signal-side lepton flavour (electron in left column, muon in right column).

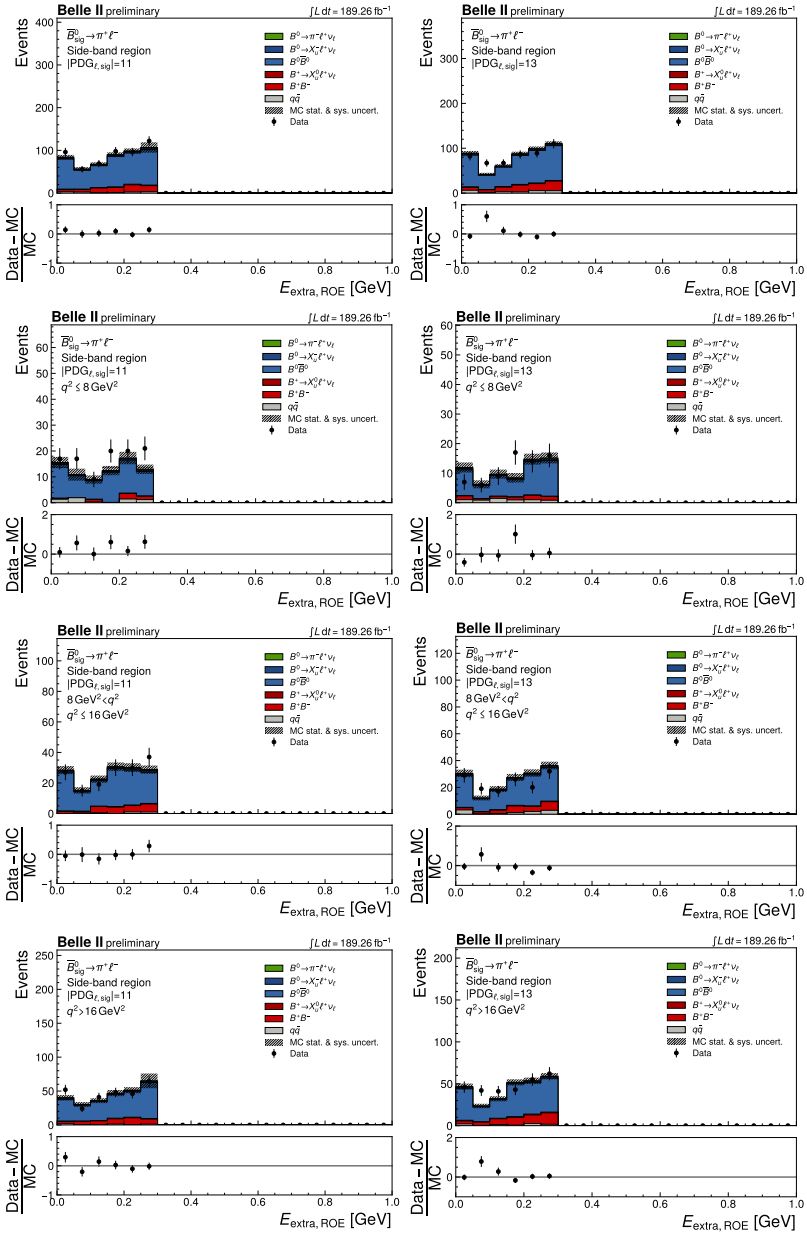


Figure D.9: $E_{\text{extra,ROE}}$ for B^0 signal mode, broken down by signal-side lepton flavour (electron in left column, muon in right column).

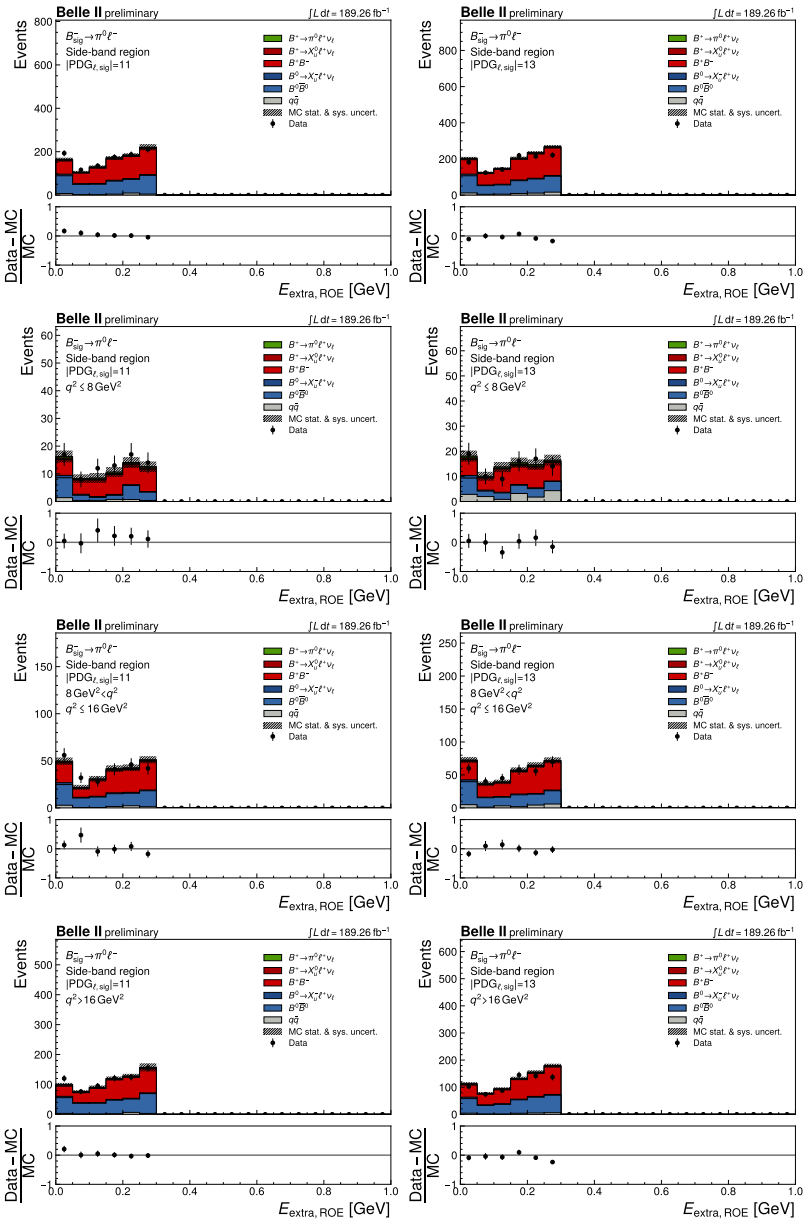


Figure D.10: $E_{\text{extra,ROE}}$ for B^+ signal mode, broken down by signal-side lepton flavour (electron in left column, muon in right column).

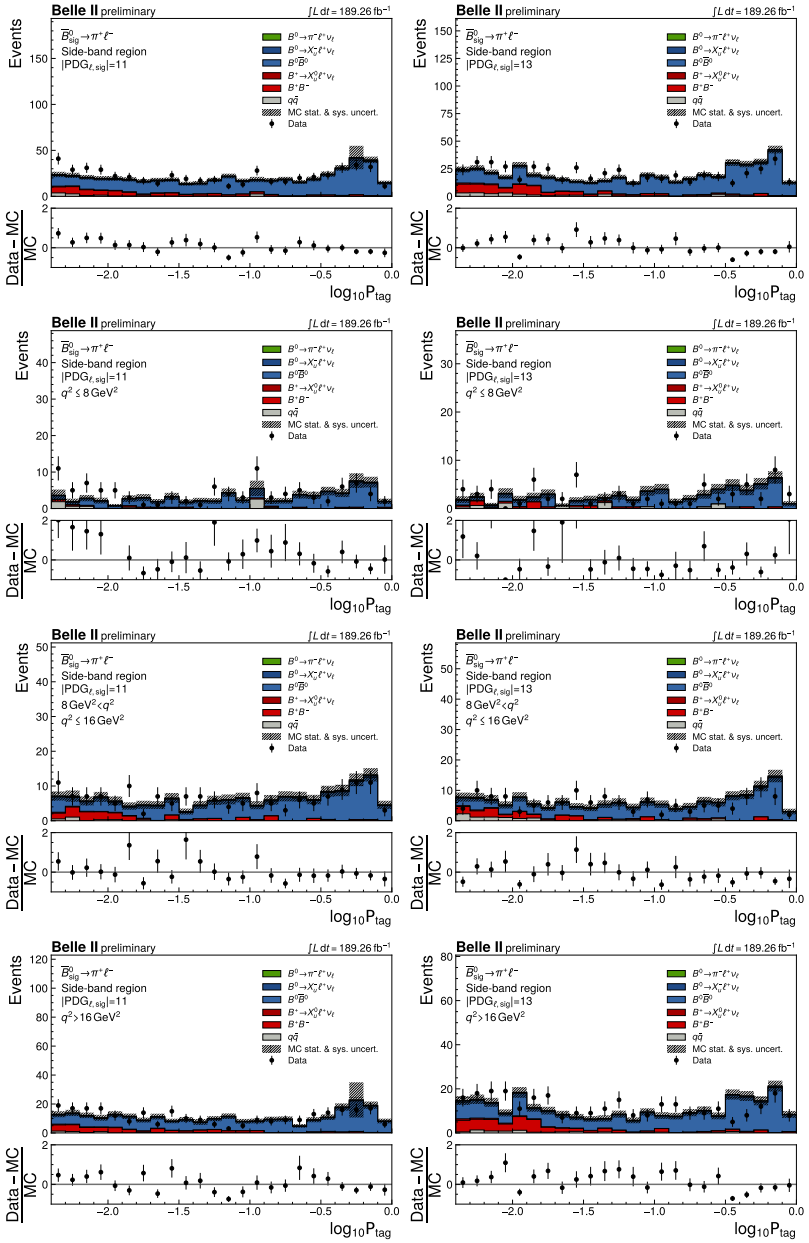


Figure D.11: $\log_{10}\mathscr{P}_{\text{tag}}$ for B^0 signal mode, broken down by signal-side lepton flavour (electron in left column, muon in right column).

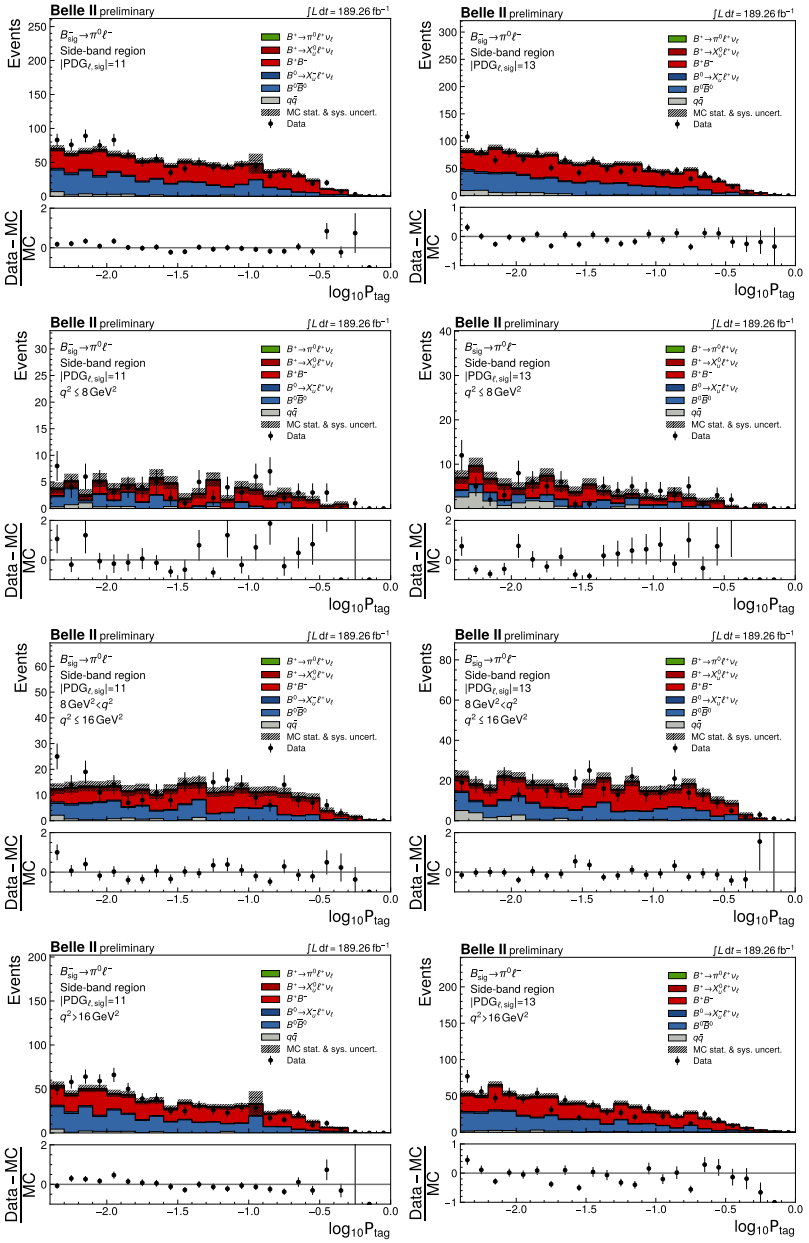


Figure D.12: $\log_{10}\mathcal{P}_{\text{tag}}$ for B^+ signal mode, broken down by signal-side lepton flavour (electron in left column, muon in right column).

D.2. Normalisation mode

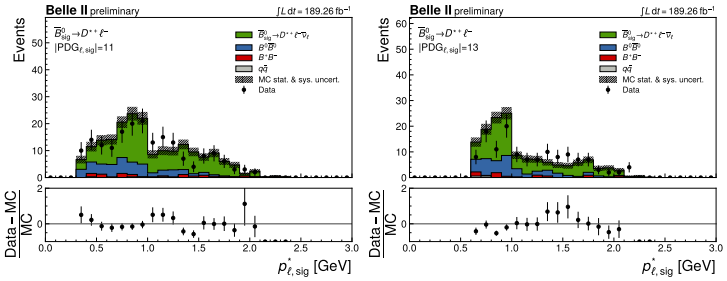


Figure D.13: Signal-side lepton momentum for B^0 normalisation mode, broken down by signal-side lepton flavour (electron on left, muon in right).

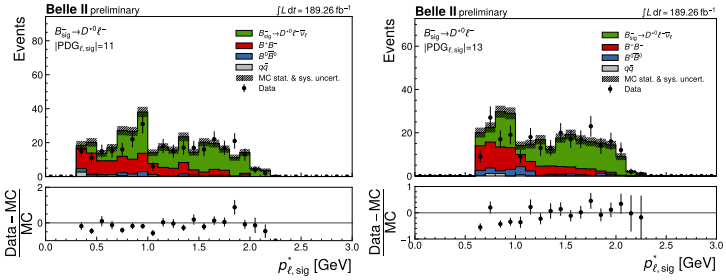


Figure D.14: Signal-side lepton momentum for B^+ normalisation mode, broken down by signal-side lepton flavour (electron on left, muon in right).

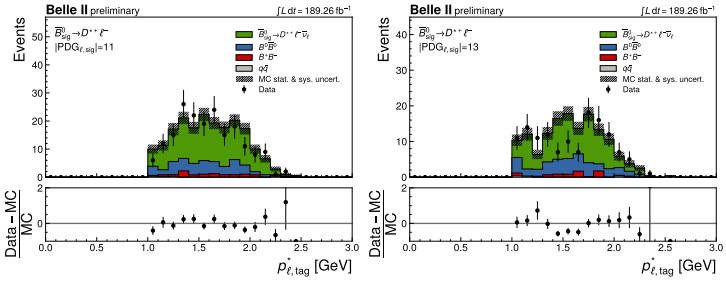


Figure D.15: Tag-side lepton momentum for B^0 normalisation mode, broken down by signal-side lepton flavour (electron on left, muon in right).

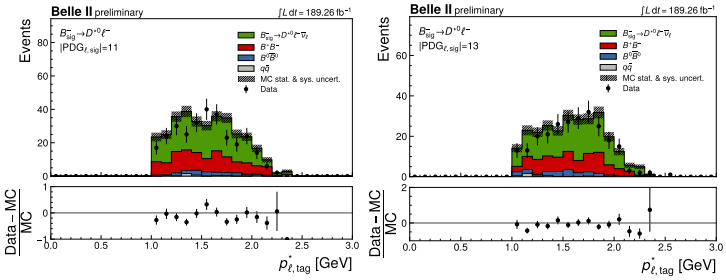


Figure D.16: Tag-side lepton momentum for B^+ normalisation mode, broken down by signal-side lepton flavour (electron on left, muon in right).

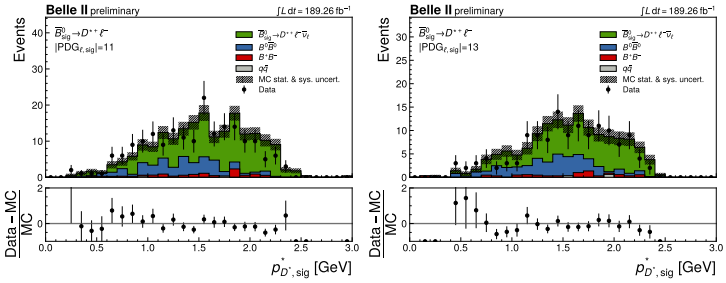


Figure D.17: Signal-side D^* momentum for B^0 normalisation mode, broken down by signal-side lepton flavour (electron on left, muon in right).

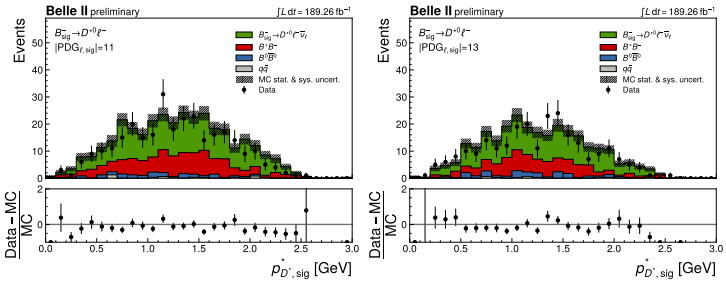


Figure D.18: Signal-side D^* momentum for B^+ normalisation mode, broken down by signal-side lepton flavour (electron on left, muon in right).

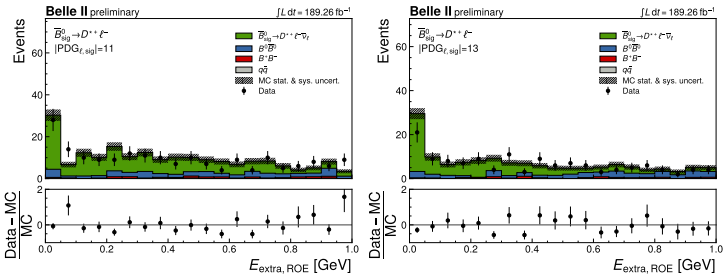


Figure D.19: $E_{\text{extra,ROE}}$ for B^0 normalisation mode, broken down by signal-side lepton flavour (electron on left, muon in right).

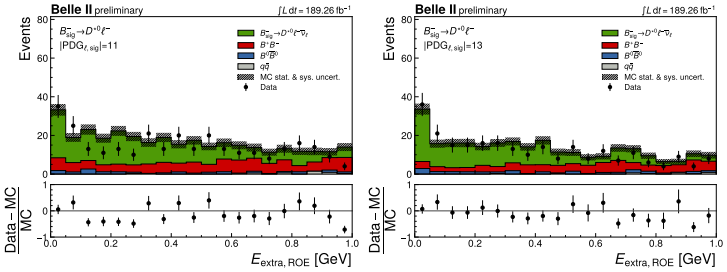


Figure D.20: $E_{\text{extra,ROE}}$ for B^+ normalisation mode, broken down by signal-side lepton flavour (electron on left, muon in right).

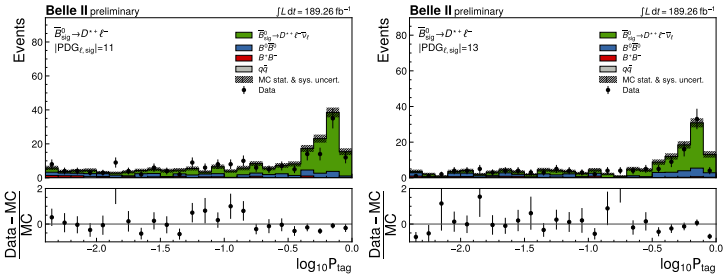


Figure D.21: $\log_{10}\mathscr{P}_{\text{tag}}$ for B^0 normalisation mode, broken down by signal-side lepton flavour (electron on left, muon in right).

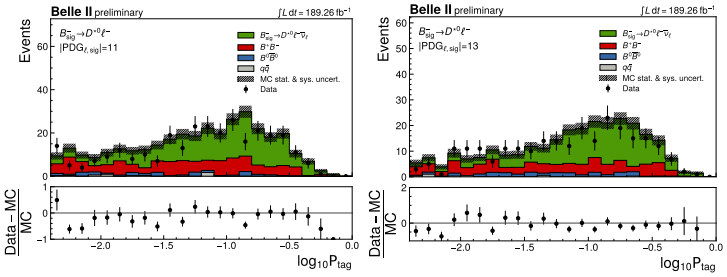


Figure D.22: $\log_{10}\mathscr{P}_{\text{tag}}$ for B^+ normalisation mode, broken down by signal-side lepton flavour (electron on left, muon in right).

Supplementary plots of LQ global fit results

E.

This appendix contains plots of the GAMBIT scan results from Chapter 6. The plots here are of parameter combinations for which there is very little constraint, and were thus not included in the main text.

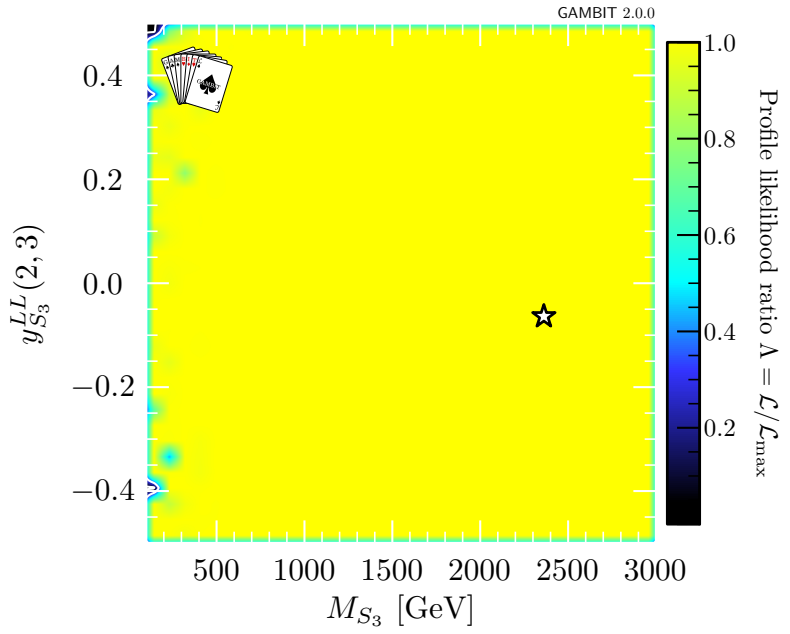


Figure E.1: Profile likelihood plot of M_{S_3} and $(2,3)$ element of $y^{LL}_{S_3}$ in $S_1 + S_3$ scan.

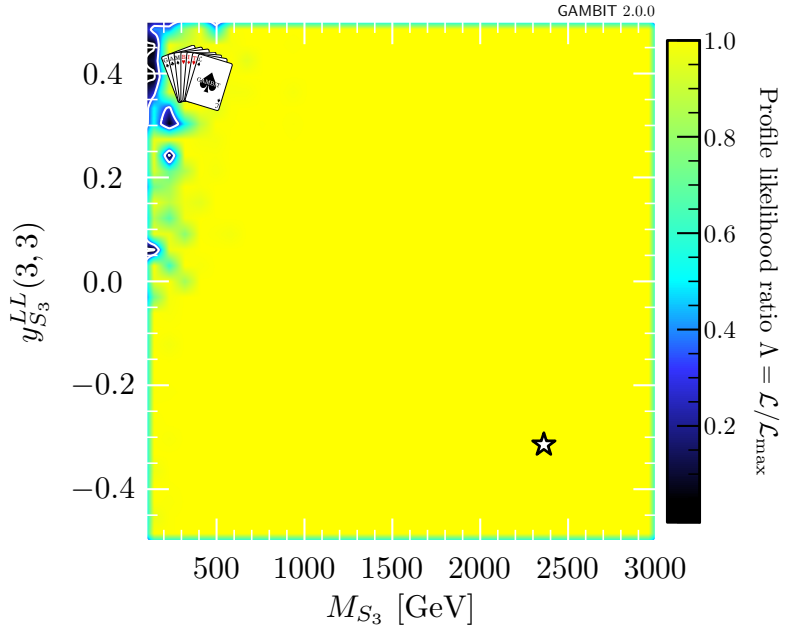


Figure E.2: Profile likelihood plot of M_{S_3} and $(3,3)$ element of $y^{LL}_{S_3}$ in $S_1 + S_3$ scan.

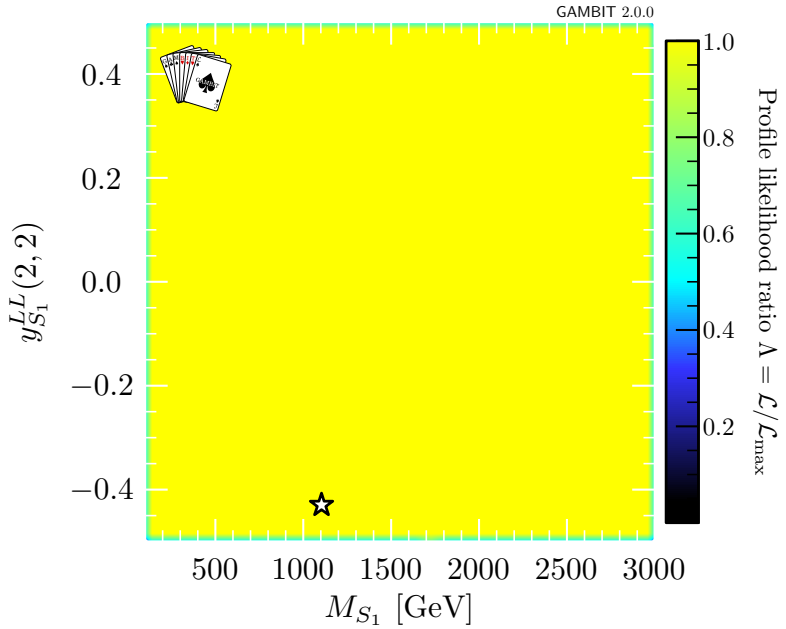


Figure E.3: Profile likelihood plot of M_{S_1} and $(2,2)$ element of $y^{LL}_{S_1}$ in $S_1 + S_3$ scan.

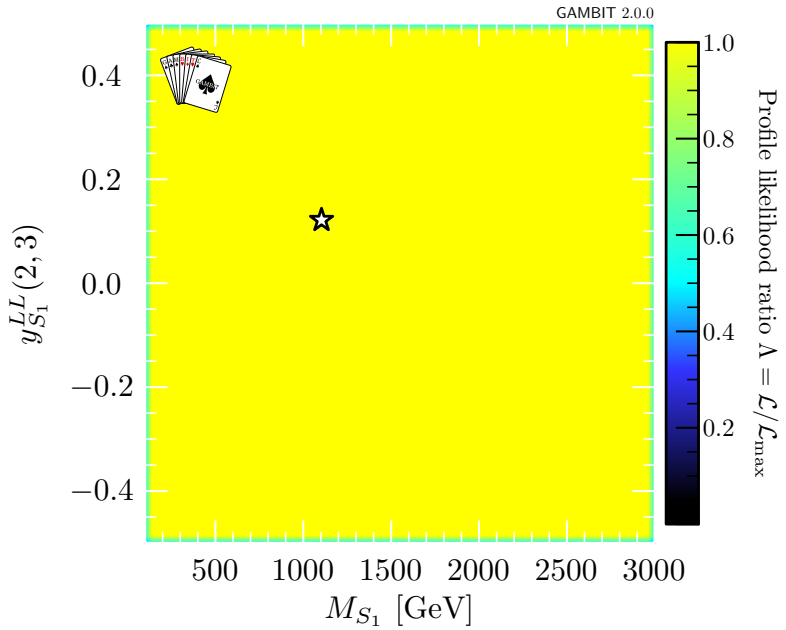


Figure E.4: Profile likelihood plot of M_{S_1} and $(2,3)$ element of $y^{LL}_{S_1}$ in $S_1 + S_3$ scan.

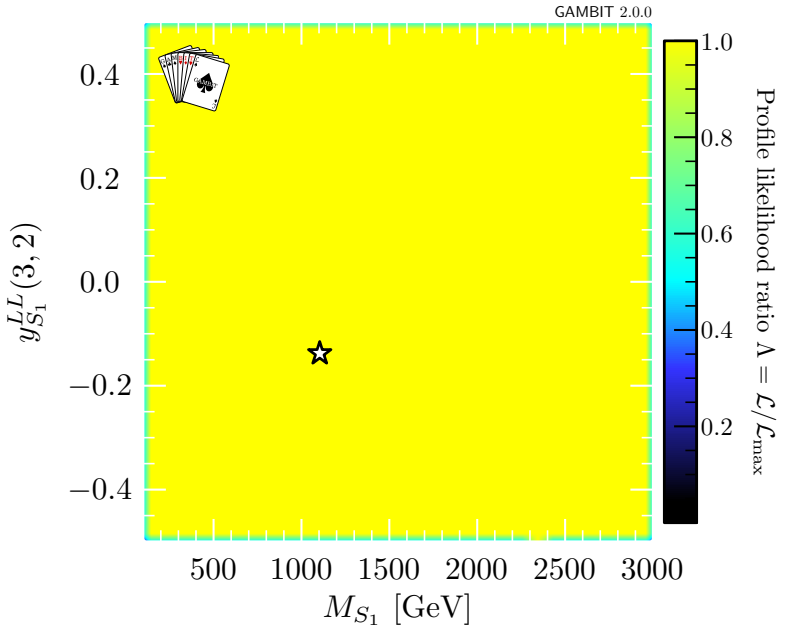


Figure E.5: Profile likelihood plot of M_{S_1} and $(3,2)$ element of $y^{LL}_{S_1}$ in $S_1 + S_3$ scan.

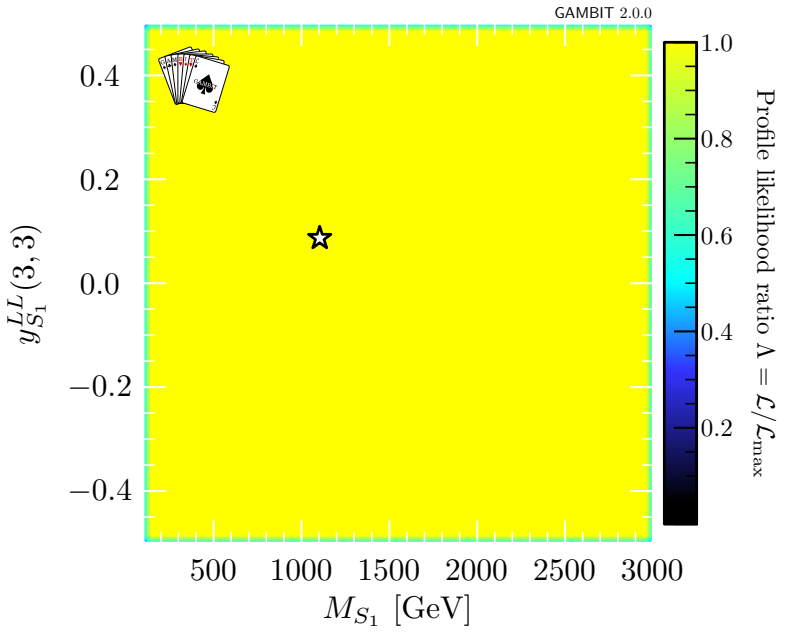


Figure E.6: Profile likelihood plot of M_{S_1} and $(3,3)$ element of $y^{LL}_{S_1}$ in $S_1 + S_3$ scan.

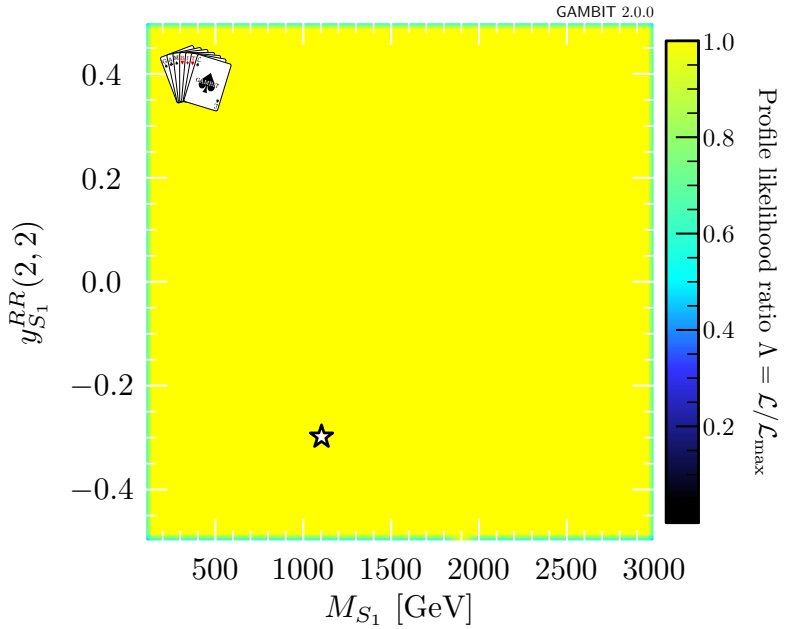


Figure E.7: Profile likelihood plot of M_{S_1} and $(2,2)$ element of $y^{RR}_{S_1}$ in $S_1 + S_3$ scan.

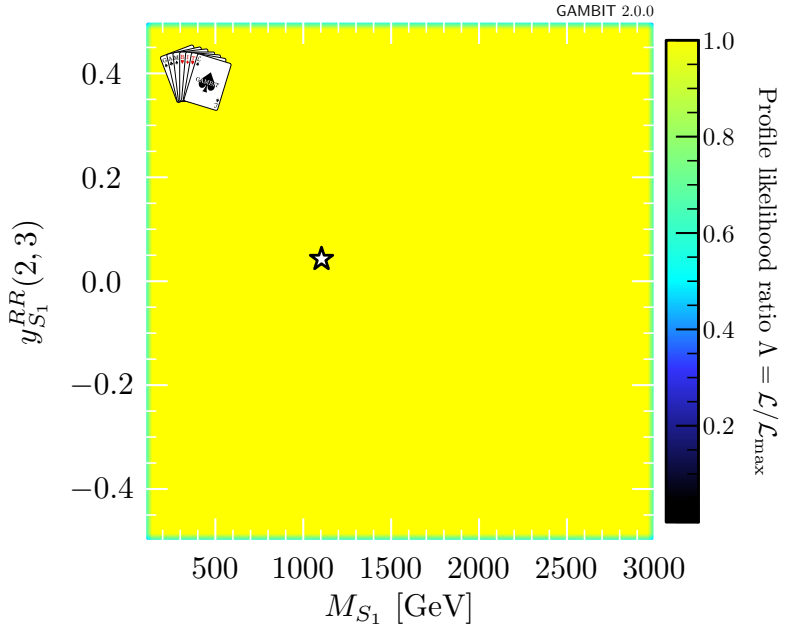


Figure E.8: Profile likelihood plot of M_{S_1} and $(2,3)$ element of $y^{RR}_{S_1}$ in $S_1 + S_3$ scan.

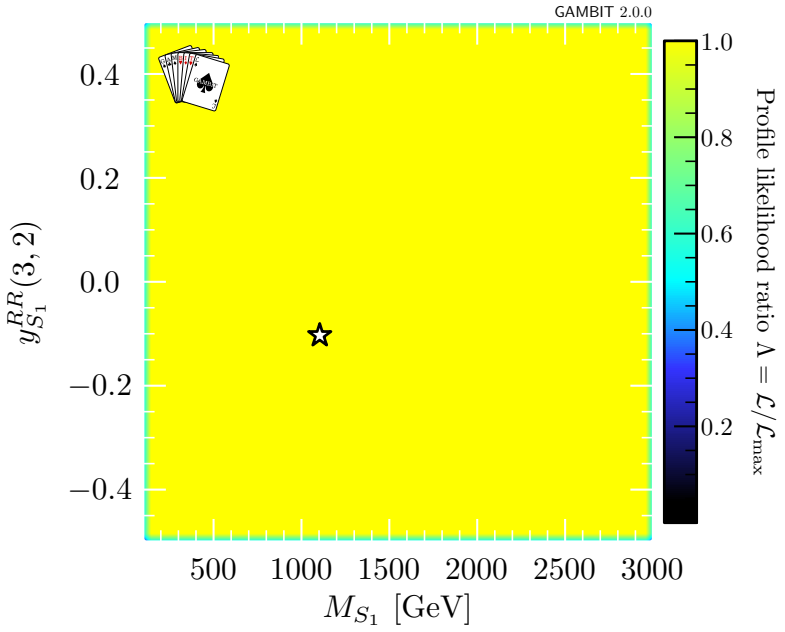


Figure E.9: Profile likelihood plot of M_{S_1} and $(3,2)$ element of $y^{RR}_{S_1}$ in $S_1 + S_3$ scan.

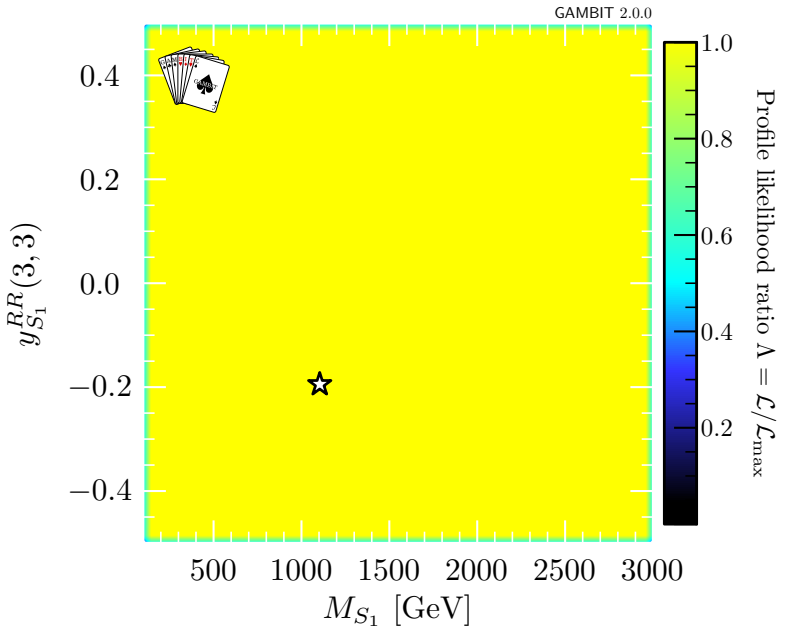


Figure E.10: Profile likelihood plot of M_{S_1} and $(3,3)$ element of $y^{RR}_{S_1}$ in $S_1 + S_3$ scan.

Backmatter

List of abbreviations

ARICH Aerogel Ring Imaging Cherenkov detector
basf2 The Belle II Analysis Software Framework
BSM Physics Beyond the Standard Model
CDB Conditions Database
CDC Central Drift Chamber
CKM Cabibbo-Kobayashi-Maskawa matrix
CMS $\Upsilon(4S)$ Centre-of-mass system
DAQ Data acquisition system
DST Data Summary Table
ECL Electronic Calorimeter
EFT Effective field theory
FCNC Flavour changing neutral current
FEI Full Event Interpretation
FR Full Reconstruction
GAMBIT The Global and Modular Beyond-the-Standard-Model Inference Tool
gbasf2 The Grid basf2 client
GUM The GAMBIT Universal Model Machine
GUT Grand Unified Theory
HEP High Energy Physics
HLT High level trigger
IP Interaction point
KEKCC KEK Central Computer system
KLM K_L^0 And Muon Detector
LCSR Light-cone sum rule
LFU Lepton flavour universality
LHC Large Hadron Collider
LHCb Large Hadron Collider beauty experiment
L1 Level-1 trigger
LPN Logical Path Name
LQ Leptoquark
MC Monte Carlo event generation
mDST Mini Data Summary Table
PID Particle identification
PXD Pixel Detector
QCD Quantum chromodynamics
QFT Quantum field theory
ROE Rest-of-event
SL Semileptonic
SM The Standard Model of Particle Physics
SVD Silicon Vertex Detector
TOP Time-of-Propagation Counter

uDST Micro Data Summary Table
UFO Universal FeynRules Output
VEV Vacuum expectation value
WG Working group

Bibliography

[1] Mark Thomson. *Modern Particle Physics*. New York: Cambridge University Press, 2013.

[2] Michael E. Peskin and Daniel V. Schroeder. *An Introduction to Quantum Field Theory*. Reading, USA: Addison-Wesley, 1995.

[3] Anthony Zee. *Quantum Field Theory in a Nutshell*. Princeton University Press, 2003.

[4] C. N. Yang and R. L. Mills. "Conservation of Isotopic Spin and Isotopic Gauge Invariance". In: *Physical Review* 96.1 (Oct. 1, 1954), pp. 191–195. DOI: 10.1103/PhysRev.96.191. URL: https://link.aps.org/doi/10.1103/PhysRev.96.191 (visited on 08/23/2023).

[5] Abdus Salam. "Weak and Electromagnetic Interactions". In: *Conf. Proc. C* 680519 (1968), pp. 367–377. DOI: 10.1142/9789812795915_0034.

[6] Steven Weinberg. "A Model of Leptons". In: *Physical Review Letters* 19.21 (Nov. 20, 1967), pp. 1264–1266. DOI: 10.1103/PhysRevLett.19.1264. URL: https://link.aps.org/doi/10.1103/PhysRevLett.19.1264 (visited on 08/23/2023).

[7] Sheldon L. Glashow. "Partial-Symmetries of Weak Interactions". In: *Nuclear Physics* 22.4 (Feb. 1, 1961), pp. 579–588. DOI: 10.1016/0029-5582(61)90469-2. URL: https://www.sciencedirect.com/science/article/pii/0029558261904692 (visited on 08/23/2023).

[8] H. David Politzer. "Reliable Perturbative Results for Strong Interactions?" In: *Physical Review Letters* 30.26 (June 25, 1973), pp. 1346–1349. DOI: 10.1103/PhysRevLett.30.1346. URL: https://link.aps.org/doi/10.1103/PhysRevLett.30.1346 (visited on 08/23/2023).

[9] Kenneth G. Wilson. "Confinement of Quarks". In: *Physical Review D* 10.8 (Oct. 15, 1974), pp. 2445–2459. DOI: 10.1103/PhysRevD.10.2445. URL: https://link.aps.org/doi/10.1103/PhysRevD.10.2445 (visited on 08/23/2023).

[10] Kenneth G. Wilson and J. Kogut. "The renormalization group and the ϵ expansion". In: *Physics Reports* 12.2 (Aug. 1, 1974), pp. 75–199. DOI: 10.1016/0370-1573(74)90023-4. URL: https://www.sciencedirect.com/science/article/pii/0370157374900234 (visited on 08/23/2023).

[11] Peter W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons". In: *Physical Review Letters* 13.16 (Oct. 19, 1964), pp. 508–509. DOI: 10.1103/PhysRevLett.13.508. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.508 (visited on 08/23/2023).

[12] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. "Global Conservation Laws and Massless Particles". In: *Physical Review Letters* 13.20 (Nov. 16, 1964), pp. 585–587. DOI: 10.1103/PhysRevLett.13.585. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.585 (visited on 08/23/2023).

[13] F. Englert and R. Brout. "Broken Symmetry and the Mass of Gauge Vector Mesons". In: *Physical Review Letters* 13.9 (Aug. 31, 1964), pp. 321–323. DOI: 10.1103/PhysRevLett.13.321. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.321 (visited on 08/23/2023).

[14] ATLAS Collaboration. "Observation of a New Particle in the Search for the Standard Model Higgs Boson with the ATLAS Detector at the LHC". In: *Physics Letters B* 716.1 (Sept. 2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020. URL: https://linkinghub.elsevier.com/retrieve/pii/S037026931200857X (visited on 12/22/2022).

[15] CMS Collaboration. “Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC”. In: *Physics Letters B* 716.1 (Sept. 2012), pp. 30–61. DOI: 10.1016/j.physletb.2012.08.021. URL: http://arxiv.org/abs/1207.7235 (visited on 12/22/2022).

[16] Nicola Cabibbo. “Unitary Symmetry and Leptonic Decays”. In: *Physical Review Letters* 10.12 (June 15, 1963), pp. 531–533. DOI: 10.1103/PhysRevLett.10.531. URL: https://link.aps.org/doi/10.1103/PhysRevLett.10.531 (visited on 01/30/2023).

[17] Makoto Kobayashi and Toshihide Maskawa. “CP-Violation in the Renormalizable Theory of Weak Interaction”. In: *Progress of Theoretical Physics* 49.2 (Feb. 1, 1973), pp. 652–657. DOI: 10.1143/PTP.49.652. URL: https://doi.org/10.1143/PTP.49.652 (visited on 01/30/2023).

[18] Ling-Lie Chau and Wai-Yee Keung. “Comments on the Parametrization of the Kobayashi-Maskawa Matrix”. In: *Physical Review Letters* 53.19 (Nov. 5, 1984), pp. 1802–1805. DOI: 10.1103/PhysRevLett.53.1802. URL: https://link.aps.org/doi/10.1103/PhysRevLett.53.1802 (visited on 08/22/2023).

[19] Lincoln Wolfenstein. “Parametrization of the Kobayashi-Maskawa Matrix”. In: *Physical Review Letters* 51.21 (Nov. 21, 1983), pp. 1945–1947. DOI: 10.1103/PhysRevLett.51.1945. URL: https://link.aps.org/doi/10.1103/PhysRevLett.51.1945 (visited on 08/22/2023).

[20] Particle Data Group. “Review of Particle Physics”. In: *Progress of Theoretical and Experimental Physics* 2022.8 (Aug. 8, 2022), p. 083C01. DOI: 10.1093/ptep/ptac097. URL: https://academic.oup.com/ptep/article/doi/10.1093/ptep/ptac097/6651666 (visited on 01/16/2023).

[21] CKMfitter Collaboration. *Preliminary Results as of Spring 2021*. Jan. 2021. URL: http://ckmfitter.in2p3.fr/www/results/plots_spring21/ckm_res_spring21.html (visited on 01/30/2023).

[22] UTfit Collaboration et al. *New UTfit Analysis of the Unitarity Triangle in the Cabibbo-Kobayashi-Maskawa Scheme*. Dec. 7, 2022. DOI: 10.48550/arXiv.2212.03894. URL: http://arxiv.org/abs/2212.03894 (visited on 01/30/2023). preprint.

[23] E. Kou et al. “The Belle II Physics Book”. In: (Aug. 31, 2018). URL: https://arxiv.org/abs/1808.10567v3 (visited on 02/05/2019).

[24] A. J. Bevan et al. “The Physics of the B Factories”. In: *The European Physical Journal C* 74.11 (Nov. 2014). DOI: 10.1140/epjc/s10052-014-3026-9. URL: http://link.springer.com/10.1140/epjc/s10052-014-3026-9 (visited on 12/04/2018).

[25] Belle Collaboration. “Observation of Large CP Violation and Evidence for Direct CP Violation in $B^0 \to \pi^+\pi^-$ Decays”. In: *Physical Review Letters* 93.2 (July 6, 2004), p. 021601. DOI: 10.1103/PhysRevLett.93.021601. URL: https://link.aps.org/doi/10.1103/PhysRevLett.93.021601 (visited on 12/22/2022).

[26] Belle Collaboration. “Evidence for Direct CP Violation in $B^0 \to K^+\pi^-$ Decays”. In: *Physical Review Letters* 93.19 (Nov. 5, 2004), p. 191802. DOI: 10.1103/PhysRevLett.93.191802. URL: https://link.aps.org/doi/10.1103/PhysRevLett.93.191802 (visited on 12/22/2022).

[27] BaBar Collaboration. “Direct CP Violating Asymmetry in $B^0 \to K^+\pi^-$ Decays”. In: *Physical Review Letters* 93.13 (Sept. 20, 2004), p. 131801. DOI: 10.1103/PhysRevLett.93.131801. URL: http://arxiv.org/abs/hep-ex/0407057 (visited on 08/22/2023).

[28] S. Aoki et al. “FLAG Review 2019”. Version 3. In: (2019). DOI: 10.48550/ARXIV.1902.08191. URL: https://arxiv.org/abs/1902.08191 (visited on 01/11/2023).

[29] BaBar Collaboration. "Study of $B \to \pi \ell \nu$ and $B \to \rho \ell \nu$ Decays and Determination of $|V_{ub}|$". In: *Physical Review D* 83.3 (Feb. 7, 2011), p. 032007. DOI: 10.1103/PhysRevD.83.032007. URL: http://arxiv.org/abs/1005.3288 (visited on 07/15/2019).

[30] Fermilab Lattice Collaboration et al. "$|V_{ub}|$ from $B \to \pi \ell \nu$ decays and (2+1)-flavor lattice QCD". In: *Physical Review D* 92.1 (July 23, 2015), p. 014024. DOI: 10.1103/PhysRevD.92.014024. URL: http://arxiv.org/abs/1503.07839 (visited on 10/04/2022).

[31] J. M. Flynn et al. "$B \to \pi \ell \nu$ and $B_s \to K \ell \nu$ form factors and $|V_{ub}|$ from 2+1-flavor lattice QCD with domain-wall light quarks and relativistic heavy quarks". In: *Physical Review D* 91.7 (Apr. 14, 2015), p. 074510. DOI: 10.1103/PhysRevD.91.074510. URL: http://arxiv.org/abs/1501.05373 (visited on 08/22/2023).

[32] Alexander Khodjamirian, Thomas Mannel, and Nils Offen. "B-Meson Distribution Amplitude from the $B \to \pi$ Form Factor". In: *Physics Letters B* 620.1-2 (July 2005), pp. 52–60. DOI: 10.1016/j.physletb.2005.06.021. URL: http://arxiv.org/abs/hep-ph/0504091 (visited on 08/22/2023).

[33] Nico Gubernari, Ahmet Kokulu, and Danny van Dyk. "$B \to P$ and $B \to V$ Form Factors from B-Meson Light-Cone Sum Rules beyond Leading Twist". In: *Journal of High Energy Physics* 2019.1 (Jan. 2019), p. 150. DOI: 10.1007/JHEP01(2019)150. URL: http://arxiv.org/abs/1811.00983 (visited on 08/22/2023).

[34] UTfit Collaboration et al. "An Improved Standard Model Prediction Of BR($B \to \tau \nu$) And Its Implications For New Physics". In: *Physics Letters B* 687.1 (Apr. 2010), pp. 61–69. DOI: 10.1016/j.physletb.2010.02.063. URL: http://arxiv.org/abs/0908.3470 (visited on 12/20/2022).

[35] David London and Joaquim Matias. "B Flavour Anomalies: 2021 Theoretical Status Report". In: *Annual Review of Nuclear and Particle Science* 72.1 (Sept. 26, 2022), pp. 37–68. DOI: 10.1146/annurev-nucl-102020-090209. URL: http://arxiv.org/abs/2110.13270 (visited on 12/20/2022).

[36] Patrick Koppenburg. *Anomalies Plots*. Apr. 25, 2023. URL: https://www.nikhef.nl/~pkoppenb/anomalies.html (visited on 12/15/2022).

[37] LHCb Collaboration. "Measurement of lepton universality parameters in $B^+ \to K^+ \ell^+ \ell^-$ and $B^0 \to K^{*0} \ell^+ \ell^-$ decays". Version 1. In: (2022). DOI: 10.48550/ARXIV.2212.09153. URL: https://arxiv.org/abs/2212.09153 (visited on 12/22/2022).

[38] HFLAV Collaboration. *Averages of b-hadron, c-hadron, and τ-lepton properties as of 2021*. Jan. 23, 2023. URL: http://arxiv.org/abs/2206.07501 (visited on 01/30/2023). preprint.

[39] Belle Collaboration. "Measurement of the τ lepton polarization and $R(D^*)$ in the decay $\bar{B} \to D^* \tau^- \bar{\nu}_\tau$". In: *Physical Review Letters* 118.21 (May 26, 2017). DOI: 10.1103/PhysRevLett.118.211801. URL: http://arxiv.org/abs/1612.00529 (visited on 03/22/2019).

[40] Antonio Petrella. "Inclusive and Exclusive $|V_{ub}|$". Mar. 30, 2009. URL: http://arxiv.org/abs/0903.5180 (visited on 03/19/2019).

[41] Christopher Bouchard, Lu Cao, and Patrick Owen. *Summary of the 2018 CKM working group on semileptonic and leptonic b-hadron decays*. Feb. 25, 2019. DOI: 10.48550/arXiv.1902.09412. URL: http://arxiv.org/abs/1902.09412 (visited on 12/18/2022). preprint.

[42] E. McLean et al. *$B_s \to D_s^{(*)} l \nu$ Form Factors with Heavy HISQ Quarks*. Jan. 15, 2019. DOI: 10.48550/arXiv.1901.04979. URL: http://arxiv.org/abs/1901.04979 (visited on 12/22/2022). preprint.

[43] Muon g−2 Collaboration. "Final Report of the Muon E821 Anomalous Magnetic Moment Measurement at BNL". In: *Physical Review D* 73.7 (Apr. 7, 2006), p. 072003. DOI: 10.1103/PhysRevD.73.072003. URL: http://arxiv.org/abs/hep-ex/0602035 (visited on 12/19/2022).

[44] B. Abi et al. “Measurement of the Positive Muon Anomalous Magnetic Moment to 0.46 Ppm”. In: *Physical Review Letters* 126.14 (Apr. 7, 2021), p. 141801. DOI: 10.1103/PhysRevLett.126.141801. URL: http://arxiv.org/abs/2104.03281 (visited on 12/19/2022).

[45] Andreas Crivellin and Martin Hoferichter. *Combined explanations of $(g-2)_\mu$, $(g-2)_e$ and implications for a large muon EDM*. May 9, 2019. DOI: 10.48550/arXiv.1905.03789. URL: http://arxiv.org/abs/1905.03789 (visited on 01/30/2023). preprint.

[46] Peter Athron et al. “New physics explanations of a_μ in light of the FNAL muon $g-2$ measurement”. In: *Journal of High Energy Physics* 2021.9 (Sept. 2021), p. 80. DOI: 10.1007/JHEP09(2021)080. URL: http://arxiv.org/abs/2104.03691 (visited on 01/30/2023).

[47] Sz Borsanyi et al. “Leading Hadronic Contribution to the Muon Magnetic Moment from Lattice QCD”. In: *Nature* 593.7857 (7857 May 2021), pp. 51–55. DOI: 10.1038/s41586-021-03418-1. URL: https://www.nature.com/articles/s41586-021-03418-1 (visited on 04/25/2023).

[48] Howard Georgi and S. L. Glashow. “Unity of All Elementary-Particle Forces”. In: *Physical Review Letters* 32.8 (Feb. 25, 1974), pp. 438–441. DOI: 10.1103/PhysRevLett.32.438. URL: https://link.aps.org/doi/10.1103/PhysRevLett.32.438 (visited on 02/25/2020).

[49] Edward Witten. “Quest For Unification”. July 9, 2002. URL: http://arxiv.org/abs/hep-ph/0207124 (visited on 02/25/2020).

[50] JoAnne L. Hewett and Thomas G. Rizzo. “Much Ado About Leptoquarks: A Comprehensive Analysis”. In: *Physical Review D* 56.9 (Nov. 1, 1997), pp. 5709–5724. DOI: 10.1103/PhysRevD.56.5709. URL: http://arxiv.org/abs/hep-ph/9703337 (visited on 08/27/2019).

[51] I. Doršner et al. “Physics of leptoquarks in precision experiments and at particle colliders”. In: *Physics Reports* 641 (June 2016), pp. 1–68. DOI: 10.1016/j.physrep.2016.06.001. URL: http://arxiv.org/abs/1603.04993 (visited on 08/27/2019).

[52] Super-Kamiokande Collaboration. “Search for proton decay via $p \to e^+\pi^0$ and $p \to \mu^+\pi^0$ with an enlarged fiducial volume in Super-Kamiokande I-IV”. In: *Physical Review D* 102.11 (Dec. 22, 2020), p. 112011. DOI: 10.1103/PhysRevD.102.112011. URL: http://arxiv.org/abs/2010.16098 (visited on 02/10/2023).

[53] Andrei Angelescu et al. “Closing the window on single leptoquark solutions to the B-physics anomalies”. In: *Journal of High Energy Physics* 2018.10 (Oct. 2018), p. 183. DOI: 10.1007/JHEP10(2018)183. URL: http://arxiv.org/abs/1808.08179 (visited on 07/07/2022).

[54] Martin Bauer and Matthias Neubert. “One Leptoquark to Rule Them All: A Minimal Explanation for $R_{D^{(*)}}$, R_K and $(g-2)_\mu$”. In: (Nov. 5, 2015). DOI: 10.1103/PhysRevLett.116.141802. URL: https://arxiv.org/abs/1511.01900v2 (visited on 02/10/2023).

[55] Bernat Capdevila et al. “Patterns of New Physics in $b \to s\ell^+\ell^-$ transitions in the light of recent data”. In: *Journal of High Energy Physics* 2018.1 (Jan. 2018). DOI: 10.1007/JHEP01(2018)093. URL: http://arxiv.org/abs/1704.05340 (visited on 01/17/2019).

[56] Yasuhito Sakaki et al. “Testing leptoquark models in $\bar{B} \to D^{(*)}\tau\bar{\nu}$”. In: (Sept. 2, 2013). DOI: 10.1103/PhysRevD.88.094012. URL: https://arxiv.org/abs/1309.0301v3 (visited on 02/10/2023).

[57] Guido D’Amico et al. “Flavour anomalies after the R_{K^*} measurement”. In: *Journal of High Energy Physics* 2017.9 (Sept. 2017). DOI: 10.1007/JHEP09(2017)010. URL: http://arxiv.org/abs/1704.05438 (visited on 01/17/2019).

[58] ATLAS Collaboration. *Search for pair production of third-generation scalar leptoquarks decaying into a top quark and a τ-lepton in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*. July 20, 2021. DOI: 10.1007/JHEP06(2021)179. URL: http://arxiv.org/abs/2101.11582 (visited on 02/10/2023). preprint.

[59] ATLAS Collaboration. "Search for new phenomena in pp collisions in final states with tau leptons, b-jets, and missing transverse momentum with the ATLAS detector". In: *Physical Review D* 104.11 (Dec. 21, 2021), p. 112005. DOI: 10.1103/PhysRevD.104.112005. URL: http://arxiv.org/abs/2108.07665 (visited on 02/10/2023).

[60] ATLAS Collaboration. *Search for new phenomena in final states with b-jets and missing transverse momentum in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector*. May 28, 2021. DOI: 10.1007/JHEP05(2021)093. URL: http://arxiv.org/abs/2101.12527 (visited on 02/10/2023). preprint.

[61] CMS Collaboration. *Search for singly and pair-produced leptoquarks coupling to third-generation fermions in proton-proton collisions at $\sqrt{s} = 13$ TeV*. June 19, 2021. DOI: 10.1016/j.physletb.2021.136446. URL: http://arxiv.org/abs/2012.04178 (visited on 02/10/2023). preprint.

[62] Matthias Neubert. "Effective Field Theory and Heavy Quark Physics". In: *Physics In D ≥ 4 Tasi 2004* (July 2006), pp. 149–194. DOI: 10.1142/9789812773579_0004. URL: http://arxiv.org/abs/hep-ph/0512222 (visited on 01/29/2021).

[63] Thomas Keck. *The Full Event Interpretation for Belle II*. 2015. URL: https://docs.belle2.org/record/275.

[64] Thomas Keck et al. "The Full Event Interpretation – An exclusive tagging algorithm for the Belle II experiment". July 23, 2018. URL: http://arxiv.org/abs/1807.08680 (visited on 01/17/2019).

[65] T. Abe et al. *Belle II Technical Design Report*. Nov. 1, 2010. DOI: 10.48550/arXiv.1011.0352. URL: http://arxiv.org/abs/1011.0352 (visited on 11/09/2022). preprint.

[66] S. Kurokawa and E. Kikutani. "Overview of the KEKB Accelerators". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*. KEKB: The KEK B-factory 499.1 (Feb. 21, 2003), pp. 1–7. DOI: 10.1016/S0168-9002(02)01771-0. URL: https://www.sciencedirect.com/science/article/pii/S0168900202017710 (visited on 11/09/2022).

[67] Kazunori Akai, Kazuro Furukawa, and Haruyo Koiso. "SuperKEKB Collider". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*. Advances in Instrumentation and Experimental Methods (Special Issue in Honour of Kai Siegbahn) 907 (Nov. 1, 2018), pp. 188–199. DOI: 10.1016/j.nima.2018.08.017. URL: https://www.sciencedirect.com/science/article/pii/S0168900218309616 (visited on 02/02/2023).

[68] Mikhail Zobov. "Crab Waist collision scheme: a novel approach for particle colliders". In: *Journal of Physics: Conference Series* 747 (Sept. 2016), p. 012090. DOI: 10.1088/1742-6596/747/1/012090. URL: http://arxiv.org/abs/1608.06150 (visited on 02/22/2019).

[69] Y. Ohnishi et al. "SuperKEKB Operation Using Crab Waist Collision Scheme". In: *The European Physical Journal Plus* 136.10 (Oct. 12, 2021), p. 1023. DOI: 10.1140/epjp/s13360-021-01979-8. URL: https://doi.org/10.1140/epjp/s13360-021-01979-8 (visited on 11/10/2022).

[70] KEK. "SuperKEKB Collider Achieves the World's Highest Luminosity". In: *Brookhaven National Laboratory* (June 26, 2022). URL: https://www.bnl.gov/newsroom/news.php?a=117285 (visited on 11/10/2022).

[71] Yoshiyuki Onuki. "Belle II Status and Prospect". In: *Nuclear and Particle Physics Proceedings*. QCD 21 Is the 24th International Conference on Quantum Chromodynamics 318–323 (Nov. 1, 2022), pp. 78–84. DOI: 10.1016/j.nuclphysbps.2022.09.017. URL: https://www.sciencedirect.com/science/article/pii/S2405601422000177 (visited on 02/02/2023).

[72] H. Ye et al. "Commissioning and Performance of the Belle II Pixel Detector". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 987 (Jan. 21, 2021), p. 164875. DOI: 10.1016/j.nima.2020.164875. URL: https://www.sciencedirect.com/science/article/pii/S0168900220312729 (visited on 02/02/2023).

[73] Francesco Forti. "The Design, Construction, Operation and Performance of the Belle II Silicon Vertex Detector". In: *Journal of Instrumentation* 17.11 (Nov. 1, 2022), P11042. DOI: 10.1088/1748-0221/17/11/P11042. URL: http://arxiv.org/abs/2201.09824 (visited on 02/02/2023).

[74] N. Taniguchi. "Central Drift Chamber for Belle-II". In: *Journal of Instrumentation* 12.06 (June 2017), p. C06014. DOI: 10.1088/1748-0221/12/06/C06014. URL: https://dx.doi.org/10.1088/1748-0221/12/06/C06014 (visited on 02/02/2023).

[75] J. Fast. "The Belle II Imaging Time-of-Propagation (iTOP) Detector". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*. The 9th International Workshop on Ring Imaging Cherenkov Detectors (RICH2016) 876 (Dec. 21, 2017), pp. 145–148. DOI: 10.1016/j.nima.2017.02.045. URL: https://www.sciencedirect.com/science/article/pii/S0168900217302280 (visited on 02/02/2023).

[76] S. Nishida et al. "Aerogel RICH for the Belle II Forward PID". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*. RICH2013 Proceedings of the Eighth International Workshop on Ring Imaging Cherenkov Detectors Shonan, Kanagawa, Japan, December 2-6, 2013 766 (Dec. 1, 2014), pp. 28–31. DOI: 10.1016/j.nima.2014.06.061. URL: https://www.sciencedirect.com/science/article/pii/S0168900214008080 (visited on 02/02/2023).

[77] M. Starič. "The TOP Counter and Determination of Bunch-Crossing Time at Belle II". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 1055 (Oct. 1, 2023), p. 168474. DOI: 10.1016/j.nima.2023.168474. URL: https://www.sciencedirect.com/science/article/pii/S0168900223004643 (visited on 08/26/2023).

[78] V. Aulchenko et al. "Electromagnetic Calorimeter for Belle II". In: *Journal of Physics: Conference Series* 587.1 (Feb. 2015), p. 012045. DOI: 10.1088/1742-6596/587/1/012045. URL: https://dx.doi.org/10.1088/1742-6596/587/1/012045 (visited on 02/02/2023).

[79] T. Aushev et al. "A scintillator based endcap K_L^0 and muon detector for the Belle II experiment". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 789 (July 21, 2015), pp. 134–142. DOI: 10.1016/j.nima.2015.03.060. URL: https://www.sciencedirect.com/science/article/pii/S016890021500385X (visited on 02/02/2023).

[80] Satoru Yamada et al. "Data Acquisition System for the Belle II Experiment". In: *IEEE Transactions on Nuclear Science* 62.3 (June 2015), pp. 1175–1180. DOI: 10.1109/TNS.2015.2424717.

[81] Tom Browder. "Belle II Overview". US Belle II Operations Review. Oct. 29, 2020. URL: https://indico.bnl.gov/event/9843/#3-belle-ii-overview (visited on 08/26/2023).

[82] Leo Piilonen. "Particle Identification in the Belle II Detector". IAS HEP2021 Workshop. Jan. 14, 2021. URL: https://docs.belle2.org/record/2278 (visited on 08/26/2023).

[83] Torben Ferber and Phillip Urquijo. *Overview of the Belle II Physics Generators*. June 22, 2015. URL: https://docs.belle2.org/record/282. preprint.

[84] Yo Sato et al. "Monte Carlo Matching in the Belle II Software". In: *EPJ Web of Conferences* 251 (2021), p. 03021. DOI: 10.1051/epjconf/202125103021. URL: https://www.epj-conferences.org/articles/epjconf/abs/2021/05/epjconf_chep2021_03021/epjconf_chep2021_03021.html (visited on 11/24/2022).

[85] Valerio Bertacchi et al. "Track Finding at Belle II". In: *Computer Physics Communications* 259 (Feb. 2021), p. 107610. DOI: 10.1016/j.cpc.2020.107610. URL: http://arxiv.org/abs/2003.12466 (visited on 08/26/2023).

[86] Belle II Collaboration. *Belle II Analysis Software Framework (Basf2)*. Oct. 22, 2022. DOI: 10.5281/zenodo.5574115. URL: https://github.com/belle2/basf2 (visited on 11/08/2022).

[87] T. Kuhr et al. "The Belle II Core Software". In: *Computing and Software for Big Science* 3.1 (Dec. 2019), p. 1. DOI: 10.1007/s41781-018-0017-9. URL: http://arxiv.org/abs/1809.04299 (visited on 02/02/2023).

[88] Sam Cunliffe. "High-Level Analysis Software for the Belle II Experiment". HSF Analysis WG. Feb. 13, 2019. URL: https://indico.cern.ch/event/789007/#6-basf2-the-belle-ii-analysis (visited on 09/07/2023).

[89] Hideki Miyake et al. "Belle II Production System". In: *Journal of Physics: Conference Series* 664 (Dec. 23, 2015), p. 052028. DOI: 10.1088/1742-6596/664/5/052028.

[90] Geoffrey C. Fox and Stephen Wolfram. "Observables for the Analysis of Event Shapes in e^+e^- Annihilation and Other Processes". In: *Physical Review Letters* 41.23 (Dec. 4, 1978), pp. 1581–1585. DOI: 10.1103/PhysRevLett.41.1581. URL: https://link.aps.org/doi/10.1103/PhysRevLett.41.1581 (visited on 08/01/2019).

[91] BaBar Collaboration. "Measurement of the $B \to \pi\ell\nu$ Branching Fraction and Determination of $|V_{ub}|$ with Tagged B Mesons". In: *Physical Review Letters* 97.21 (Nov. 21, 2006), p. 211801. DOI: 10.1103/PhysRevLett.97.211801. URL: http://arxiv.org/abs/hep-ex/0607089 (visited on 05/28/2020).

[92] Belle Collaboration. "Measurements of branching fractions and q^2 distributions for $B \to \pi\ell\nu$ and $B \to \rho\ell\nu$ Decays with $B \to D^{(*)}\ell\nu$ Decay Tagging". In: *Physics Letters B* 648.2-3 (May 2007), pp. 139–148. DOI: 10.1016/j.physletb.2007.02.067. URL: http://arxiv.org/abs/hep-ex/0604024 (visited on 07/15/2019).

[93] M. Feindt and U. Kerzel. "The NeuroBayes Neural Network Package". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*. Proceedings of the X International Workshop on Advanced Computing and Analysis Techniques in Physics Research 559.1 (Apr. 1, 2006), pp. 190–194. DOI: 10.1016/j.nima.2005.11.166. URL: https://www.sciencedirect.com/science/article/pii/S0168900205022679 (visited on 11/24/2022).

[94] Thomas Keck. *FastBDT: A Speed-Optimized and Cache-Friendly Implementation of Stochastic Gradient-Boosted Decision Trees for Multivariate Classification*. Sept. 20, 2016. DOI: 10.48550/arXiv.1609.06119. URL: http://arxiv.org/abs/1609.06119 (visited on 11/24/2022). preprint.

[95] Belle II Collaboration. *Exclusive $B \to X_u\ell\nu_\ell$ Decays with Hadronic Full-event-interpretation Tagging in 62.8 fb^{-1} of Belle II Data*. Nov. 1, 2021. DOI: 10.48550/arXiv.2111.00710. URL: http://arxiv.org/abs/2111.00710 (visited on 11/24/2022). preprint.

[96] Alina Manthei et al. *A first calibration of the Belle II semileptonic tag-side reconstruction algorithm using $B \to X\ell\nu$ decays with 62.8 fb^{-1}*. June 11, 2021. URL: https://docs.belle2.org/record/2421. preprint.

[97] Sam Cunliffe, Thomas Kuhr, and Martin Ritter. *Light Releases of Basf2 for High-Level Analysis*. Sept. 7, 2018. URL: https://docs.belle2.org/record/1114. preprint.

[98] Guido Van Rossum and Fred L. Drake. *Python 3 Language Reference*. 2009. URL: https://docs.python.org/3/reference.

[99] Armin Ronacher. *Jinja*. Pallets, Nov. 7, 2022. URL: https://github.com/pallets/jinja (visited on 11/08/2022).

[100] Georg Brandl et al. *The Sphinx Documentation Generator*. Version 5.0.1. Sphinx, June 2, 2022. URL: https://github.com/sphinx-doc/sphinx (visited on 11/07/2022).

[101] Alex Rudakov and Ash Berlin-Taylor. *sphinx-argparse*. Oct. 4, 2022. URL: https://github.com/ashb/sphinx-argparse (visited on 11/07/2022).

[102] JSON Schema Organisation. *JSON Schema*. URL: https://json-schema.org/.

[103] Michael Eliachevitch et al. *b2luigi*. Oct. 17, 2022. URL: https://github.com/nils-braun/b2luigi (visited on 11/08/2022).

[104] Atlassian Corporation. *BitBucket*. URL: https://bitbucket.org.

[105] CLEO Collaboration. "Study of the q^2-dependence of $B \to \pi\ell\nu$ and $B \to \rho(\omega)\ell\nu$ Decay and Extraction of $|V_{ub}|$". In: *Physical Review D* 68.7 (Oct. 9, 2003), p. 072003. DOI: 10.1103/PhysRevD.68.072003. URL: http://arxiv.org/abs/hep-ex/0304019 (visited on 01/03/2023).

[106] BaBar Collaboration. "Measurement of the $B^0 \to \pi\ell\nu$ Form-Factor Shape and Branching Fraction, and Determination of $|V_{ub}|$ with a Loose Neutrino Reconstruction Technique". In: *Physical Review Letters* 98.9 (Feb. 28, 2007), p. 091801. DOI: 10.1103/PhysRevLett.98.091801. URL: http://arxiv.org/abs/hep-ex/0612020 (visited on 01/16/2023).

[107] BaBar Collaboration. "Measurements of $B \to \{\pi, \eta, \eta'\}\ell\nu$ Branching Fractions and Determination of $|V_{ub}|$ with Semileptonically Tagged B Mesons". In: *Physical Review Letters* 101.8 (Aug. 22, 2008), p. 081801. DOI: 10.1103/PhysRevLett.101.081801. URL: http://arxiv.org/abs/0805.2408 (visited on 05/28/2020).

[108] Belle Collaboration. "Study of Exclusive $B \to X_u\ell\nu$ Decays and Extraction of $|V_{ub}|$ using Full Reconstruction Tagging at the Belle Experiment". In: *Physical Review D* 88.3 (Aug. 7, 2013), p. 032005. DOI: 10.1103/PhysRevD.88.032005. URL: http://arxiv.org/abs/1306.2781 (visited on 01/16/2023).

[109] Belle II Collaboration. *Measurement of the photon-energy spectrum in inclusive $B \to X_s\gamma$ decays identified using hadronic decays of the recoil B meson in 2019-2021 Belle II data*. Oct. 18, 2022. DOI: 10.48550/arXiv.2210.10220. URL: http://arxiv.org/abs/2210.10220 (visited on 01/16/2023). preprint.

[110] Belle II Collaboration. *Reconstruction of $B \to \rho\ell\nu_\ell$ decays identified using hadronic decays of the recoil B meson in 2019 – 2021 Belle II data*. Nov. 28, 2022. DOI: 10.48550/arXiv.2211.15270. URL: http://arxiv.org/abs/2211.15270 (visited on 01/16/2023). preprint.

[111] Belle II Collaboration. *Study of Exclusive $B \to \pi e^+ \nu_e$ Decays with Hadronic Full-event-interpretation Tagging in 189.3 fb^{-1} of Belle II Data*. Sept. 21, 2022. DOI: 10.48550/arXiv.2206.08102. URL: http://arxiv.org/abs/2206.08102 (visited on 01/16/2023). preprint.

[112] Björn O. Lange, Matthias Neubert, and Gil Paz. "Theory of charmless inclusive B decays and the extraction of V_{ub}". In: *Physical Review D* 72.7 (Oct. 31, 2005), p. 073006. DOI: 10.1103/PhysRevD.72.073006. URL: https://link.aps.org/doi/10.1103/PhysRevD.72.073006 (visited on 01/16/2023).

[113] Carlos Ramirez, John F. Donoghue, and Gustavo Burdman. "Semileptonic $b \to u$ decay". In: *Physical Review D* 41.5 (Mar. 1, 1990), pp. 1496–1503. DOI: 10.1103/PhysRevD.41.1496. URL: https://link.aps.org/doi/10.1103/PhysRevD.41.1496 (visited on 08/28/2023).

[114] Markus Prim. *eFFORT*. Zenodo, July 29, 2020. DOI: 10.5281/zenodo.3965699. URL: https://zenodo.org/record/3965699 (visited on 04/29/2022).

[115] Markus Röhrken. "Retraining of the FEI: Unbiasing the $\cos\theta_{BD\ell}$ distributions". Semileptonic and leptonic working group meeting. Oct. 5, 2021. URL: https://indico.belle2.org/event/5483/#6-sl-fei-retraining-to-unbias.

[116] Alon Hershenhorn, Torben Ferber, and Christopher Hearty. *ECL Shower Shape Variables Based on Zernike Moments*. Jan. 22, 2017. URL: https://docs.belle2.org/record/454. preprint.

[117] Patrick Koppenburg. "Statistical biases in measurements with multiple candidates". Aug. 21, 2019. URL: http://arxiv.org/abs/1703.01128 (visited on 01/30/2020).

[118] William Sutcliffe. *PIDvar*. Version 0.1.1. Nov. 30, 2021. URL: https://stash.desy.de/users/sutclw/repos/pidvar.

[119] Sviatoslav Bilokin. *Systematic Corrections Framework*. July 3, 2019. URL: https://syscorrfw.readthedocs.io.

[120] Belle II Lepton ID Group. *Muon and electron identification performance with* 189 fb^{-1} *of Belle II data*. Apr. 21, 2021. URL: https://docs.belle2.org/record/2340 (visited on 09/05/2023). preprint.

[121] Saurabh Sandilya and Alan Schwartz. *Study of Kaon and Pion Identification Performances in Phase III data with* D^{*+} *sample*. July 19, 2019. URL: https://docs.belle2.org/record/1558. preprint.

[122] Taichiro Koga. *Momentum dependent* π^0 *efficiency measurement with D decays*. Sept. 3, 2020. URL: https://docs.belle2.org/record/2096. preprint.

[123] Chaoyi Lyu. "Slow π Tracking Efficiency". May 2022 Belle II General Meeting. May 12, 2022. URL: https://indico.belle2.org/event/6621/contributions/35173/attachments/16622/24779/Face-to-Face%20tracking%20meeting%20202022.pdf.

[124] C. Glenn Boyd, Benjamin Grinstein, and Richard F. Lebed. "Constraints on Form Factors For Exclusive Semileptonic Heavy to Light Meson Decays". In: *Physical Review Letters* 74.23 (June 5, 1995), pp. 4603–4606. DOI: 10.1103/PhysRevLett.74.4603. URL: http://arxiv.org/abs/hep-ph/9412324 (visited on 01/16/2023).

[125] Daniel Ferlewicz, Phillip Urquijo, and Eiasha Waheed. "Revisiting fits to $B^0 \to D^{*-}\ell^+\nu_\ell$ to measure $|V_{cb}|$ with novel methods and preliminary LQCD data at non-zero recoil". In: *Physical Review D* 103.7 (Apr. 29, 2021), p. 073005. DOI: 10.1103/PhysRevD.103.073005. URL: http://arxiv.org/abs/2008.09341 (visited on 05/03/2022).

[126] Belle Collaboration. "Measurement of the decay $B \to D\ell\nu_\ell$ in fully reconstructed events and determination of the Cabibbo-Kobayashi-Maskawa matrix element $|V_{cb}|$". In: *Physical Review D* 93.3 (Feb. 25, 2016), p. 032006. DOI: 10.1103/PhysRevD.93.032006. URL: http://arxiv.org/abs/1510.03657 (visited on 05/03/2022).

[127] Florian Bernlochner. *Private Communication*. E-mail.

[128] Belle II Collaboration. *A calibration of the Belle II hadronic tag-side reconstruction algorithm with $B \to X\ell\nu$ decays*. Aug. 17, 2020. DOI: 10.48550/arXiv.2008.06096. URL: http://arxiv.org/abs/2008.06096 (visited on 01/16/2023). preprint.

[129] Florian Bernlochner et al. *Tests of lepton universality in angular observables of hadronically tagged $B \to D^*\ell\nu$ decays at Belle II*. May 25, 2022. URL: https://docs.belle2.org/record/3023. preprint.

[130] Roger Barlow and Christine Beeston. "Fitting using finite Monte Carlo samples". In: *Computer Physics Communications* 77.2 (Oct. 1993), pp. 219–228. DOI: 10.1016/0010-4655(93)90005-W. URL: https://linkinghub.elsevier.com/retrieve/pii/001046559390005W (visited on 10/28/2019).

[131] William Sutcliffe. *BinFit*. Jan. 15, 2020. URL: https://stash.desy.de/users/sutclw/repos/binfit.

[132] Claude Bourrely, Irinel Caprini, and Laurent Lellouch. "Model-independent description of $B \to \pi l\nu$ decays and a determination of $|V_{ub}|$". In: *Physical Review D* 82.9 (Nov. 24, 2010), p. 099902. DOI: 10.1103/PhysRevD.82.099902. URL: http://arxiv.org/abs/0807.2722 (visited on 01/25/2023).

[133] I. Caprini, L. Lellouch, and M. Neubert. "Dispersive Bounds on the Shape of $B \to D^{(*)}\ell\nu$ Form Factors". In: *Nuclear Physics B* 530.1-2 (Oct. 1998), pp. 153–181. DOI: 10.1016/S0550-3213(98)00350-2. URL: http://arxiv.org/abs/hep-ph/9712417 (visited on 01/16/2023).

[134] Eric O. Lebigot. *uncertainties*. Jan. 8, 2023. URL: https://github.com/lebigot/uncertainties (visited on 01/16/2023).

[135] The DarkMachines High Dimensional Sampling Group et al. "A Comparison of Optimisation Algorithms for High-Dimensional Particle and Astrophysics Applications". In: *Journal of High Energy Physics* 2021.5 (May 2021), p. 108. DOI: 10.1007/JHEP05(2021)108. URL: http://arxiv.org/abs/2101.04525 (visited on 03/08/2023).

[136] Glen Cowan et al. "Asymptotic Formulae for Likelihood-Based Tests of New Physics". In: *The European Physical Journal C* 71.2 (Feb. 2011), p. 1554. DOI: 10.1140/epjc/s10052-011-1554-0. URL: http://arxiv.org/abs/1007.1727 (visited on 02/23/2023).

[137] GAMBIT Collaboration. "GAMBIT: The Global and Modular Beyond-the-Standard-Model Inference Tool". In: *The European Physical Journal C* 77.11 (Nov. 2017), p. 784. DOI: 10.1140/epjc/s10052-017-5321-8. URL: http://arxiv.org/abs/1705.07908 (visited on 07/10/2019).

[138] GAMBIT Collaboration. "GAMBIT and Its Application in the Search for Physics Beyond the Standard Model". In: *Progress in Particle and Nuclear Physics* 113 (July 2020), p. 103769. DOI: 10.1016/j.ppnp.2020.103769. URL: http://arxiv.org/abs/1912.04079 (visited on 02/23/2023).

[139] Csaba Balázs et al. "ColliderBit: A GAMBIT Module for the Calculation of High-Energy Collider Observables and Likelihoods". In: *The European Physical Journal C* 77.11 (Nov. 2017), p. 795. DOI: 10.1140/epjc/s10052-017-5285-8. URL: http://arxiv.org/abs/1705.07919 (visited on 02/21/2023).

[140] Florian U. Bernlochner et al. "FlavBit: A GAMBIT Module for Computing Flavour Observables and Likelihoods". In: *The European Physical Journal C* 77.11 (Nov. 2017), p. 786. DOI: 10.1140/epjc/s10052-017-5157-2. URL: http://arxiv.org/abs/1705.07933 (visited on 02/03/2020).

[141] The GAMBIT Models Workgroup: et al. "SpecBit, DecayBit and PrecisionBit: GAMBIT Modules for Computing Mass Spectra, Particle Decay Rates and Precision Observables". In: *The European Physical Journal C* 78.1 (Jan. 2018), p. 22. DOI: 10.1140/epjc/s10052-017-5390-8. URL: http://link.springer.com/10.1140/epjc/s10052-017-5390-8 (visited on 02/23/2023).

[142] The GAMBIT Scanner Workgroup et al. "Comparison of Statistical Sampling Methods with ScannerBit, the GAMBIT Scanning Module". In: *The European Physical Journal C* 77.11 (Nov. 2017), p. 761. DOI: 10.1140/epjc/s10052-017-5274-y. URL: http://arxiv.org/abs/1705.07959 (visited on 02/21/2023).

[143] Torbjörn Sjöstrand et al. "An Introduction to PYTHIA 8.2". In: *Computer Physics Communications* 191 (June 2015), pp. 159–177. DOI: 10.1016/j.cpc.2015.01.024. URL: http://arxiv.org/abs/1410.3012 (visited on 02/23/2023).

[144] F. Mahmoudi. "SuperIso: A program for calculating the isospin asymmetry of $B \to K^*\gamma$ in the MSSM". In: *Computer Physics Communications* 178.10 (May 2008), pp. 745–754. DOI: 10.1016/j.cpc.2007.12.006. URL: http://arxiv.org/abs/0710.2067 (visited on 02/23/2023).

[145] Torsten Bringmann et al. "DarkBit: A GAMBIT Module for Computing Dark Matter Observables and Likelihoods". In: *The European Physical Journal C* 77.12 (Dec. 2017), p. 831. DOI: 10.1140/epjc/s10052-017-5155-4. URL: http://arxiv.org/abs/1705.07920 (visited on 02/23/2023).

[146] Andrei Angelescu et al. "On the single leptoquark solutions to the B-physics anomalies". In: *Physical Review D* 104.5 (Sept. 14, 2021), p. 055017. DOI: 10.1103/PhysRevD.104.055017. URL: http://arxiv.org/abs/2103.12504 (visited on 05/29/2023).

[147] Svjetlana Fajfer and Nejc Košnik. "Vector leptoquark resolution of R_K and $R_{D^{(*)}}$ puzzles". In: *Physics Letters B* 755 (Apr. 2016), pp. 270–274. DOI: 10.1016/j.physletb.2016.02.018. URL: http://arxiv.org/abs/1511.06024 (visited on 05/29/2023).

[148] Sanjay Bloor et al. "The GAMBIT Universal Model Machine: From Lagrangians to Likelihoods". In: *The European Physical Journal C* 81.12 (Dec. 2021), p. 1103. DOI: 10.1140/epjc/s10052-021-09828-9. URL: http://arxiv.org/abs/2107.00030 (visited on 02/21/2023).

[149] Wolfram Research, Inc. *Mathematica*. Version 13.2. 2022. URL: https://www.wolfram.com/mathematica.

[150] Adam Alloul et al. "FeynRules 2.0 - A Complete Toolbox for Tree-Level Phenomenology". In: *Computer Physics Communications* 185.8 (Aug. 2014), pp. 2250–2300. DOI: 10.1016/j.cpc.2014.04.012. URL: http://arxiv.org/abs/1310.1921 (visited on 02/23/2023).

[151] F. Staub. *Sarah*. Nov. 6, 2012. DOI: 10.48550/arXiv.0806.0538. URL: http://arxiv.org/abs/0806.0538 (visited on 02/23/2023). preprint.

[152] Johan Alwall et al. "MadGraph 5: Going Beyond". In: *Journal of High Energy Physics* 2011.6 (June 2011), p. 128. DOI: 10.1007/JHEP06(2011)128. URL: http://arxiv.org/abs/1106.0522 (visited on 02/23/2023).

[153] Alexander Belyaev, Neil D. Christensen, and Alexander Pukhov. "CalcHEP 3.4 for Collider Physics within and beyond the Standard Model". In: *Computer Physics Communications* 184.7 (July 2013), pp. 1729–1769. DOI: 10.1016/j.cpc.2013.01.014. URL: http://arxiv.org/abs/1207.6082 (visited on 02/23/2023).

[154] Jihyun Bhom and Marcin Chrzaszcz. "HEPLike: An Open Source Framework for Experimental Likelihood Evaluation". In: *Computer Physics Communications* 254 (Sept. 2020), p. 107235. DOI: 10.1016/j.cpc.2020.107235. URL: http://arxiv.org/abs/2003.03956 (visited on 02/23/2023).

[155] LHCb Collaboration. "Angular analysis and differential branching fraction of the decay $B_s^0 \to \phi\mu^+\mu^-$". In: *Journal of High Energy Physics* 2015.9 (Sept. 2015), p. 179. DOI: 10.1007/JHEP09(2015)179. URL: http://arxiv.org/abs/1506.08777 (visited on 02/24/2023).

[156] LHCb Collaboration. "Measurement of the $B_s^0 \to \mu^+\mu^-$ branching fraction and effective lifetime and search for $B^0 \to \mu^+\mu^-$ decays". In: *Physical Review Letters* 118.19 (May 11, 2017), p. 191801. DOI: 10.1103/PhysRevLett.118.191801. URL: http://arxiv.org/abs/1703.05747 (visited on 02/24/2023).

[157] CMS Collaboration. "Measurement of properties of $B_s^0 \to \mu^+\mu^-$ decays and search for $B^0 \to \mu^+\mu^-$ with the CMS experiment". In: *Journal of High Energy Physics* 2020.4 (Apr. 2020), p. 188. DOI: 10.1007/JHEP04(2020)188. URL: http://arxiv.org/abs/1910.12127 (visited on 02/24/2023).

[158] ATLAS Collaboration. "Study of the rare decays of B_s^0 and B^0 mesons into muon pairs using data collected during 2015 and 2016 with the ATLAS detector". In: *Journal of High Energy Physics* 2019.4 (Apr. 2019), p. 98. DOI: 10.1007/JHEP04(2019)098. URL: http://arxiv.org/abs/1812.03017 (visited on 02/24/2023).

[159] Janez Brest et al. "Self-Adapting Control Parameters in Differential Evolution: A Comparative Study on Numerical Benchmark Problems". In: *IEEE Transactions on Evolutionary Computation* 10.6 (Dec. 2006), pp. 646–657. DOI: 10.1109/TEVC.2006.872133.