B-Physics in Lattice QCD and at Belle II

BY

Sophie Hollitt

Supervisors:

Paul Jackson

Ross Young
James Zanotti

A thesis submitted towards the degree of

Doctor of Philosophy

at

The University of Adelaide
School of Physical Sciences

October, 2019

Contents

Abstract

Theory and experiment both have key roles to play in our understanding of the Universe. In flavour physics, semileptonic and leptonic decays of B mesons let us access CKM matrix elements, and anomalies in these decay processes offer tantalising hints of new physics. New measurements of these processes with improved statistical and systematic errors are expected over the coming years using data from the Belle II experiment. New lattice calculations (with a greater understanding of QCD effects in the Standard Model) will be needed to support this increase in experimental precision.

Precise measurement of rare processes with unreconstructed energy from neutrinos—such as $B \to D^{(*)}\tau\nu$ and $B \to \tau\nu$—requires the entire $B\overline{B}$ event to be constrained. This is achieved via reconstruction of the companion B meson in the decay, the so-called tag B. In this work, we describe Belle II's reconstruction process for the tag B, and prepare the skimming to collect data for analysts ahead of Belle II data production.

Measurements of $B \to \tau\nu$ can be used to resolve the anomaly between inclusive and exclusive measurements of CKM matrix elements, if a sufficiently precise value of f_B is available from the lattice QCD community. As f_B is often calculated on the lattice via the ratio f_{B_s}/f_B, it is important to understand and control SU(3) breaking effects in the light and strange quarks, and study how these affect extrapolations of f_{B_s}/f_B. In this work, we compute f_B and f_{B_s} using a set of gauge field configurations that break SU(3) flavour in a controlled way, keeping the average of the lighter quark masses held fixed at the physical value.

Thesis Declaration

I certify that this work contains no material which has been accepted for the award of any other agree 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 acknowledge that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

I also 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 seacrh 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.

SIGNED:

DATE: 30/08/2019

Acknowledgements

First and foremost, this work would not be possible without my supervisors Paul Jackson, Ross Young, and James Zanotti. Thank you for the bright idea to combine experiment and lattice into one project so I didn't have to choose. It's been an absolute pleasure working with the three of you these last few years.

I am grateful for the support of the global communities I am a part of. In lattice QCD, I must acknowledge the work of the other members of QCDSF, and all the hard work that's come before me in generating these lattice configurations. In the Belle II Collaboration, a sincere thanks to everyone else who has been developing the software, and to our accelerator- and detector-building colleagues.

More personally, thank you to everyone who's been part of my development as a researcher, both close to home and further away. Thank you to the NLab group at KMI and Nagoya University for hosting me in 2018, and thank you to Racha and Hannah in Canada for working alongside me in our skimming studies. Thank you to the University of Sydney group for our weekly meetings. Thank you to Alex, Kimmy, Rob, Riley, and everyone else that has kept our Adelaide weekly student meetings running. This thesis is dedicated to all of you, to my mentors and to my peers that are like mentors to me, and to the people I mentor back.

This work also wouldn't be possible without support from outside of physics. Thanks to my board game friends who make sure I get some socialisation every week. Thanks to dance communities around the world, who ensure that I'm not alone anywhere I go. Nagoya Blues, I'll miss you.

Lastly, thanks to my family who have kept me fed through this last stretch of writing, to Jaime who puts up with me not pulling my weight as a housemate during writing time. To Alex L who keeps me sane by letting me get away from it a little, to Finn who keeps me sane by sharing the load.

Thank you.

x

List of Figures

List of Tables

Chapter 1

Introduction

Particle physics is one of our key tools for learning more about the universe and our place in it. Despite all we have learned so far, we still have many unanswered questions about our beginnings and about the way the universe is structured. One of these fundamental questions is about the dominance of matter over antimatter in the universe, and trying to find the source of this large asymmetry.

One of the best sources so far in investigating this question has been the study of charge-parity (CP) violation. CP symmetry suggests that replacing a particle with its antiparticle should result in the same physics if its spatial coordinates are also inverted. Notably, this symmetry can be broken in weak interactions. By noticing and studying these CP violations in weak decays, we have discovered the third generation of quarks in the Standard Model [1, 2], and ultimately entered an era of high luminosity precision tests of physics.

In this introductory chapter we explore the scope of flavour physics with a focus on the CKM matrix elements and decays of the B meson. An overview of the physics relevant to studies of flavour is presented, leading into a discussion of the status of flavour in experiment. Two flavour anomalies—$R(D^{(*)})$ and the difference between inclusive and exclusive measurements of CKM elements—are used to motivate measurements at Belle II and the need for further precision in relevant theory quantities calculated with Lattice QCD.

1.1   The Physics of Flavour

The work in this thesis is primarily concerned with the weak and strong interactions, and the mathematical formulation of the Standard Model in regards to these interactions will be discussed further in Chapter 5. In the context of introducing flavour physics as an area of study, in this Section we provide a brief overview of the particles of the Standard Model and summarise the physics in a more mechanical way. A diagram of the particles of the Standard Model is provided in Figure 1.1.
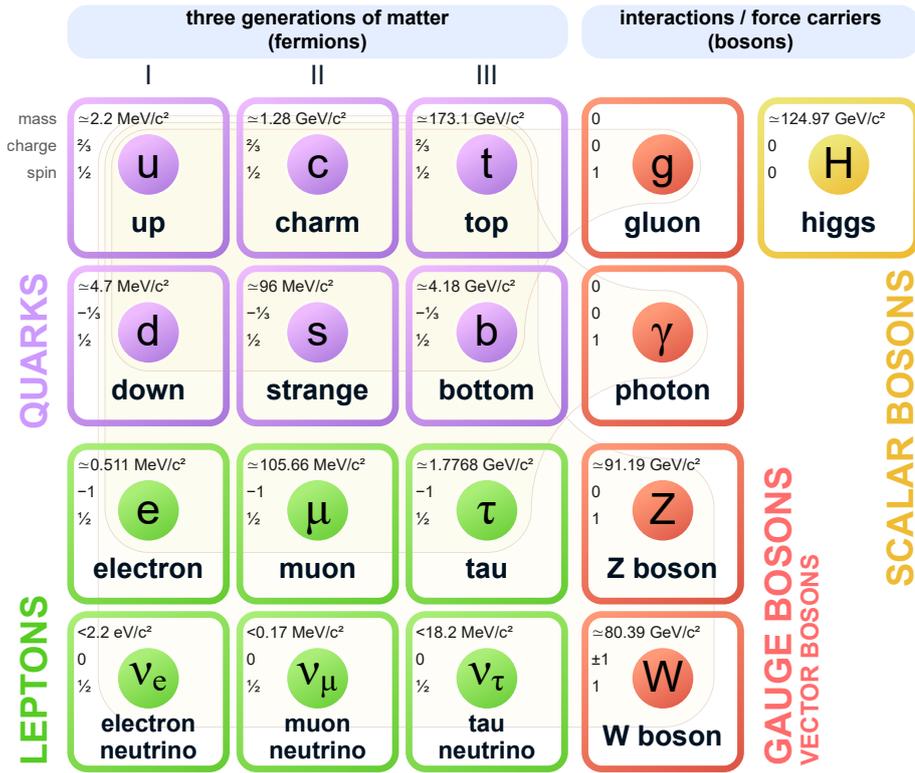
Figure 1.1: Particles of the Standard Model, graphic produced by the Particle Data Group [3].

There are five bosons in the Standard Model, four of which mediate fundamental interactions. The photon is the mediator of electromagnetic interactions based on particle charge, while the Higgs particle and the associated Higgs mechanism are involved in the generation of the mass of the massive Z and W bosons. The Z boson in particular is the primary way that the very light neutral neutrinos interact with other types of matter. Other weak interactions with the W boson will be discussed later in this Section.

The quarks are shown in the upper left corner of Figure 1.1. Each quark possesses a colour charge of the strong interaction, not shown on chart. Each quark can have any one of the three available colour charges, and can change their colour charge by interacting with gluons which also carry colour and anti-colour. Similarly, anti-quarks may have anti-colour charge. The theory of quark and gluon interactions, Quantum Chromodynamics, has two key properties relevant to observation of matter containing quarks: that quarks are only found in groupings that are colourless or have neutral colour charge (colour confinement), and that the strength of interactions in QCD decreases as the energy scale of the interactions increases (asymptotic freedom). In general, this means that at low energies quarks are only found in tightly-bound particles called hadrons, usually a quark-antiquark pair with a colour-anticolour combination (mesons) or a set of three quarks each with a different colour (baryons). As shown, there are a total of six quark types or flavours,

divided into up-type quarks (u, c, t) and down-type quarks (d, s, b). Of the six quark types, the top quark is too heavy and has too short a lifetime compared to the characteristic timescale of the strong interaction to form hadrons, and instead quickly decays via the weak force.

The weak and electromagnetic interactions in the Standard Model can be considered as a unified $SU(2)_L \times U(1)_4$ theory, combining the vector $U(1)_4$ acting on all spinors with the $SU(2)_L$ component acting only on left-handed fermions or right-handed antifermions (see Chapter 5). This $SU(2)_L$ form means that our observable mass eigenstates—which may have mixed chirality—also consist of a mixture of different weak eigenstates. In the lepton sector, the charged leptons are entirely left-handed with right-handed antiparticles, while the neutrinos have mixed mass eigenstates relative to their weak eigenstates. In the quark sector, weak mixing is governed by the Cabibbo-Kobayashi-Maskawa (CKM) matrix

$$\begin{bmatrix} d' \\ s' \\ b' \end{bmatrix} = \begin{bmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{bmatrix} \begin{bmatrix} d \\ s \\ b \end{bmatrix} \tag{1.1}$$

where d, s, b are the observable quark mass eigenstates, and d', s', b' are the weak eigenstates. We note that each CKM element is complex-valued in this formulation.

Our current understanding of quark flavour in the Standard Model is the result of combined efforts of theorists and experimentalists working together since the 1960s. Flavour and indeed quark theories were developed before the charm or heavier quarks had been observed. Studies of strangeness in weak decays led to the proposal of the Glashow-Iliopoulos-Maiani mechanism (GIM mechanism) to explain suppression of tree-level flavour changing neutral currents [4]. CP violation in kaon decays remained unexplained until a third generation of quarks were proposed by Kobayashi and Maskawa in 1973 [1]: the 3×3 unitary quark-mixing matrix (now the CKM matrix) can have a CP-violating phase not possible with a 2×2 matrix in a 4-flavour model. Indeed, the CKM matrix can be parameterised into three pairs of mixing angles (one of which is the Cabibbo angle for mixing between down and strange quarks) plus an additional CP-violating phase, though the Wolfenstein parameterisation (shown below in Equation 1.2) is more common.

$$\begin{bmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3 (\rho - i\eta) \\ -\lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3 (1 - \rho - i\eta) & -A\lambda^2 & 1 \end{bmatrix} + \mathcal{O}(\lambda^4) \tag{1.2}$$

This parametrisation makes explicit the usual approximation that $|V_{tb}| \simeq 1$ to

sufficient precision for most cases. To make the parameterisation exact at all orders, we instead select new parameters $\bar{\rho}$ and $\bar{\eta}$ such that

$$\bar{\rho} + i\bar{\eta} = -\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}. \tag{1.3}$$

Via this parameterisation, the six unitarity relations for the CKM matrix are reduced to one single relation where all terms are the same order in λ and this relation can be visualised as a triangle in the $\bar{\rho}\bar{\eta}$ plane. This triangle is shown in Figure 1.2. We note that the measured sides of the triangle are

$$\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} \text{ and } \frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*}$$

and that the non-vanishing area of the triangle is a result of the presence of the CP violating phase parameter in the CKM matrix. The angles of the triangle can be measured directly in experiment, although each final state comes from a mixture of multiple tree-level Feynman diagrams: $B_s \to K_s^0\rho^0$ or D_sK for the angle γ; $B^0 \to \pi\pi, \pi\rho,$ or $\rho\rho$ for α; $B^0 \to J/\Psi K_s^0$ for β.
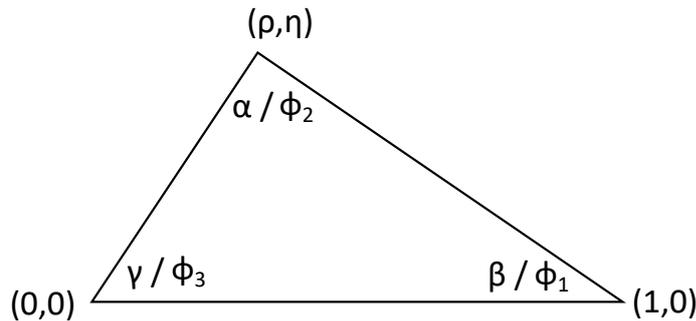
Figure 1.2: A diagram of the CKM unitarity triangle. Angles can be labelled with α, β, γ (BABAR Collaboration) or ϕ_1, ϕ_2, ϕ_3 (Belle Collaboration).

The Summer 2018 combined fit of experimental results to the CKM triangle is shown in Figure 1.3 (reproduced from [5]). The current global fits of CKM matrix elements will be described in further detail in Section 1.3.2, as any discrepancy between observed and predicted values in the CKM triangle may be considered an indication of beyond-Standard-Model effects.

1.1.1 Measuring CKM elements

Four of the nine elements of the CKM matrix can also be measured from processes involving B mesons. The most well-known or most-used decay channel for each is shown in Figure 1.4.

In the context of B Factory experiments, we often consider the "Penguin" decay channel $B^0 \to K^{*0}\gamma$ as our main way to measure V_{ts} and consider $B^0\bar{B}^0$ for V_{td}.
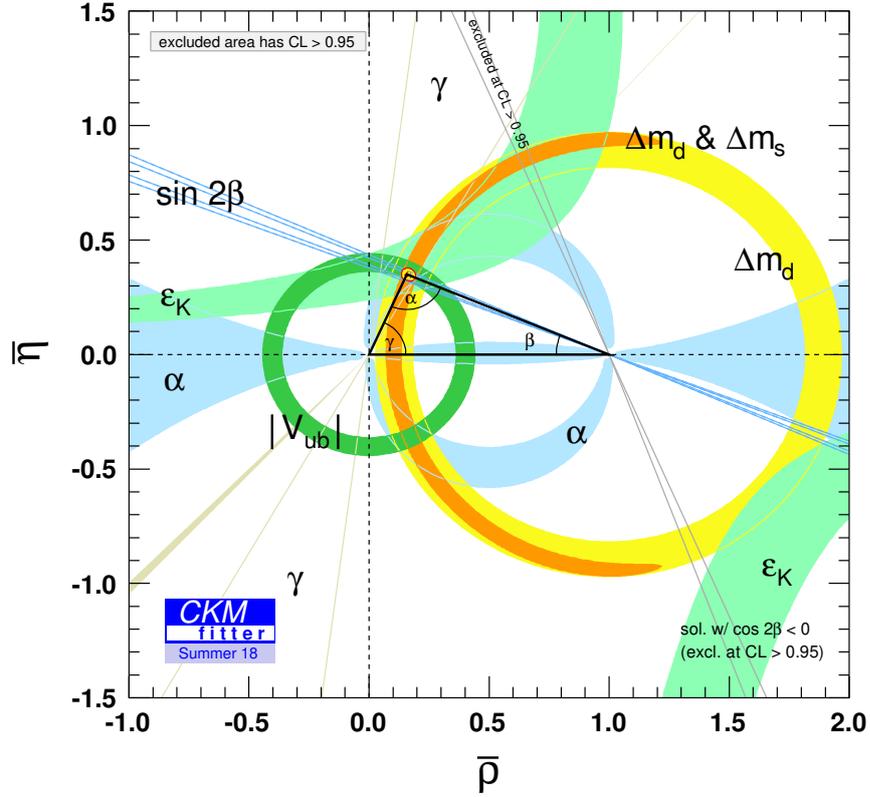
Figure 1.3: Summer 2018 global fit for the CKM unitarity triangle [5]. Different experimental constraints on the position of the apex of the triangle, and on some of the angles, are shown.
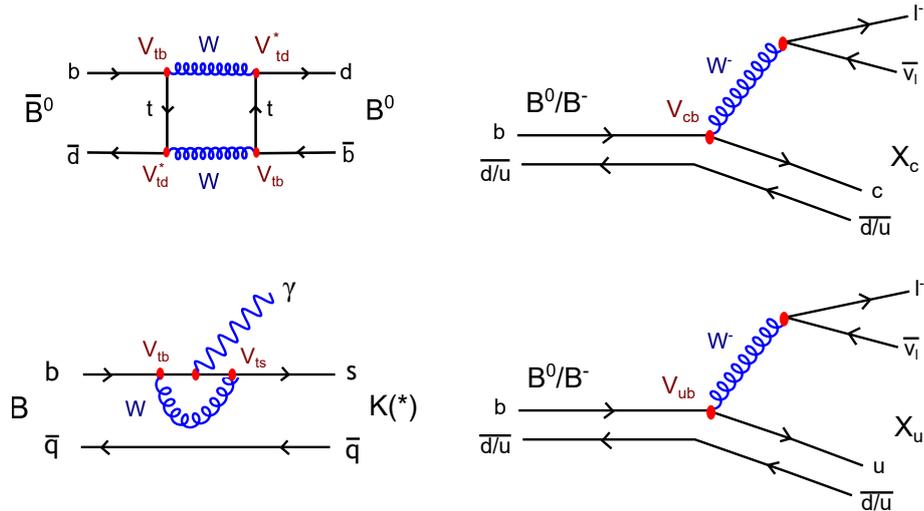
Figure 1.4: Feynman diagrams of decay channels commonly used to calculate CKM matrix elements. X_c and X_u denote any appropriate meson or excited state containing a charm quark or no charm quark respectively. Although not directly shown, quarks in these diagrams are understood to have additional QCD interactions to form bound hadrons.

In general, however, $B_s^0 \bar{B}_s^0$ mixing can also be used to measure V_{ts} and $B^0 \to \pi\gamma$ channels are used alongside $B^0 \to K^{*0}\gamma$ to constrain $|V_{td}/V_{ts}|$. While penguin decay channels are not a key consideration in this thesis, the $B^0\bar{B}^0$ mixing channel is relevant to both the broader physics context of CP violation and the CKM matrix elements. The oscillation frequency Δm_b between the B^0 and $\bar{B}^0$ mesons can be described by

$$\Delta m_b = \underbrace{\frac{G_F^2}{6\pi^2} m_B M_W^2 \eta_B \ S_0(\frac{m_t^2}{M_W^2})}_{\substack{\text{measured values or} \\ \text{known perturbative functions}}} \underbrace{f_B^2 \hat{B}_B}_{\substack{\text{nonperturbative} \\ \text{component}}} |V_{tb}^* V_{td}|^2 \qquad (1.4)$$

where S_0 is a function of the top and W masses, m_t and M_W, respectively; G_F and η_B are other perturbative QCD parameters; f_B is the B meson decay constant and $\hat{B}_B$ its associated Bag parameter [6]. Hence, we see that the measurement depends on both perturbative and nonperturbative theory inputs.

Continuing the discussion of CKM elements, V_{cb} and V_{ub} can both be measured from semileptonic decays of the B meson. In fact, these measurements can be performed using either the branching ratio of an individual B decay channel (exclusive measurement), or by measuring all semileptonic decays without explicitly reconstructing the hadron involved (inclusive measurement). Inclusive channels are often labelled $B \to X_c \ell \nu$ where X_c represents all charmed mesons. The physics of extracting a V_{qb} from an inclusive measurement is often much more complicated than its matching exclusive equivalents. The transition amplitude must be calculated in a way that is sensitive to long range nonperturbative effects of the B meson, rather than the unknown final particle state, and this is achieved using Operator Product Expansion (OPE) methods. This is discussed further in Section 1.3.2.

In the exclusive case, we consider the $B \to \pi\ell\nu$ channel as a simplified example. $|V_{ub}|$ can be extracted from the measured decay rate, with $\ell = e, \mu$, via the simplified relation [6]

$$\frac{d\Gamma}{dq^2} = \frac{G_F^2 |V_{ub}|^2}{24\pi^3} \mathbf{p}_\pi^3 |f_+^{B\pi}(q^2)|^2, \qquad (1.5)$$

where $\mathbf{p}$ is the pion's 3-momentum in the B rest frame, and $f_+^{B\pi}(q^2)$ is one of the two $B \to \pi$ form factors calculated in Lattice QCD. The form factor in particular encodes all of the dynamics of binding quarks into hadrons, and at this level, electromagnetic interactions between the hadron and lepton are not considered. For an inclusive decay, the form factor(s) must be replaced by shape functions that encode the effect of multiple different hadronic final states and resonances.

We can, of course, also calculate the CKM matrix elements from other B decays. Of particular interest to this thesis, leptonic decays such as $B \to \tau\nu$ can also be used to independently determine $|V_{ub}|$ in the absence of new physics. A Feynman

diagram for the $B \to \tau\nu$ decay channel is shown in Figure 1.5, and the branching ratio for this decay is given by

$$\mathcal{BR}(B^+ \to \ell^+\nu) = \frac{G_F^2 M_B M_\ell^2}{8\pi} \left(1 - \frac{M_\ell^2}{M_B^2}\right)^2 \tau_B f_B^2 |V_{ub}|^2 \tag{1.6}$$

for f_B the decay constant and τ_B the B meson lifetime. Much like the semileptonic decay rate, the leptonic branching ratio consists mostly of simple perturbative quantities or experimentally measured values, with all of the QCD information encoded in the single parameter f_B.
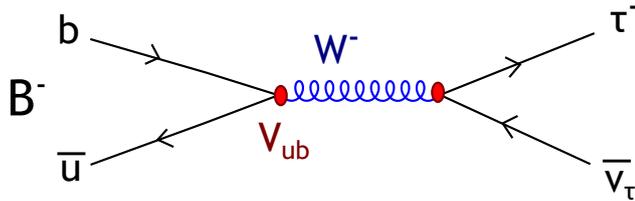
Figure 1.5: Feynman diagram of $B \to \tau\nu$ decay channel. We note that the τ undergoes further weak decay to produce leptonic or hadronic final states.

Having summarised some of the physics of weak decays and CKM elements, further context on the current status of experimental measurements is required to tie these concepts to the motivation of this thesis.

1.2 *B* mesons in experiment

As briefly mentioned in the previous Section, the discovery of the b quark is relatively recent, and studies of the b quark and of flavour physics in general have always encouraged a strong link between theory and experiment. By 1973, a total of three new quarks had been theoretically proposed, but had not yet been observed or discovered.

While the charm quark was first seen in 1974 [7,8], the first sign of the bottom quark in experiment was a few years later at Fermilab in 1977 [2]. A resonance bump for the ground state Υ confirmed the presence of b quarks, and launched studies of the b in earnest. Three years later, the $\Upsilon(4S)$ resonance was discovered by the CLEO collaboration [9], alongside signs that B mesons consisting of one $\bar{b}$ quark and one u or d quark had been produced.

CLEO was joined by the ARGUS experiment [10] in 1982. Both ARGUS and CLEO are experiments at symmetric e^+e^- colliders, able to use the known total energy of the collision to constrain the missing energy and momentum of unreconstructed particles when all other final state particles are collected. Between the two experiments, physicists were able to make first measurements of the CKM matrix

elements and of B decays in general. Many branching ratios were directly measured and branching ratios of rarer decays could be constrained. The first measurement of $B^0\bar{B}^0$ mixing was made at ARGUS [11,12], the larger mixing rate than expected suggesting a larger top quark mass contribution than the 40 GeV expected at that time. Key physics results from CLEO include the first measurement of semileptonic B to charmless decays [13], and the first measurement of the flavour-changing neutral current $b \to s\gamma$ [14]. Both experiments made crucial strides into the measurement of B mesons and the new growing field of flavour physics.

In particular, many experimental researchers collaborated with theorists to develop new ways to explain the behaviour of B mesons as observed. Heavy Quark Effective Theory (HQET) is used to explain the B meson behaviour and particularly $b \to c$ transitions in terms of a screening effect between the heavy and light quarks. In particular, because the Compton wavelength $1/m_Q$ of the heavy quark is smaller than the approximate 1 fm size of the hadron, the low energy properties of the heavy-light hadron are somewhat independent of the exact mass and spin properties of the heavy quark.

These early experiments were key to mapping the physics of the B meson and measuring the CKM matrix elements. Unfortunately, with the low luminosity of these experiments, direct studies of CP violation on low-branching-ratio channels such as $B^0 \to J/\Psi K_s^0$ could not be made. Instead, a higher-luminosity collider would be needed to measure these channels. The time-dependent behaviour of CP violating decays could also be observed by using an asymmetric e^+e^- collision to allow motion of the B mesons and decay products in the detector and improve vertexing in the time domain. Two such higher-luminosity asymmetric collider experiments were planned and constructed: BABAR at SLAC [15], and Belle at KEK [16]. Together, these two experiments are often referred to as the B Factories [6].

Before the B Factories, it was unknown if the lifetime of the B^0 or B^+ meson was longer, and this measurement was crucial for studies of the $B^0\bar{B}^0$ oscillation frequency or time dependent CP measurements. With this lifetime measurement in hand, both experiments were able to proceed with measurements of time dependent CP violation, including measurement of each of the CKM unitarity triangle angles. First results for both experiments were presented at the ICHEP conference in 2000 [17,18], with first publications from both experiments following soon after. In 2001, both experiments published results for $\sin(2\beta)$ within a month of each other [19,20], clearly establishing CP violation in the B meson system.

Over their decade of operation, both experiments advanced the physics of flavour in the B meson sector and contributed large amounts of new B branching ratio information for hadronic, semileptonic, and leptonic decays, but also made several other key measurements in related areas. $D^0\bar{D}^0$ mixing was first observed at the

B factories due to the high number of D mesons produced from $e^+e^- \to c\bar{c}$ inter-actions, and studies of quarkonia in general were an important part of B Factory physics despite not being part of the original physics program. Later in the runs of both experiments, the e^+e^- collisions were also tuned to other energies, lead-ing to the discovery of the lowest-energy $b\bar{b}$ state, the η_b, at BABAR in 2008 [21]. Belle continued until 2010 with studies of the $\Upsilon(5S)$ resonance, before finishing its operations so that the upgrades for Belle II [22] could begin.

Despite these great strides by Belle and BABAR in understanding flavour and CP violations, we are left with several unanswered questions. Not all of the measured branching ratios match with SM expectations, and in fact, some rare decays seem to be suppressed. In the next Section, we will focus on two particular anomalies in semileptonic B decays.

1.3 Flavour anomalies

The flavour sector of the Standard Model currently contains a large number of anomalies. Individual results and the world averages may be 2-4σ from the Stan-dard Model: not significant enough for discovery but significant enough to question our understanding and suggest possible new physics. At this stage, no individual explanation has clear advantage in explaining all of the anomalies seen in b-quark hadronic decays at the B Factories and LHCb [23]. While LHCb is also able to study B_s and B_c mesons as well as b-baryons, in this thesis we will remain focused on anomalies in semileptonic B decays.

Flavour anomalies may suggest new particles involved in B decays at either loop or tree level, but there have also been signs of possible Lepton Flavour Universality violation (LFV). In semileptonic decay channels, we expect the hadronic and leptonic components to factorise so that they are independent of each other, so by taking the ratio of two decay channels with the same hadronic components but different leptons, our Standard Model expectation is that the difference in decay rate should only be affected by lepton mass differences. The ratios $R(D^{(*)})$ and $R(K^{(*)})$ are defined:

$$R(D^{(*)}) = \frac{\mathcal{BR}(B \to D^{(*)}\tau\nu)}{\mathcal{BR}(B \to D^{(*)}\ell\nu)}, \ R(K^{(*)}) = \frac{\mathcal{BR}(B \to K^{(*)}\mu\mu)}{\mathcal{BR}(B \to K^{(*)}ee)} \tag{1.7}$$

where $\ell = e$ or μ.

As both leptons are light, $R(K^{(*)})$ is expected to be 1 (or very close to 1) in the Standard Model, but this is not necessarily the case in experiment. Recent LHCb results for $R(K^{*0})$ are 2.1-2.5 σ below the Standard Model expectation [24]. Further measurements are needed to reveal whether these anomalies persist, and

greater understanding of Standard Model inputs to experiment is needed to make precision measurements of these quantities in future. Measurements of $R(K^{(*)})$ and the rare decay channel $B \to K\nu\nu$ at Belle II may be able to shed further light on this possible LFV in the coming years. The experimental status of $R(D^{(*)})$ is discussed in further detail below.

1.3.1 $R(D^{(*)})$

Measurements of $R(D^{(*)})$ at B factory experiments depend on the ability to reconstruct not only the semileptonic or semitauonic B decays, but also the other remaining B meson in the event, often called B_{tag}. The known centre-of-mass energy of e^+e^- collisions at Belle or BABAR is close to the $\Upsilon(4S)$ mass, and can be used to calculate the missing energy and momentum of the unreconstructed neutrinos in semileptonic or semitauonic decays when all other final state particles are reconstructed. Kinematic information and missing energy information is then used to separate $D\tau\nu$ from $D\ell\nu$ decays, and fits are made in comparison to Monte Carlo shapes to determine yield. The status of experimental measurements for $R(D^{(*)})$ at the B factories and also at LHCb is shown in Figure 1.6 and Figure 1.7.
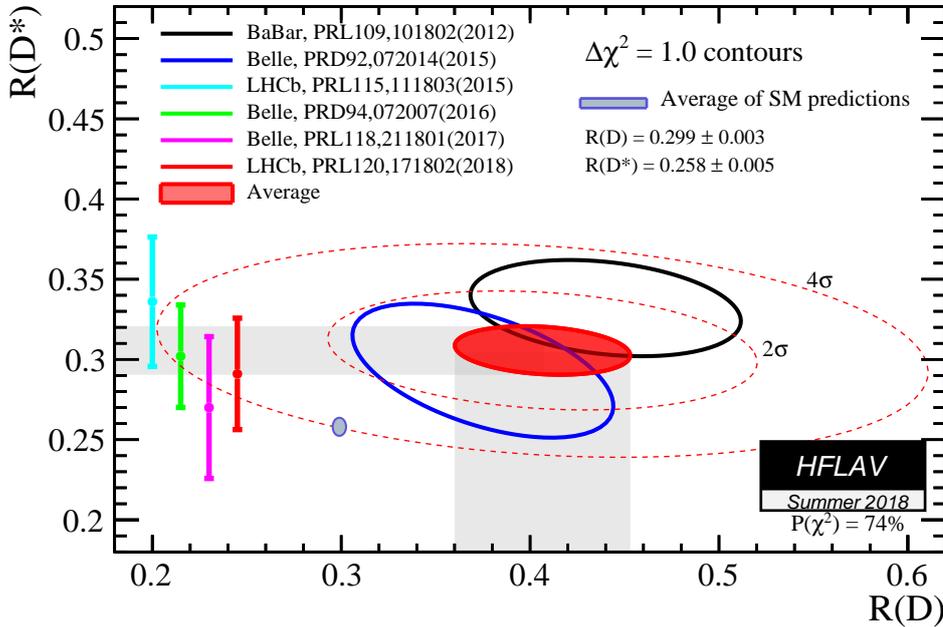
Figure 1.6: Combined measurements of $R(D)$ and $R(D^*)$ by Heavy Flavour Averaging Group (HFLAV) [25]. Status of world average experimental combination and theory combination in Summer 2018.

In 2018, HFLAV presents the tension between the Standard Model average and experimental average as 3.78σ. The main two ellipses from the Belle and BABAR experiments are both from simultaneous measurements of $R(D)$ and $R(D^*)$ using
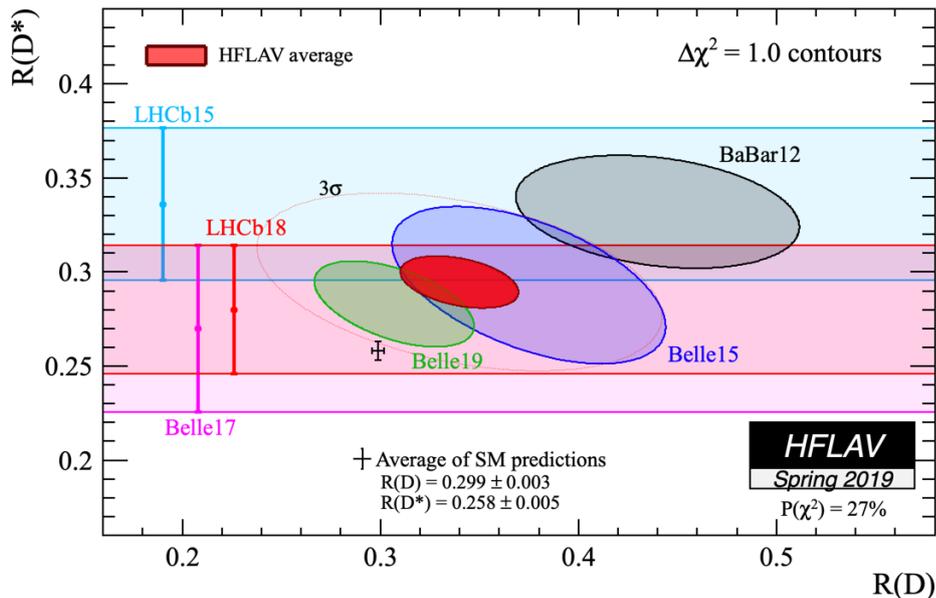
Figure 1.7: Combined measurements of $R(D)$ and $R(D^*)$ by Heavy Flavour Averaging Group [25], from Spring 2019. As in Figure 1.6 except the 2016 $R(D^*)$-only Belle result is updated to the 2019 simultaneous result for $R(D)$ and $R(D^*)$ using semileptonic tags [26].

hadronic B decay modes as tags. In the Spring 2019 average (see Figure 1.7), the new simultaneous $R(D)$ and $R(D^*)$ measurement from Belle using semileptonic B tags is also included. This reduces the tension with the Standard Model somewhat, but the deviation is still $> 3\sigma$. Further experimental input from the Belle II experiment will be crucial to investigate this anomaly further, and if there is new physics, learning more about what that new physics might be. In particular, improving tagging efficiency to collect larger samples of $B \rightarrow D\tau\nu$ decays will be a key aspect for increasing statistics in new measurements.

Many new physics models have been proposed to address flavour anomalies in general. The simplest types of Two Higgs Doublet Models (2HDM) have already had much of their parameter space ruled out by existing measurements at Belle, BABAR, and LHCb [6, 27]. If anomalies in $R(D^{(*)})$ do persist, measurement of the tau polarisation $P_\tau(D^*)$ at Belle II may be able to identify whether the new physics is scalar, vector, or tensor (see prospects in the Belle II Physics Book [28]). Belle II prospects in general will be further discussed in Chapter 8.

1.3.2 V_{ub} and V_{cb}

Measurement of the CKM matrix elements $|V_{ub}|$ and $|V_{cb}|$ at the B factories is mostly performed using semileptonic B decays, as discussed in Section 1.1. Both inclusive

and exclusive measurements are used in the determination of CKM elements, and there is some tension between these two methods, as seen in Figure 1.8.
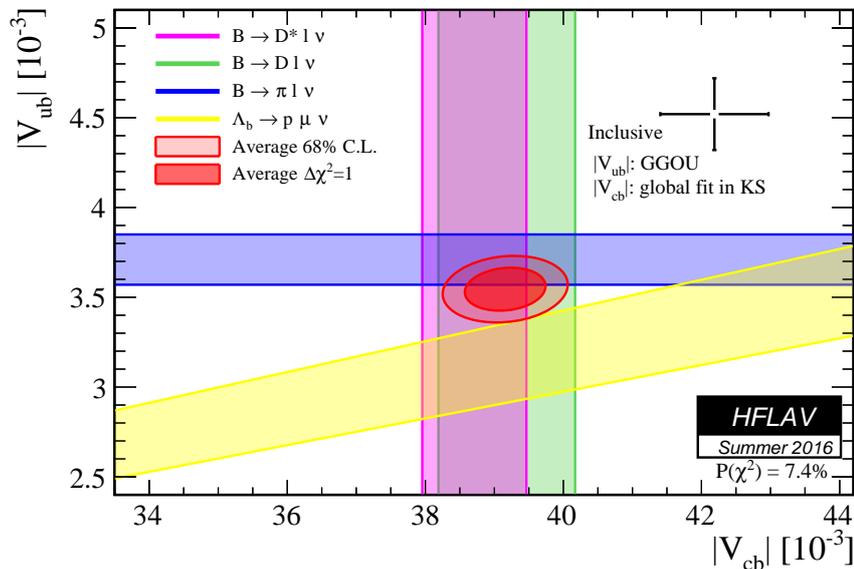
Figure 1.8: Comparison of world-average fit of exclusive measurements of V_{ub} and V_{cb} to world average inclusive measurement by Heavy Flavour Averaging Group [25] in Summer 2016.

Let us first consider the tension between inclusive and exclusive measurements of $|V_{cb}|$. In order to make an exclusive decay measurement (using a similar form to that of the exclusive $B \to \pi\ell\nu$ in Equation 1.5), the relevant form factor(s) must first be extended into a suitable kinematic region by using an appropriate parameterisation of HQET. In the presented HFLAV result, the CLN parameterisation is used to adjust the form factors for $B \to D^*\ell\nu$. An updated result using newer form factors for $B \to D\ell\nu$ and the BGL parameterisation matches better with the inclusive measurement [29], but as the parameterisation is new, further research is needed to confirm the result. In fact, studies using fits to the unfolded Belle data [30] indicate that the BGL parameterisation has unexpectedly large $\mathcal{O}(1/m)$ corrections, and that the CLN parameterisation is better matched to current theory results [31].

Aside from theoretical systematics, these measurements of $B \to D^{(*)}\ell\nu$ also have experimental systematics from the calibration of hadronic tag modes and from the unknown branching ratio to $B \to D^{**}\ell\nu$ modes. At Belle II, these effects could be reduced via direct measurement of and calibration for these effects. In addition, some of the HQET expansion coefficients could be directly measured with sufficient data, which may reveal a solution to the parameterisation problem.

Turning to $|V_{ub}|$, both the inclusive and exclusive measurements are difficult due to the lower branching fraction of $B \to X_u\ell\nu$ modes compared to X_c modes, and the X_c mode background in much of the X_u parameter space. By selecting

only events with $m_{X_u} < m_D$ to avoid this experimental background, the inclusive study is conducted in a difficult kinematic region theoretically. The nonperturbative dynamics and momentum distribution of the b quark—represented by one or more shape functions—are crucial to the extraction of $|V_{ub}|$, but at the time of Belle these dynamics could not be experimentally measured. There are currently four different models for the shape functions, but at Belle II, the leading order shape function could be extracted from fits to $B \to X_s\gamma$.

As briefly mentioned in Section 1.1, another way to obtain an independent measurement of V_{ub} is to instead measure the branching ratio of the $B \to \tau\nu$ rare decay. With increased statistics at Belle II, a competitive measurement of V_{ub} from $B \to \tau\nu$ should be possible in the absence of new physics. Compared to the inclusive semileptonic decay, this is a much simpler measurement from a theory perspective as we do not need to take into account multiple excited hadron states. Such a measurement is dependent on a sufficiently precise determination of the Standard Model B meson decay constant, most reliably derived from Lattice QCD.

1.4 Theory and Flavour

Increased statistics in measurements of branching ratios can only take us so far in terms of improved precision in CKM matrix elements. At some point, theory systematics must be addressed. The Belle II Theory Interface Platform [28] has allowed collaboration between theorists and experimentalists in order to set targets for advances in theory to match the future statistical error from experiment.

Non-perturbative effects in QCD—and soon also QED contributions—will need to be understood in greater detail in order to search for new physics deviations from the Standard Model. Many theory parameters used in experiment now have a target of around 1% error with good control of systematics to reduce these errors below expected experimental precision.

Lattice QCD is a key method for simulating nonperturbative effects in QCD. Space and time is divided into a Euclidean grid or 'lattice', and calculations of hadronic matrix elements (including form factors or other physically-relevant quantities) are averaged over multiple different configurations of quarks and gluons in the vacuum. The specific construction of Lattice QCD will be discussed further in Chapter 5. In the past, Lattice QCD simulations have often used heavier up and down quarks than is physical and hence required extrapolations toward the physical mass. Improvements in computing power have led to more realistic simulations being possible, including lighter pion masses but also finer lattice spacings. Crucially, modern lattice QCD calculations can also quantify all sources of uncertainty.

At present, multiple research groups are involved in the study of B mesons on

the lattice, but we focus on developments toward $B \to D^{(*)}$ form factors and B decay constants relevant to $R(D^{(*)})$ and $|V_{qb}|$. The status of various collaborations is presented in Table 1.1, in part collected via the Flavour Lattice Averaging Group review for 2019 [32].

Table 1.1: Current status of lattice QCD calculations of quantities relevant to B meson decays.

Collaboration	lighter quarks	heavy quarks	Observables
RBC/UKQCD	domain wall fermions	domain wall fermions, extrapolation of m_h	f_{B_s}/f_B and $\hat{B}_{B_s}/\hat{B}_{B_d}$ [33], $B_{(s)} \to D^{(*)}_{(s)}\ell\nu$ FF [34]
Fermilab/MILC	HISQ	HISQ	f_B, f_{B_s}, and f_{B_s}/f_B [35]
	asqtad	Fermilab action	$B \to D^*\ell\nu$ at nonzero recoil [36]
HPQCD	HISQ	NRQCD	f_B, f_{B_s}, and f_{B_s}/f_B [37]
LANL-SWME	HISQ	Oktay-Kronfeld (OK) action (improved Fermilab action)	$B \to D^*\ell\nu$ [38]
JLQCD	domain wall fermions	domain wall fermions	$B \to D^{(*)}\ell\nu$ form factor [39]
ETM	twisted mass, Osterwalder-Seiler (OS)	Osterwalder-Seiler (OS)	f_B, f_{B_s}, and f_{B_s}/f_B [40]
QCDSF	clover-Wilson fermions	RHQ Wilson action/Fermilab action	f_{B_s}/f_B (this work)

Notably, a variety of b quark actions and other methodologies are used, producing a variety of different systematics in each calculation. While this makes research into methods of controlling systematics more scattered, variety of methods producing similar results can be thought of as a way of increasing the robustness of the world average result.

Research into new methods to continue increasing precision in lattice QCD estimates is ongoing. As outlined by B2TiP [28], introducing QED corrections into lattice QCD calculations is expected to improve the agreement between theory and experiment in future. Research in this area is just beginning, with the first QED corrections to leptonic decay rates being published fairly recently [41]. Additional methods to directly calculate inclusive quantities in Lattice QCD are also being explored (see overview in [42]).

1.5 Thesis Overview

Flavour physics depends on the relationship between theory and experiment, and further development of theory precision in Lattice QCD is required in preparation for upcoming and future experimental studies. Additionally, updated experimental analysis techniques are needed to gather as many rare B decay modes as possible to attempt to resolve current anomalies. In this thesis, I take a multipronged approach to addressing these issues.

Improving efficiency and purity in B meson reconstruction is a critical part of the Belle II physics program, and I focus on the reconstruction of the tag B for rare decay channels. In Chapter 2, I begin with an overview of the Belle II detector and SuperKEKB collider, and discuss the way that particle decays are reconstructed from detector information. In Chapter 3, the principles behind Belle II's tag B reconstruction software are summarised, while in Chapter 4 we discuss the progress of the development of this software and its performance in simulated data studies as Belle II software progresses.

To support this improved experimental precision for studies of $B \to \tau\nu$ decays in particular, I also present Lattice QCD calculations of the B meson decay constants f_B and f_{B_s} at a variety of lattice spacings and pion masses, with a focus on SU(3) symmetry breaking effects. In Chapter 5, I discuss the theoretical formulation of the Standard Model and provide a background in the principles of Lattice QCD, including a discussion of the quark action chosen to produce b quarks in this research. In Chapter 6, we tune this b quark action and combine lattice correlation functions in order to calculate the decay constant. Results for SU(3) flavour symmetry breaking in the decay constants are presented in Chapter 7.

Finally, in Chapter 8 I summarise the work presented in this thesis and provide an overview of possible extensions.

Chapter 2

Belle II Overview

The Belle II detector [22] is built on the site of the original Belle experiment at the High Energy Accelerator Research Organisation (KEK) site in Tsukuba, Japan. Data collection at Belle ceased in 2010 to allow the KEKB accelerator to be upgraded to SuperKEKB [43] ready for Belle II. Despite being built in the same location, Belle II is almost completely new, designed to operate at a much higher luminosity.

In the past year, the final commissioning and testing of Belle II has been completed and data taking has begun. Phase 2 testing of the detector including first collisions of the SuperKEKB beams began in April 2018 [44] and ended in July 2018, providing valuable testing of the individual subdetectors, data acquisition systems and triggers under e^+e^- collision conditions ahead of planned Phase 3 operation. Phase 3 began in March 2019.

In this chapter, we begin with an overview of the $\Upsilon(4S)$ resonance and background conditions at SuperKEKB, before discussing the parts of the Belle II detector and basic particle reconstruction software.

2.1 The Upsilon 4S resonance

For the semileptonic and leptonic decay channels of interest to this thesis, data is most optimally collected at the $\Upsilon(4S)$ resonance. The production cross section of hadrons from the e^+e^- collision is shown in Figure 2.1, with the $B\bar{B}$ threshold energy shown. The $\Upsilon(4S)$ itself is just above this $B\bar{B}$ threshold, so that B mesons from the $\Upsilon(4S)$ decay are produced with very little momentum above that of their shared centre of mass. This resonance decays to $B\bar{B}$ 96% of the time [3], split approximately equally between B^+B^- and $B^0\bar{B}^0$ pairs.

We note that these Υ resonances sit on top of a background of $e^+e^- \to q\bar{q}$ production, referred to as continuum background. Although some Belle II analyses specifically study decays in the continuum, for example from $c\bar{c}$, for most studies the continuum is considered an additional background which must be suppressed. In
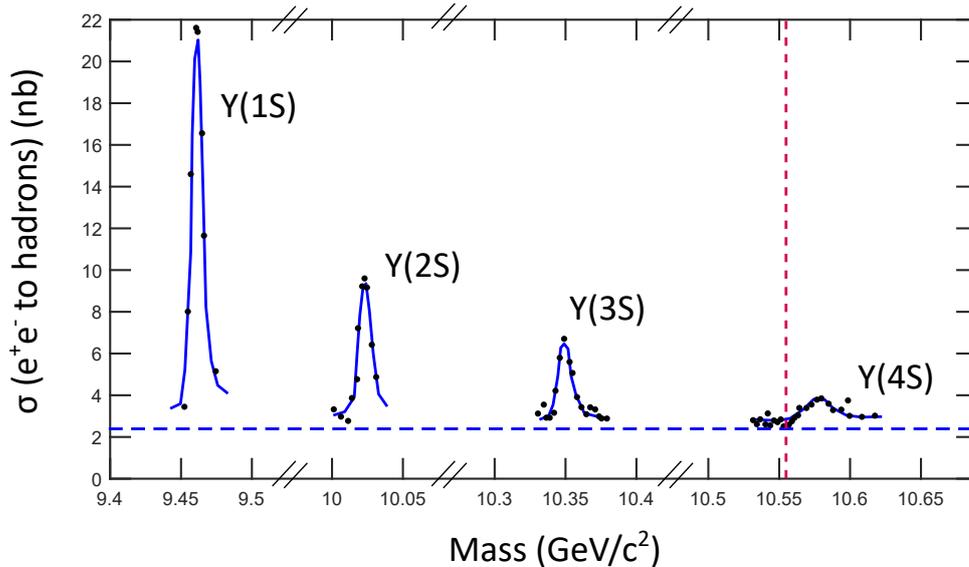
Figure 2.1: Production cross section of hadrons from e^+e^- collision. The first four Υ resonances are shown, above the so-called $q\bar{q}$ continuum background. The $B\bar{B}$ threshold is also shown. Plot is a reproduction of CESR data [9, 45–47] based on a similar reproduction shown in [48].

order to calibrate the amount of continuum in the data, the SuperKEKB collision energy may be tuned just under the resonance, so that data with and without $\Upsilon(4S)$ may be compared.

2.2   SuperKEKB

The SuperKEKB accelerator consists of rings of radius 3 km and some initial linear accelerator apparatus, shown in Figure 2.2. Multiple linear accelerator units are used in the injector linear accelerator (LINAC), and these units can be swapped in and out of use to change the energies of each beam. For operation at the $\Upsilon(4S)$ resonance, the electron and positron beams are accelerated to energies of 7 GeV and 4 GeV respectively. The electron and positron beams have a single overlap point at the site of the Belle II detector.

Compared to the previous KEKB accelerator, SuperKEKB is designed to have a 40× higher luminosity. This will be achieved using the 'Nano-beam' scheme [49], where the size of the overlap of the electron and positron beams is greatly reduced. The size of this overlap region is much smaller than the bunch length produced by the accelerator. A visual comparison of the KEKB scheme and the Nano-Beam scheme is shown in Figure 2.3.

The 40× luminosity figure compared to KEKB can be inferred from the lumi-
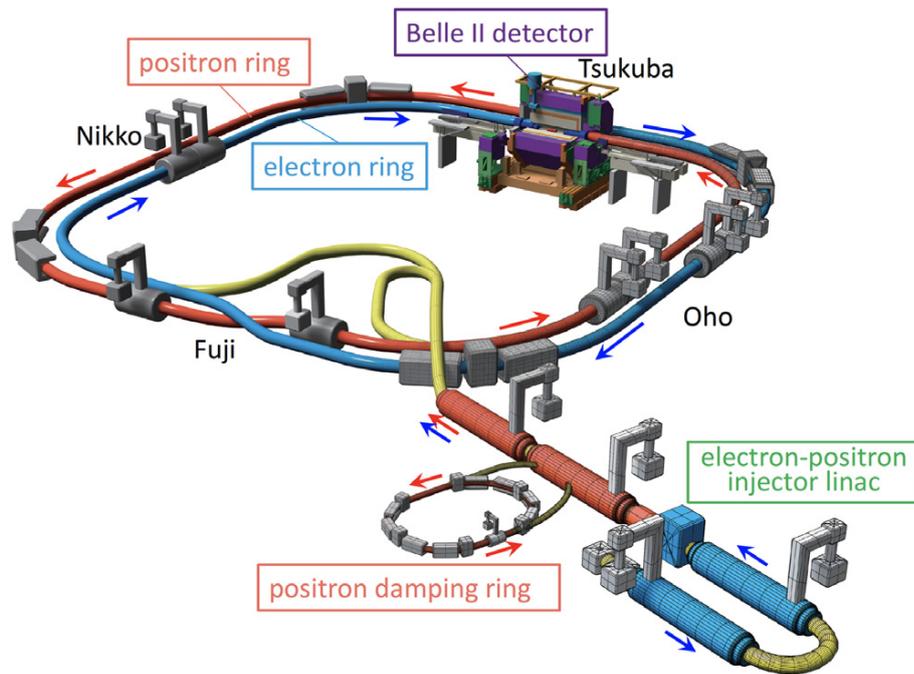
Figure 2.2: Schematic view of SuperKEKB, sourced from [43]
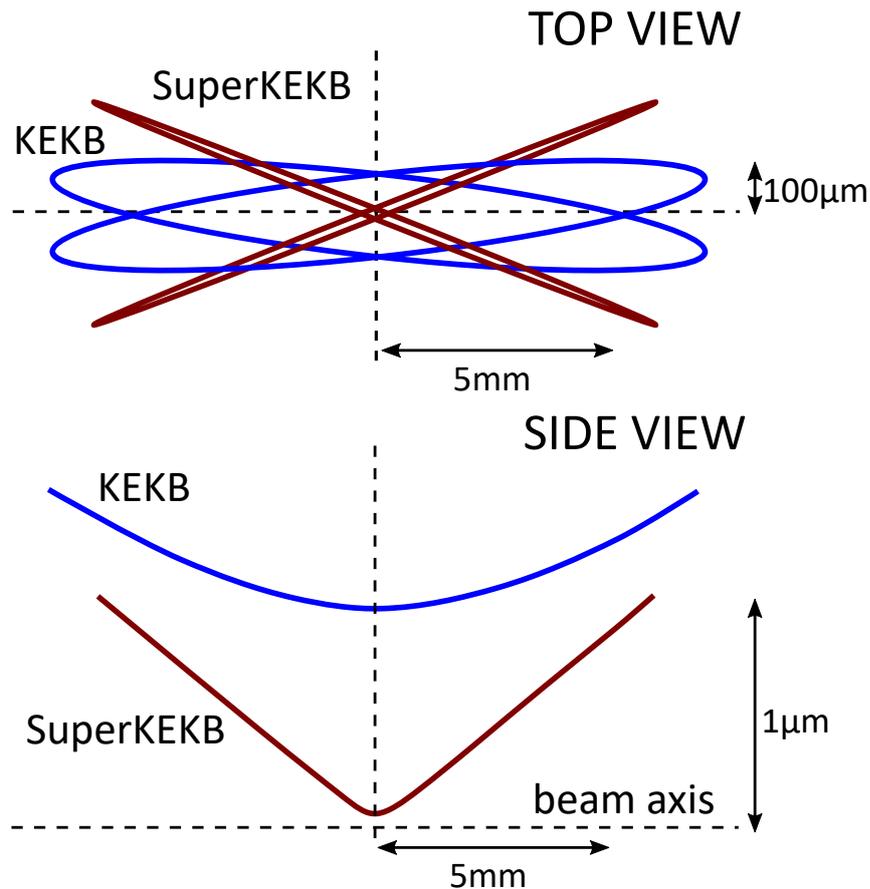
Figure 2.3: Comparison of KEKB beam scheme and SuperKEKB Nano-beam scheme, adapted from [43]. Top: The crossing angle between the SuperKEKB beams is larger than for KEKB, while the bunches themselves are much narrower. Bottom: Comparison of the vertical beam size at KEKB and SuperKEKB. The SuperKEKB beam is much more tightly focused at the interaction point.

nosity equation

$$L = \frac{\gamma_\pm}{2er_e} \left(\frac{I_\pm \xi_\pm}{\beta^*_{y\pm}} \right) \left(\frac{R_L}{R_{\xi_y}} \right) \tag{2.1}$$

for I the beam current, e the electron charge, r_e the classical electron radius, γ the Lorentz factor, and β^*_y and $\xi_\pm$ the vertical beta function or beam-beam parameter respectively for either the electron $(-)$ or positron $(+)$ beam at the interaction point. The ratio R_L/R_{ξ_y} of reduction factors from the crossing angle and hourglass effects is close to 1. Using the Nano-beam scheme, the β^*_y for SuperKEKB will be $20\times$ smaller than the β^*_y at KEKB, with the additional factor of 2 in luminosity supplied by increased beam currents.

Unfortunately, this increase in luminosity is also expected to increase the amount of background in the Belle II detector, via a number of different background processes. This will be discussed in the next Section.

2.3 Beam-induced background

Beam background is especially relevant to decays including neutrinos. At Belle and BABAR, the extra energy remaining in the calorimeter after both B mesons in the $\Upsilon(4S)$ event are reconstructed peaks at zero for correctly reconstructed events due to the low number of background hits, while incorrectly-reconstructed events may have some extra energy after reconstruction. Instead, due to the large beam background at Belle II, the calorimeter energy will not peak at zero unless background photons can be effectively removed.

While the Belle II detector and SuperKEKB beam pipes are designed to prevent beam background photons from reaching the detector wherever possible, a certain amount of background is unavoidable at high luminosity. Further development of trigger rules or analysis techniques can help to mitigate the impact of background on final analysis results.

In this Section, we discuss the five major sources of beam background at Belle II. Of these background sources, three are related to scattering of the electron and positron beams (Touschek scattering, Beam gas scattering, and radiative Bhabha scattering) and can cause showers in the detector due to the interaction of beam electrons or positrons with the pipe walls or focusing magnets.

Touschek scattering

Touschek scattering is fundamentally a Coulomb scattering between two particles in the same bunch, causing one to gain momentum and one to lose momentum. Under the magnetic acceleration of the beam pipe, these different momenta mean different

curvature and the scattered particles will hit the beam pipe walls and shower. If these showers occur too close to the experiment, background photons from these longitudinally-scattered particles will reach the detector.

As Touschek scattering depends on bunch size and beam size, overall we expect 20x higher rates of Touschek scattering with SuperKEKB than with KEKB. To mitigate the effects of Touschek backgrounds on the Belle II detector, collimators are used to separate out particles with non-matching trajectories before they reach Belle II. Heavy-metal shielding will also be used around the vertex detector to prevent shower particles reaching the detector acceptance.

Beam gas scattering

The effect of beam gas scattering is somewhat similar to Touschek scattering, although the means of production of electrons or positrons with deviant momenta is different. In beam gas scattering, electrons or positrons in a bunch may scatter in the transverse direction from residual gas molecules in the beam pipe, via a Coulomb or Bremsstrahlung process.

Radiative Bhabha scattering

Bhabha scattering is the umbrella term for electromagnetic electron-positron scattering processes. In a radiative Bhabha process, the electron and positron interaction produces a photon which usually travels along the beam pipe. Through emission of photons, radiative Bhabha processes also reduce the energy of the beams, causing them to divert inside the beam-shaping magnets of the detector and produce further photons in showers as an additional effect.

When these photons interact with the iron in the detector or beam pipe, neutrons are produced via the (luminosity-dependent) giant photo-nuclear resonance mechanism. Additional shielding is required around the detector and beam pipes to block this neutron radiation.

Synchrotron radiation

Synchrotron radiation is emitted by either the electron or positron beam under radial acceleration, though as the synchrotron radiation power is dependent on the beam energy and magnetic field strength, this background comes mostly from the electron beam. The beam pipes leading into the Belle II detector are shaped with fins and baffles to block direct synchrotron radiation and scattered photons from reaching the inner detector.

Two photon processes

The final background to consider is two photon processes, as seen in Figure 2.4.
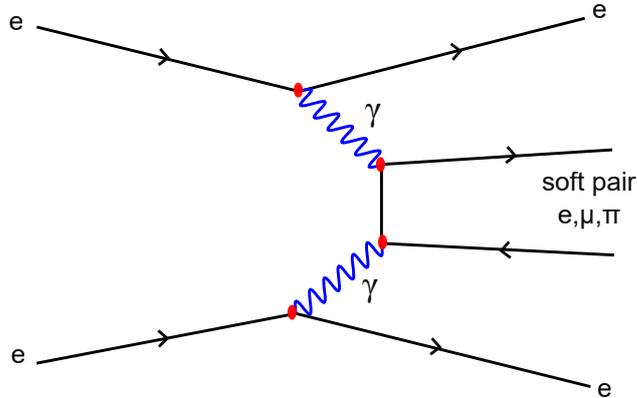
Figure 2.4: Feynman diagram of a two-photon process

While the original electron and positron often have a high enough energy to continue inside the beam pipe instead of interacting directly with the detector, the low momentum (or "soft") fermions produced will spiral through the detector as a result of the constant magnetic field, producing multiple background detector hits.

2.4 Parts of the Belle II detector

A schematic of the full Belle II detector is shown in Figure 2.5. In this section, we present an overview of the Belle II detector as a whole, starting from the outermost parts of the detector and working our way toward the centre. The way that these detectors are used in the context of particle reconstruction is discussed further in Section 2.5.

2.4.1 Outer detector: KLM and Solenoid

The outermost Belle II detector unit is the K_L^0 and muon detector (KLM). The accurate detection of K_L^0 in particular is useful for CP violation studies in $B \rightarrow J/\Psi K_L^0$. High momentum muons are also difficult to accurately measure in other parts of the detector.

The KLM detector unit consists of alternating layers of iron plates and scintillator strips, the scintillator strips being used to collect ionisation tracks and showers from interactions in the iron layers. The Belle detector used glass-electrode resistive plate chamber technology in a similar role, but the increased background rate at Belle II from radiative Bhabha neutron showers in particular, favours scintillators. Aside from providing a large number of interaction lengths to ensure that K_L^0 do decay, the iron plates also serve as a magnetic flux return for the solenoid magnet.
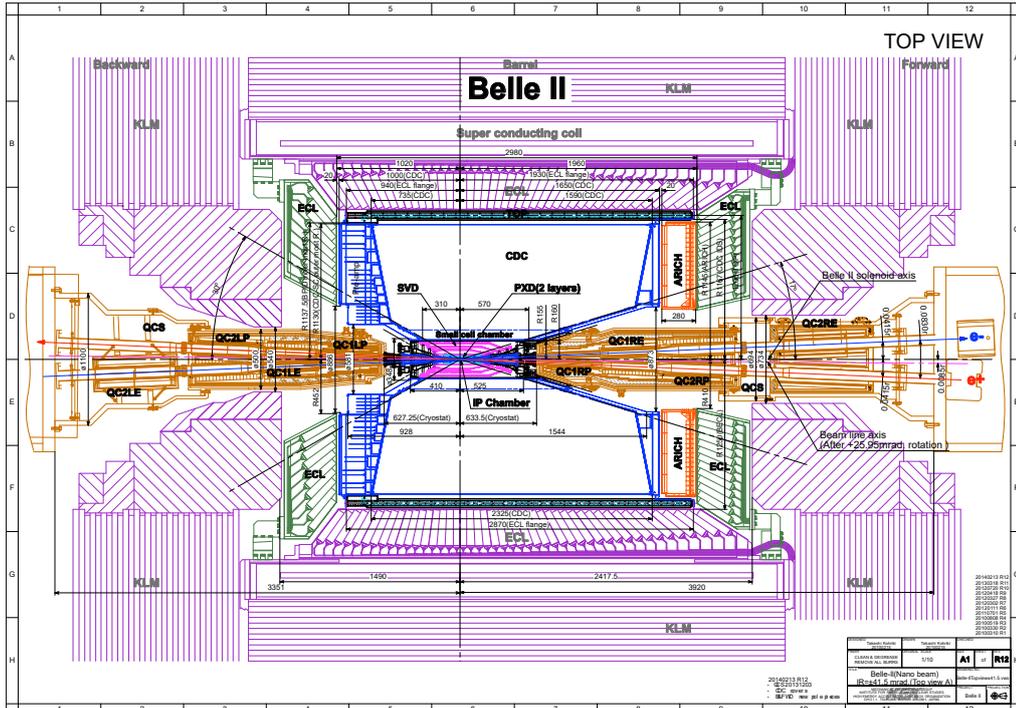
Figure 2.5: Schematic of the Belle II detector, also seen in [28].

The superconducting solenoid itself is inside the yoke of the KLM and iron plates, and is designed to create a constant 1.5 T magnetic field at the centre of the detector.

2.4.2 ECL

The main purpose of the electromagnetic calorimeter (ECL) is to separate electrons from hadrons, and also record photon impacts. The ECL is split into barrel, forward endcap, and backward endcap regions, and together covers approximately 90% of the solid angle of the centre-of-mass system. Each individual segment of the ECL consists of a thallium-doped caesium iodide CsI(Tl) crystal, and while these crystals are re-used from Belle, the read out electronics are completely new to deal with the larger number of background photons and increased luminosity at Belle II.

2.4.3 ARICH and TOP

The Aerogel Ring Imaging Cherenkov detector (ARICH) and the Time Of Propagation counter (TOP) are the two primary particle identifying detectors at Belle II. Both detectors use the shape and intensity of Cherenkov rings to identify particles, although the method used to do so varies between the endcap and barrel detectors.

In the barrel region, the TOP is used to help identify particles. TOP consists of 16 quartz bar detector modules arranged around the barrel, with photosensor equipment at one end. When particles are incident on the TOP, their Cherenkov ring

image propagates along the quartz bar to rows of fast, specially-developed photodetectors. The time and position of arrival of Cherenkov photons at the photodetectors provides information about the particle type, momentum, and position of impact on the TOP.

In ARICH, the Cherenkov rings are detected directly using 2D sensor arrays. To increase the number of Cherenkov photons detected, two 2 cm-thick aerogel layers with slightly different refractive indices are used, such that the Cherenkov rings from both layers form the same ring in the detector.

2.4.4 Central Drift Chamber

The purpose of the Central Drift Chamber (CDC) is to identify the momentum of charged particles by reconstructing their curved path due to the constant magnetic field of the solenoid magnet. The CDC itself consists of 56 layers of wires in either an axial orientation (aligned with the magnetic field) or a stereo (skewed) orientation, inside a chamber filled with a helium and ethane (C_2H_6) gas mixture. By combining information from both types of wires, 3D helical particle tracks may be reconstructed. The wires are more closely-spaced toward the centre of the detector, where a higher density of tracks are expected.

2.4.5 Vertex detector

The Belle II vertex detector (VXD) consists of two parts: the inner silicon Pixel Detector (PXD) and the Silicon Vertex Detector (SVD) for a total of six detector layers (see Figure 2.6). The PXD has two layers of pixelated sensors, and is mounted very close to the interaction point and beam pipe: the first PXD layer starts at a radius of 14 mm from the centre, while the beam pipe's radius is 10 mm. For comparison, this is much closer to the interaction point than the vertex detector in the Belle experiment, and is expected to improve both vertex resolution and reconstruction efficiency.
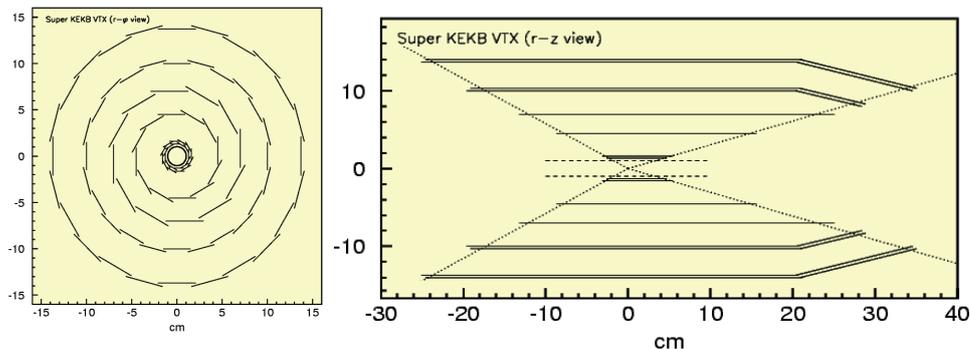
Figure 2.6: VXD schematics from [28].

The four SVD layers in the VXD are silicon strip detectors. Each layer consists of a number of SVD ladders, with the length of the ladders being determined by the radius of the layer (see Figure 2.6). During Belle IIs Phase 2 operation, a test version of the VXD using 1/8th of the SVD ladders was used to assist in testing and calibration, with the remainder of the VXD to be added in Phase 3. Unfortunately due to time constraints, the full PXD was not able to be installed before Phase 3 operation began this year. Instead, the full PXD inner layer and 2/12 of the ladders in the PXD second layer are installed in Belle II at present, and a new full PXD will be assembled and installed in the coming years. [50]

2.5 Particle detection

2.5.1 Triggers

The Belle II trigger system has two levels: the low level hardware-based trigger (L1) and the high level trigger in software (HLT). The L1 trigger copes with the increased event rate at Belle II by implementing firmware level reconstruction of events. Full 3D tracking is used to identify the vertex position, so that tracks originating outside of the interaction region can be suppressed. Particle ID at the trigger level is also implemented, as 3D CDC tracks can be matched to ECL clusters. Additional firmware is used in the ECL trigger to suppress photons from radiative Bhabha events.

$B\bar{B}$ and other $q\bar{q}$ events are triggered with > 3 tracks in the event and a large deposit of energy in the ECL, and this trigger is close to 100 % efficient on $B\bar{B}$ while also suppressing non-radiative Bhabha scattering and dimuon events. However, Belle II's physics program also includes studies of low multiplicity and dark sector events which may have smaller numbers of tracks, and additional triggers are deployed to study these processes while still excluding backgrounds.

After the L1 trigger, the HLT uses offline reconstruction algorithms on a CPU farm to reduce the number of events that will be stored to disk. All detector information is used except the PXD: the full readout from the PXD is too large to be stored or processed, so regions of interest must be identified with HLT.

2.5.2 Tracks and Hits

In order to reconstruct particles in the detector, VXD and CDC hits must be identified as non-background and fit with possible tracks. Most tracks will originate inside the beam pipe, apart from charged tracks coming from neutral particles decaying inside the detector volume (eg. K_S^0). Tracks fit to external particles such as cosmic

ray muons are used to calibrate the alignment of the detector subunits relative to each other.

The first stage of fitting a track is identifying the hits that could belong to possible track candidates. This track finding uses a cellular automata model on hits in the VXD to iterate from short track pieces based on hits in adjacent layers of the VXD into full tracks. Track fitting is also applied to the CDC, and track candidates are merged. With candidates found, tracks are fit with an (approximately) helical path. In order to make fits, a particle mass hypothesis must be applied.

Reconstruction of the decay vertices of neutral particles (called V^0 reconstruction) is applied after the main charged particle track fits. This algorithm pairs oppositely-charged particles and extrapolates back to the innermost detector hit of each track. A vertex reconstruction is then applied. In practice, neutral particles not reconstructed in this stage may also be reconstructed with analysis-level software manually.

2.5.3 Calorimeter reconstruction

Accurately reconstructing neutral particles from deposits in the ECL is a crucial part of analyses involving neutrinos. Photons from $\pi^0 \to \gamma\gamma$ must be separated from beam background photons in order to ensure all particles from the $B\bar{B}$ are correctly reconstructed. In addition, the shapes of the detector showers in the ECL are valuable for determining whether a track matching an ECL deposit is from an electron, muon, hadron, or photon.

An ECL cluster is identified by a central high energy deposit (local maximum among nearest-neighbour crystals), and then the neighbouring crystals are included in the cluster in a 5x5 grid provided there is no overlap with another nearby cluster. Cluster energy is the linear sum of the energies of all included crystals. Particle ID information is collected from the shape of the energy deposits within the cluster.

2.5.4 Particle ID

Particle ID (PID) of an unknown charged particle is determined by combining information from each subdetector. In each of the subdetectors, we fit the distribution of hits with Probability Distribution Functions (PDFs) for each type of possible final state particle: pions, kaons, electrons, muons, protons, and deuterons (d). The quality of the PDF fit is quantified by a log likelihood, and the log likelihood for each subdetector is summed to produce an overall likelihood for each particle type. This value can be used directly to compare hypotheses via a simple log likelihood difference, for example the difference $\log \mathcal{L}_\pi - \log \mathcal{L}_\mu$ can be used to tell us if the pion hypothesis or muon hypothesis is more likely.

In general, however, we don't just need to know which particle hypothesis is most likely, we also want to quantify the probability of whether or not a particle hypothesis is correct. Two different forms of the PID probability have been defined and used within Belle II at different stages of the software. The first PID to consider is the binary PID, defined

$$P(A, B) = \frac{\log \mathcal{L}_A}{\log \mathcal{L}_A + \log \mathcal{L}_B} \tag{2.2}$$

for two mass hypotheses A and B. This binary PID was used for particle ID at Belle II in early software, and was defined to always use the pion for comparison when no particle B was specified (though in the case of PID for the pion, the kaon was used as hypothesis B). Currently, the collaboration uses global PID, defined

$$P(A) = \frac{\log \mathcal{L}_A}{\log \mathcal{L}_\pi + \log \mathcal{L}_K + \log \mathcal{L}_e + \log \mathcal{L}_\mu + \log \mathcal{L}_p + \log \mathcal{L}_d} \tag{2.3}$$

The advantage of global PID is that the relative likelihood of a particle being an electron compared to a muon is more easily available to non-expert analysis users, and that the sum of all possible PID hypotheses for a given particle is 1.

2.6 Belle II Simulation

Belle II simulation and reconstruction is facilitated via the Belle II Analysis Software Framework (`basf2`). This framework is designed to manage the loading of multiple individual modules and other software libraries behind the scenes, so that analysts interact with the software only via python steering files. [51] In this section, we focus on the generation of Monte Carlo (MC) samples of simulated data, while further analysis and reconstruction of B mesons using `basf2` is discussed in Chapter 3.

A MC analysis begins with generation of particles, and continues through detector simulation to reconstruction and final analysis. The first two steps, generation and simulation, are performed together to create .mDST (mini Data Summary Table) files containing track, vertex, PID, and detector cluster information. Particle decays are generated using a combination of decay generators. EvtGen [52] - developed at CLEO and BABAR - is used to generate exclusive final states, while PYTHIA [53] is used for remaining inclusive decays that cannot be modelled as a combination of exclusive channels. The generated decays are supplied to Geant4 [54] and the specialised `basf2` software for propagation inside the detector volume.

Simulation of beam-induced background is performed separately from the generation and simulation of particle events. Transport of electrons and positrons in the accelerator is simulated, and any particles leaving the nominal beam trajec-

tory are recorded so that shower events can be calculated. This data is supplied to the Geant4 simulated reconstruction step so that additional particle effects can be added.

For generated MC events, generator-level information about which particles are present is stored alongside the simulated track hits and clusters in the .mDST files. While .mDST files can be used by analysts directly, it's preferred for .mDST files to be reduced to .uDST (micro DST) files containing only events relevant to analysts in order to reduce the number of times that large data files are accessed. The skimming process will be discussed further in Chapters 3 and 4.

Monte Carlo data samples are regularly generated for the Belle II collaboration using distributed GRID computing. Each individual Monte Carlo 'campaign' includes generation of $B\bar{B}$, continuum, and individual signal mode samples using the most recent stable software version and simulated detector configuration.

2.6.1 Monte Carlo samples

The Monte Carlo campaigns used herein are shown in Table 2.1.

Table 2.1: Monte Carlo datasets used in this thesis.

	Software version	Production period
MC5	release-00-05-03	August 2015 - January 2016
MC6	release-00-07-00	June-July 2016
MC7	release-00-07-02	Nov 2016 - Jan 2017
MC8	release-00-08-00	Feb 2017 - June 2017
MC9	release-00-09-00/release-00-09-01	July 2017 - Dec 2017
MC10	release-01-00-02 / release-01-00-03	Jan 2018 - July 2018
MC11	release-02-00-01	August 2018 - January 2019
MC12	release-03-00-00	January 2019

In general, analysts may use any BASF2 software release to study data from any MC campaign, provided the matching magnetic field and beam parameters are applied. The MC version for the data is specified at .mDST load time, and mismatches in the magnetic field can cause significant particle misidentification.

In the context of more complicated analysis software packages, the distinction between release-01 and release-02 is crucial. Release-02 was the first software release to use binary PID, and as a result, PID cut levels (and thus analysis scripts) from before and after this change are incompatible. This thesis will include some studies from before release-02 (MC5, MC6, MC7, MC9) and after release-02 (re-analysis of MC9 and MC10, MC11).

For each MC production, we will be referring to several different types of data samples:

BGx0 Data samples produced with no beam background effects included.

BGx1 Data samples produced with the 'expected' amount of beam background. This includes Touschek, Coulomb, and Radiative Bhabha backgrounds (see Section 2.3). For some analyses, BGx2 or BGx5 samples are also produced.

charged An MC sample of $\Upsilon(4S) \to B^+B^-$ pairs. These B mesons decay generically using PDG branching fractions where available. The remaining parts are then best estimates of the maximum branching fractions of some rare channels, with the remaining difference between the inclusive fraction and the sum of exclusive fractions filled in with hypothetical unmeasured decays.

mixed An MC sample of generic $\Upsilon(4S) \to B^0\bar{B}^0$ pairs that also includes $B^0\bar{B}^0$ mixing.

2.7 Summary

The Belle II experiment continues the legacy of the B Factories further into the luminosity frontier with the new SuperKEKB accelerator expected to achieve $40\times$ the luminosity of Belle using the new Nano-Beam scheme [49]. This increased luminosity comes at the cost of additional luminosity-dependent beam backgrounds, which are addressed as part of the Belle II detector design.

Each of the subdetectors of Belle II is described, as is the way that subdetector information is combined to generate tracks, secondary vertices, and particle ID information. Finally, the process of simulating Belle II data for Monte Carlo studies was outlined and some information about individual Belle II MC datasets provided. In the next chapter, we will build on this knowledge of the Belle II detector and `basf2` software to discuss the techniques used to reconstruct $B\bar{B}$ events for semileptonic and leptonic analyses.

Chapter 3

B meson reconstruction

Having already discussed the hardware of the Belle II detector in Chapter 2, in this Chapter we will further discuss the software used to reconstruct signal events, including those with one or more neutrinos. The hermitic nature of the Belle II detector allows $B\bar{B}$ signal events to be well constrained by reconstructing all visible final state particles. We can consider a signal event as consisting of a signal B meson in our chosen channel of study and an additional "tag" B meson (see Figure 3.1), though it's important to note that due to the asymmetric nature of the e^+e^- collision, for the most part both B mesons and their products will travel in the same direction along the detector and their products may not be spatially separated. For signal decay modes with missing energy due to neutrinos, tagging can be crucial to reduce background from missing or incorrect final state particles in the signal reconstruction. Studies of rare modes in particular benefit from a larger selection of possible tag modes to increase the number of signal candidates retained.
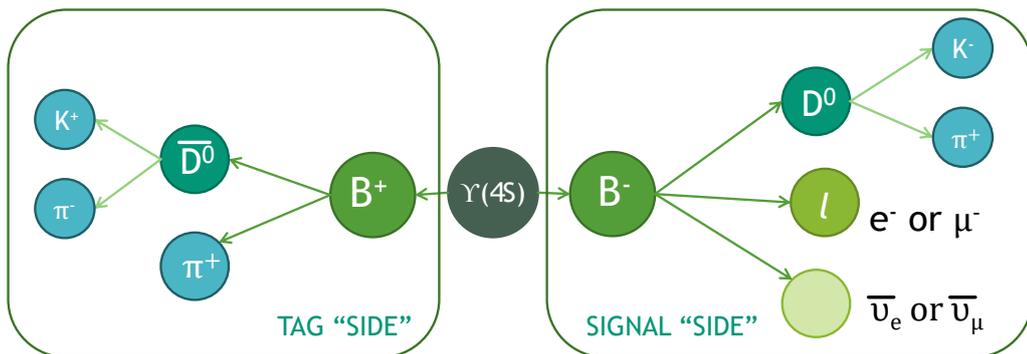
Figure 3.1: Schematic diagram of a possible $B \to D\ell\nu$ event. In this example, the tag B is reconstructed in a hadronic mode.

This Chapter will consider reconstruction of both signal and tag B mesons. For reconstruction of signal, we consider the tools available in the `basf2` software for assembling final state particles into B mesons, and discuss key analysis variables to ensure reconstruction quality. Lastly, we will discuss the reconstruction of tag

modes using specialised software.

3.1   Reconstructing exclusive modes with `basf2`

The main `basf2` software is designed to be simple enough to use at the analysis level [51]. Steering scripts are used to specify how particles are combined in a way transparent to the end user, with the implementation of constructing particles from detector tracks kept in self-contained modules. In this thesis, we focus on reconstruction of exclusive modes, primarily the $B \to D\ell\nu$ and $B \to D\tau\nu$ modes used in calculation of $R(D^{(*)})$. The process of reconstructing a $B \to D^*\ell\nu$ will be discussed step by step.

Firstly, final state particles from the event must be collected. The Belle II detector is able to collect $e^\pm$, $\mu^\pm$, $\pi^\pm$, $K^\pm$, proton, antiproton, γ, and K_L^0 particle candidates directly, while π^0 candidates can be reconstructed from photon pairs using dedicated software. In practice, analysts are able to access candidates for any of these particle channels by using Standard Lists with default selections for particle ID (PID) or momentum requirements.

Consider reconstructing the D^* meson, say in a $D^* \to D\pi$ mode with $D \to K\pi\pi^0$. We will use the charged pion, charged kaon, and neutral pion default lists. We note that the same physical track may be present in both the pion and kaon lists at this stage if the particle ID is sufficiently ambiguous and the reconstruction sufficiently loose. However, when we begin combining particles into intermediate particle candidates, the `basf2` software will ensure that the same track is not used more than once in a single candidate.

Vertex fitting can be used to eliminate many incorrectly reconstructed candidates by constraining the mass and momentum of the reconstructed particle. This reconstruction is a mixture of both finding vertex candidates and fitting parameters to the interaction vertex based on the daughter particle information. The vertex position, and the particle mass and invariant momentum can be calculated, and the daughter particle momenta can also be corrected to account for the mother particle information.

At Belle II, these fits can be performed using one of several possible fitting packages: KFit (a port of the Belle vertex fitter KFitter), RAVE [55], or TreeFitter [56]. A kinematic fit (eg KFit) uses the known properties of the user-specified decay chain to improve the fit, and these kinematic constraints from the decay chain are applied using Lagrange multiplier methods. In contrast, RAVE uses a process of weighting and reweighting possible vertex fit properties. The weight can be interpreted as a probability of assigning a track to a particular vertex, and this is iterated until a minimum weighted least squares fit is found. Lastly, TreeFitter is

designed to fit the entire decay chain simultaneously rather than requiring multiple intermediate fit steps.

Returning to our example, candidate D mesons are reconstructed by combining K, π, and π^0 candidates. The 4-momentum of the D meson is calculated from the 4-momentum of the candidate daughter particles, and then low quality D meson candidates can be eliminated via selections on the D invariant mass. Similarly the D^* meson is reconstructed by combining D meson candidates (which have been re-fitted to have the nominal D mass) with an additional low momentum (or "slow") pion. In practice, as reconstruction of a D^* meson requires two fitting steps, the second with only a low-momentum particle, selection criteria on the mass difference $m(D^*) - m(D)$ are much more effective as constraints than selection criteria on $m(D^*)$ itself.

Of course, not all combinations of reconstructed K, π, and π^0 mesons truly come from a D or D^* decay, or from a D^* meson decaying in our chosen mode. The incorrect D^* meson candidates where the final state particles are correctly identified but assigned to the D^* decay tree incorrectly (for example by switching the direct pion from the D^* meson with one from the D meson decay), or where correct final state particles from two B mesons in the event are mixed to form a single candidate, are called combinatorial background. This category can also include real D or D^* mesons from a $c\bar{c}$ event mistaken for a $B\overline{B}$ event, though this is often referred to as continuum background.

The other major type of background in reconstruction is physics background, where incorrect particle ID or final state particles that aren't reconstructed in the detector can cause another B decay mode to mimic our signal decay of interest. For this reason, analysts will often run their reconstruction scripts on samples of suspected physics background modes, to check how frequently these decays occur. Unfortunately, in the context of rare decay studies not all sources of background will be well known. At Belle and BABAR, studies of $R(D^{(*)})$ and exclusive V_{cb} included un-modelled background from additional excited D meson modes (referred to collectively as D^{**}). At Belle II, branching ratios of D^{**} modes are intended to be measured directly in order to further constrain their contribution to desired signal measurements.

While selections on particle masses and momenta during the reconstruction process can be effective at eliminating some combinatorial background, specialist analysis variables taking into account the whole event can be especially effective. In the next Section, we discuss variables useful to $B \to D^{(*)}\ell\nu$ and $B \to D^{(*)}\tau\nu$ analyses.

3.2 *B* meson analysis variables

Aside from the reconstructed mass and momentum of the B meson, there are a wide variety of other variables that can be used in analysis. The beam constraints from the e^+e^- collision can be used with a single B meson to extract additional information, especially when the specific signal mode of interest is known. By including a tag B meson as well, the full event can be further constrained and signal events may be separated from continuum background.

For Monte Carlo samples, we also have access to variables that assess the correctness of reconstruction. The definitions of these variables and their relevance to reconstruction analyses will also be discussed in this Section.

3.2.1 Missing energy analysis variables

In the context of the $R(D^{(*)})$ analysis, $B \to D\tau\nu$ and $B \to D\ell\nu$ decays need to be reconstructed but also differentiated from each other, especially when we consider only the $\tau \to \ell\nu$ decay channels. The decay geometry of these two modes is compared in Figure 3.2.
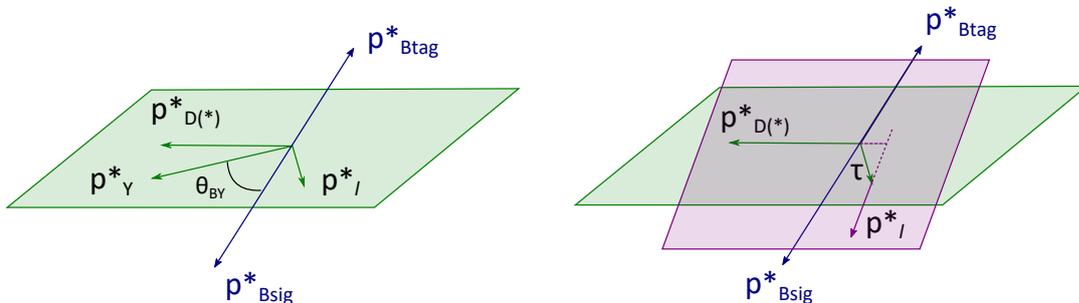
Figure 3.2: Decay geometry of semileptonic (left) and semitauonic (right) decay modes. Only visible final state particles are shown. In a semileptonic decay, the $D^{(*)}$ and lepton momentum may be combined (marked as Y), and the angle between this combined momentum and the expected $\mathbf{p}^*_{\text{sig}}$ is a real angle. In the semitauonic decay, a Y reconstructed from the $D^{(*)}$ and lepton lies on a different plane and has a different vertex to the true B.

As neutrinos cannot be directly reconstructed, the same final state particles are present in both $B \to D\tau\nu$ and $B \to D\ell\nu$ decays, although the momentum of the lepton in the centre of mass frame ($\mathbf{p}^*_\ell$) has a different distribution in both cases.

Instead of considering the lepton alone, a key variable for distinguishing well-reconstructed semileptonic decays is $\cos\theta_{BY}$, or the cosine of the angle between the B meson and the combined $D^{(*)}\ell$ system. The angle is labelled θ_{BY} in Figure 3.2.

In the reconstruction, this angle is defined

$$\cos\theta_{BY} = \frac{2E^*_{\text{beam}}E^*_{D^{(*)}\ell} - m^2_B - m^2_{D^{(*)}\ell}}{2|\mathbf{p}^*_B||\mathbf{p}^*_{D^{(*)}\ell}|} \tag{3.1}$$

where * is used to denote quantities in the centre-of-mass frame, and $\mathbf{p}$ is the 3-momentum of a system. E^*_{beam} is half the energy of the system in the CMS frame.

Signal $B \to D^{(*)}\ell\nu$ decays have $\cos\theta_{BY}$ between -1 and 1 in almost all cases, with those outside the range due to mis-reconstruction or detector resolution effects. As the $B \to D\tau\nu$ decays have different kinematics, these are reconstructed with larger negative values of $\cos\theta_{BY}$.

For decays with a single neutrino, after the reconstruction of the visible particles from both B mesons in the event, the 4-momentum 'missing' from the e^+e^- collision can be assigned directly to this neutrino. The invariant mass of this 4-vector can be considered as the amount of missing mass in the reconstruction, expected to peak at 0 if there was only one neutrino in the event. At Belle II, this missing mass squared (M^2_{missing}) is defined

$$p_{\nu_\ell} = p_{e^+} + p_{e^-} - p_{B_{\text{tag}}} - p_{D^{(*)}} - p_\ell, \tag{3.2}$$

$$M^2_{\text{missing}} = |p_\nu|^2. \tag{3.3}$$

We notice that this definition depends on the tag reconstruction, and also on the reconstruction of the $B \to D^{(*)}\ell\nu$ signal mode. In more general circumstances when the precise decay channel of the signal B is unknown, we can write

$$M^2_{\text{missing}} = \left(E^*_{\text{beam}} - E^*_{\text{signal}}\right)^2 - |\mathbf{p}^*_{\text{tag}} + \mathbf{p}^*_{\text{signal}}|^2 \tag{3.4}$$

to calculate the missing mass.

For modes containing leptonic tau decays and thus multiple neutrinos, the missing mass is often nonzero as the missing 4-vector is a sum of particles from different vertices rather than a single particle.

Variables that use detector-level information more directly are often also useful in separating signal reconstructions from combinatorial background. The label E_{extra} is used to refer to the sum of extra energy deposits left in the calorimeter after both B in the event have been reconstructed. At Belle or BABAR, the distribution of the sum across events is expected to peak at 0 if all final state particles in the event have been included into one of the B mesons. That is, no neutral particles such as π^0 have left deposits in the calorimeter without being reconstructed. Unfortunately the background at Belle II is much larger, and significant suppression of beam back-

ground photons is required to achieve a peak at 0 and reduce the smearing of the distribution.

3.2.2 Other significant variables

The most commonly used variables for B meson reconstruction are the beam-constrained mass (M_{bc}) and the energy difference between the reconstructed B meson and the expectation from the beam energy (ΔE). These variables are the most powerful for hadronic modes where all of the energy and momentum of the B meson is expected to be reconstructed, such as for B_{tag} modes, but they do have some applicability even in other situations. The key advantage of M_{bc} and ΔE as a pair is that they are mostly uncorrelated.

In the centre of mass frame, ΔE is defined by

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

and is expected to be zero for correctly-reconstructed B decays in channels without neutrinos.

The beam constrained mass variable was developed at CLEO, and is defined as follows:

$$M_{bc} = \sqrt{E_{\text{beam}}^{*2} - p_B^{*2}} \tag{3.6}$$

Notably in a symmetric collider, this formulation of M_{bc} is independent of the mass hypothesis for the particles used to reconstruct the B. In an asymmetric collider, some dependence on the mass hypothesis is introduced in the conversion of particles to the centre of mass frame. At BABAR, independence from the mass hypothesis was re-introduced by adjusting the formulation of the constrained mass, but at Belle M_{bc} was considered independent enough for most analyses.

An additional variable used to separate signal from background is a ratio of Fox-Wolfram moments. In practice, many variables are used in continuum suppression, but this particular variable specifically separates $B\overline{B}$ events from continuum background events by parameterising their different shapes. Since the $B\overline{B}$ threshold is just below the $\Upsilon(4S)$ resonance, $B\overline{B}$ pairs are produced with very little momentum relative to the centre of mass while other $q\overline{q}$ events have more momentum and a more jet-like structure.

The most common implementation of Fox Wolfram moments is the ratio R2, defined

$$R2 = H_2/H_0 \tag{3.7}$$

for the Fox Wolfram moments

$$H_l = \sum_{i,j} \frac{|p_i||p_j|}{E_{\text{vis}}^2} P_l(\cos\theta_{ij}) \tag{3.8}$$

where $P_l(x)$ are the Legendre polynomials, p are the particle momenta, θ_{ij} is the angle between the particle momenta, and the sum over i,j represents all pairs of particles including autocorrelations. In `basf2` software, R2 is implemented as `R2EventLevel` and constructed from all ECL clusters and tracks assigned their most likely particle reconstruction hypothesis. For tracks, this uses the largest global PID value to set the mass hypothesis in the track fit. The `R2EventLevel` has low values for more isotropic events and high numbers for more jet-like events.

3.2.3 Truth and reconstruction quality

For all analyses, it is desirable to have an indication of correctness of reconstruction. On MC samples, generator-level information about which particles were created can be compared to the particles reconstructed from tracks and clusters. In `basf2` this is performed by the MCMatcherParticles module, with a function `matchMCTruth()` that can match reconstructed particles to their generator counterparts. In this comparison process, a series of error bits are returned to describe different kinds of incorrect reconstructions. An expectation that there be no error bits between the MC B meson and final reconstructed B meson turns out to be very strict, and is not commonly used in analysis. Instead, the state of multiple error bits are combined to produce various flags indicating the quality of reconstruction:

isSignal A reconstructed particle is marked with isSignal if all its daughter final state particles are reconstructed with the correct mass hypotheses, and if the daughter particles are assigned to their mothers correctly. Missing final state radiation photons, other missing generator photons that are not direct parts of a decay process, and missing resonances are permitted in this status.

isSignalAcceptMissingNeutrino As above, but we also ignore the flag for generated neutrinos that have not been reconstructed.

isExtendedSignal As for isSignal, but a misidentification of a final state particle is allowed (eg. a pion reconstructed as a kaon). Combining the particles incorrectly is also allowed, for example switching pions from D or K decays in a $D \rightarrow K(\rightarrow \pi\pi)\pi$ process.

In most B_{sig} analyses, `isSignal` and `isSignalAcceptMissingNeutrino` are used as the indicator of reconstruction quality for individual particles. Quantitative ways

of describing reconstruction quality on different samples of multiple events are also required. Definitions for terms used in this thesis are presented below:

Purity

Purity is defined

$$\text{purity} = \frac{\text{number of correct objects}}{\text{number of objects collected}} \tag{3.9}$$

where "correctness" of an object often implies it has `isSignal==1`.

Retention

Retention is often a very mechanical measure of reconstruction. The retention is defined

$$\text{retention} = \frac{\text{number of events or candidates surviving a process}}{\text{number of events or candidates submitted to the process}} \tag{3.10}$$

Typically, the "correctness" of the candidates or objects in question is not considered for retention. It's typical for this definition to also be called "tagging efficiency" or just "efficiency" in literature or in discussion, but in this thesis we will be careful to always use "retention" in this case to avoid confusion.

Efficiency

There are many definitions of efficiency, which can be considered to lie on somewhat of a spectrum between the definition for purity and the definition for retention. Due to the large number of possible definitions, we will always be clear in the text of this thesis which efficiency is being considered in each context. In different contexts, efficiency may be defined:

- $$\text{efficiency} = \frac{\text{objects collected}}{\text{total number of input events (of any type)}}$$

  (this will always be called "retention" in this thesis)

- $$\text{efficiency} = \frac{\text{correct objects collected}}{\text{total number of input events (of any type)}}$$

  (this is will be the definition of efficiency most commonly used in this thesis)

- $$\text{efficiency} = \frac{\text{correct objects collected}}{\text{total number of input events that contained those objects}}$$

●

$$\text{efficiency} = \frac{\text{objects collected}}{\text{total number of input events that contained those objects}}$$

These last two measures require additional knowledge of the branching ratios of the process(es) we are trying to collect, or the frequency at which those processes are generated in our Monte Carlo samples. Measures of efficiency across multiple candidates as well as multiple events are possible, but in general these will not be used in this thesis as the interpretation of the numerical result tends to be difficult. Retention or purity measures including multiple candidates are almost always misleading to the casual observer, and will not be used.

For purity calculations and some formulations of efficiency, both the numerator and denominator values are collected from measurement rather than being fixed by the size of the sample. This makes the calculation of error in efficiency and purity measurements more complicated than the Poisson statistics for a retention measurement. The error in purity and efficiency is especially relevant in comparisons across different data samples. A formula for the error in purity is presented in Appendix A. Comparison of multiple data samples will be discussed in Chapter 4.

3.3 Reconstructing tag B mesons

Tag B meson modes are used to constrain events and thus improve the reconstruction quality of signal modes with missing neutrinos. If more B_{tag} modes can be reconstructed, we have more opportunities to match these tags with signal decays. In particular, the efficiency of tag reconstruction is important, combining how well tag modes are reconstructed with the actual quantity of tags that can be reconstructed based on the mode branching ratio and analysis retention. In general, the high retention aspects of tag reconstruction are more important than tag purity, as this gives us access to more signal events.

The Full Event Interpretation (FEI) [57] is Belle II's solution to reconstructing and classifying very large numbers of exclusive tag modes. This software is designed to reconstruct tracks and clusters into intermediate particles and B mesons, while providing an output value to allow the best tag channel candidate to be selected.

3.3.1 Full Event Interpretation

Multivariate classifiers are a form of machine learning that allows input information to be classified into one of multiple output categories, after a training stage in which the correct output category matching the input information is known.

These output categories may be made distinct by requiring that input is sorted into exactly one category, or a continuous numerical output may be supplied so that the category division may be applied later by the user. Continuous output can also be converted to a probability value. For the FEI, the training process is performed on Monte Carlo samples so that MC generator information can be used to check if B_{tag} candidates are correctly or incorrectly reconstructed. The `isSignal` or `isSignalAcceptMissingNeutrino` flag is used in training to identify correctly reconstructed candidates or "signal" events. Incorrectly reconstructed candidates are often referred to as (combinatorial) background. The output classifier value supplied to the user is known as the signal probability. Once trained, the FEI can be applied to data samples, where the number of correctly reconstructed B mesons can be estimated using a fit to the M_{bc} distribution.

The FEI is a hierarchical network of multivariate classifiers, shown in Figure 3.3. In the modern `basf2` software, this FEI structure is implemented using Boosted Decision Trees (BDT), specifically the FastBDT implementation developed specifically for the FEI [58].
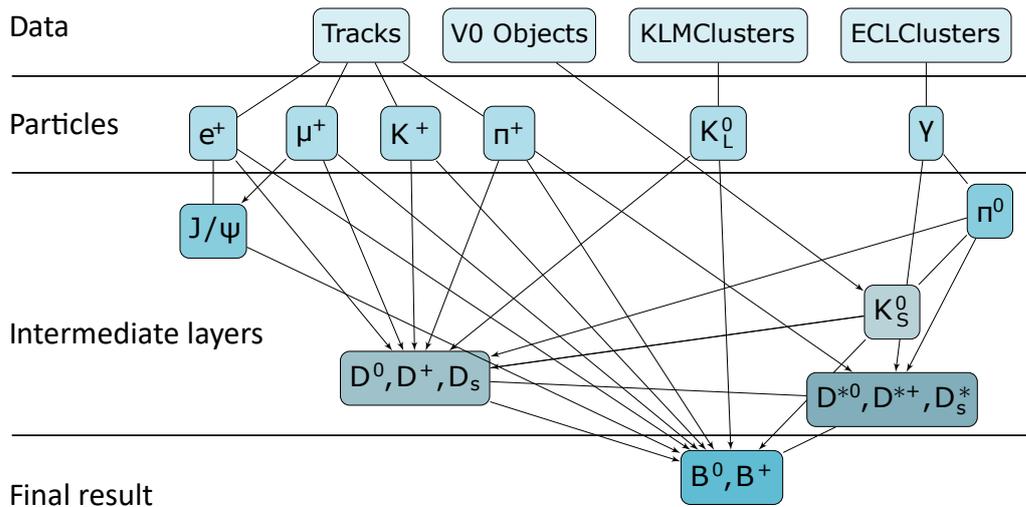
Figure 3.3: Diagram of the FEI classifier network structure. Multiple layers of intermediate classifiers are used. Adapted from [57].

Tracks and clusters are reconstructed into final state particle candidates, which are assigned a classifier value, and extremely low classifier value candidates are eliminated. Similarly, final state particles are combined into intermediate particles, again with selections on the classifier value. Finally, B mesons are assembled, suitable for analysts to use. Each intermediate particle may be reconstructed in multiple different channels. Through combinations of B, D, and other intermediate channels, the FEI has $\mathcal{O}(10,000)$ different decay chains. The default decay channels for the B are presented in Table 3.1.

The key benefit of a classifier system is that reconstructions in different channels

Table 3.1: Default B meson channels in the FEI, collected from the `basf2` software implementation. Any D_{SL} represent D mesons with semi-leptonic decay channels.

Hadronic B^+ channels	Hadronic B^0 channels	Semileptonic B^+ channels	Semileptonic B^0 channels
$\bar{D}^0\pi^+$	$D^-\pi^+$	$\bar{D}^0 e^+$	$D^- e^+$
$\bar{D}^0\pi^+\pi^0$	$D^-\pi^+\pi^0$	$\bar{D}^0\mu^+$	$D^-\mu^+$
$\bar{D}^0\pi^+\pi^0\pi^0$	$D^-\pi^+\pi^0\pi^0$	$\bar{D}^{*0} e^+$	$D^{*-} e^+$
$\bar{D}^0\pi^+\pi^+\pi^-$	$D^-\pi^+\pi^+\pi^-$	$\bar{D}^{*0}\mu^+$	$D^{*-}\mu^+$
$\bar{D}^0\pi^+\pi^+\pi^-\pi^0$	$D^-\pi^+\pi^+\pi^-\pi^0$	$D^-\pi^+ e^+$	$\bar{D}^0\pi^- e^+$
$\bar{D}^0 D^+$	$\bar{D}^0\pi^+\pi^-$	$D^-\pi^+\mu^+$	$\bar{D}^0\pi^-\mu^+$
$\bar{D}^0 D^+ K_S^0$	$D^- D^0 K^+$	$D^{*-}\pi^+ e^+$	$\bar{D}^{*0}\pi^- e^+$
$\bar{D}^{*0} D^+ K_S^0$	$D^- D^{*0} K^+$	$D^{*-}\pi^+\mu^+$	$\bar{D}^{*0}\pi^-\mu^+$
$\bar{D}^0 D^{*+} K_S^0$	$D^{*-} D^0 K^+$		
$\bar{D}^{*0} D^{*+} K_S^0$	$D^{*-} D^{*0} K^+$	$\bar{D}^0_{SL}\pi^+$	$D^-_{SL}\pi^+$
$\bar{D}^0 D^0 K^+$	$D^- D^+ K_S^0$	$\bar{D}^0_{SL}\pi^+\pi^0$	$D^-_{SL}\pi^+\pi^0$
$\bar{D}^{*0} D^0 K^+$	$D^{*-} D^+ K_S^0$	$\bar{D}^0_{SL}\pi^+\pi^0\pi^0$	$D^-_{SL}\pi^+\pi^0\pi^0$
$\bar{D}^0 D^{*0} K^+$	$D^- D^{*+} K_S^0$	$\bar{D}^0_{SL}\pi^+\pi^+\pi^-$	$D^-_{SL}\pi^+\pi^+\pi^-$
$\bar{D}^{*0} D^{*0} K^+$	$D^{*-} D^{*+} K_S^0$	$\bar{D}^0_{SL}\pi^+\pi^+\pi^-\pi^0$	$D^-_{SL}\pi^+\pi^+\pi^-\pi^0$
$D_s^+ \bar{D}^0$	$D_s^+ D^-$	$\bar{D}^0_{SL} D^+$	$\bar{D}^0_{SL}\pi^+\pi^-$
$\bar{D}^{*0}\pi^+$	$D^{*-}\pi^+$	$\bar{D}^0_{SL} D^+ K_S^0$	$D^-_{SL} D^0 K^+$
$\bar{D}^{*0}\pi^+\pi^0$	$D^{*-}\pi^+\pi^0$	$\bar{D}^{*0}_{SL} D^+ K_S^0$	$D^-_{SL} D^{*0} K^+$
$\bar{D}^{*0}\pi^+\pi^0\pi^0$	$D^{*-}\pi^+\pi^0\pi^0$	$\bar{D}^0_{SL} D^{*+} K_S^0$	$D^{*-}_{SL} D^0 K^+$
$\bar{D}^{*0}\pi^+\pi^+\pi^-$	$D^{*-}\pi^+\pi^+\pi^-$	$\bar{D}^{*0}_{SL} D^{*+} K_S^0$	$D^{*-}_{SL} D^{*0} K^+$
$\bar{D}^{*0}\pi^+\pi^+\pi^-\pi^0$	$D^{*-}\pi^+\pi^+\pi^-\pi^0$	$\bar{D}^0_{SL} D^0 K^+$	$D^-_{SL} D^+ K_S^0$
$D_s^{*+} \bar{D}^0$	$D_s^{*+} D^-$	$\bar{D}^{*0}_{SL} D^0 K^+$	$D^{*-}_{SL} D^+ K_S^0$
$D_s^+ \bar{D}^{*0}$	$D_s^+ D^{*-}$	$\bar{D}^0_{SL} D^{*0} K^+$	$D^-_{SL} D^{*+} K_S^0$
$\bar{D}^0 K^+$	$D_s^{*+} D^{*-}$	$\bar{D}^{*0}_{SL} D^{*0} K^+$	$D^{*-}_{SL} D^{*+} K_S^0$
$D^-\pi^+\pi^+$	$J/\psi K_S^0$	$\bar{D}^0 D^+_{SL}$	$D^- D0_{SL} K^+$
$D^-\pi^+\pi^+\pi^0$	$J/\psi K^+\pi^-$	$\bar{D}^0 D^+_{SL} K_S^0$	$D^- D^{*0}_{SL} K^+$
$J/\psi K^+$	$J/\psi K_S^0\pi^+\pi^-$	$\bar{D}^{*0} D^+_{SL} K_S^0$	$D^{*-} D^0_{SL} K^+$
$J/\psi K^+\pi^+\pi^-$		$\bar{D}^0 D^{*+}_{SL} K_S^0$	$D^{*-} D^{*0}_{SL} K^+$
$J/\psi K^+\pi^0$		$\bar{D}^{*0} D^{*+}_{SL} K_S^0$	$D^- D^+_{SL} K_S^0$
$J/\psi K_S^0\pi^+$		$\bar{D}^0 D^0_{SL} K^+$	$D^{*-} D^+_{SL} K_S^0$
		$\bar{D}^{*0} D^0_{SL} K^+$	$D^- D^{*+}_{SL} K_S^0$
		$\bar{D}^0 D^{*0}_{SL} K^+$	$D^{*-} D^{*+}_{SL} K_S^0$
		$\bar{D}^{*0} D^{*0}_{SL} K^+$	$D_s^+ D^-_{SL}$
		$D_s^+ \bar{D}^0_{SL}$	$D^{*-}_{SL}\pi^+$
		$\bar{D}^{*0}_{SL}\pi^+$	$D^{*-}_{SL}\pi^+\pi^0$
		$\bar{D}^{*0}_{SL}\pi^+\pi^0$	$D^{*-}_{SL}\pi^+\pi^0\pi^0$
		$\bar{D}^{*0}_{SL}\pi^+\pi^0\pi^0$	$D^{*-}_{SL}\pi^+\pi^+\pi^-$
		$\bar{D}^{*0}_{SL}\pi^+\pi^+\pi^-$	$D^{*-}_{SL}\pi^+\pi^+\pi^-\pi^0$
		$\bar{D}^{*0}_{SL}\pi^+\pi^+\pi^-\pi^0$	$D_s^{*+} D^-_{SL}$
		$D_s^{*+} \bar{D}^0_{SL}$	$D_s^+ D^{*-}_{SL}$
		$D_s^+ \bar{D}^{*0}_{SL}$	$D_s^{*+} D^{*-}_{SL}$
		$\bar{D}^0_{SL} K^+$	
		$D^-_{SL}\pi^+\pi^+$	
		$D^-_{SL}\pi^+\pi^+\pi^0$	

can be compared using the classifier output value. For the B meson, this final signal probability value combines the daughter particle signal probabilities with extra calculated information such as kinematic and vertex fit information.

Due the large number of possible candidate reconstructions with the FEI, systems are required to eliminate background candidates early before further memory-heavy layers of the classifier are applied. For example, loose selections are applied to the invariant mass of most intermediate particles, and only the best few candidates are considered in the next reconstruction stage. This number of considered particle candidates is usually between 10 and 20. Additionally, a faster version of reconstruction vertexing is used in the FEI compared to the usual methods in `basf2`. This FastFit algorithm was also developed specifically for the FEI [59].

Further information about the specific implementation of the FEI and its training variables can be found in [57, 60, 61]

3.3.2 FEI training at Belle II

The FEI training can be performed in different ways to focus on the tag reconstruction individually, or take into account information from the whole event including the signal mode. When the signal mode reconstruction information is used in the training for FEI reconstruction of tags this is referred to as a specific FEI, and the restriction that no final state particles should remain in the event after the tag reconstruction can be used in training the FEI signal probability. In contrast, the FEI trained on the single tag B meson without signal information is referred to as generic FEI. Generic trainings of the FEI can be performed centrally for all analysis users, and the performance of these generic trainings will be the focus of the next Chapter of this thesis.

Regardless of specific or generic training method, the default FEI configuration is able to train up to six different B channel sets simultaneously. These are:

- `B0:generic`, the list of strictly hadronic B^0 candidates. No channels contain K_L particles.

- `B+:generic`, the list of strictly hadronic B^+ candidates. No channels contain K_L particles.

- `B0:semileptonic`, the list of semileptonic B^0 candidates. Hadronic B^0 decays that include a semileptonic D are also included in this list. No channels contain K_L particles.

- `B+:semileptonic`, the list of semileptonic B^+ candidates. Hadronic B^+ decays that include a semileptonic D are also included in this list. No channels contain K_L particles.

- `B0:KL`, the list of hadronic B^0 candidates which include K_L^0 particles.

- `B+:KL`, the list of hadronic B^+ candidates which include K_L^0 particles.

When generating each of the FEI lists, the charge conjugate list is also created. In general, the `B0:generic` and `B+:generic` lists are referred to together as the hadronic FEI, and similarly the semileptonic lists can be grouped. The FEI trainings including K_L^0 particles are newer and relatively under-studied at this stage.

3.4 Summary

In this Chapter, we have summarised the reconstruction of B mesons in the `basf2` software. Most exclusive decay modes are reconstructed by analysis users, and specific fitting software can be selected to process combinations of tracks into intermediate particles. Common variables used to separate desired signal modes from background are discussed, and methods of quantifying reconstruction quality such as purity, efficiency, and retention variables are explained.

For reconstructing tag modes, we wish to use Belle II's FEI software. A brief overview of the FEI and its training is provided, including an overview of the FEI's default channel lists. The discussion of the FEI performance will continue in the next Chapter, where FEI results are compared across different Belle II software conditions.

Chapter 4

Measuring FEI performance

The FEI is expected to play a large role in studies of rare decays at Belle II. By collecting a large number of exclusive channels, we will be able to constrain the energy of more rare decay events. Before the FEI can be used for these studies, however, significant benchmarking and validation is required.

In this Chapter, we discuss tests of the FEI from software user perspectives as well as a physics perspective. We begin with an overview of generic FEI performance with early `basf2` software releases and early Monte Carlo samples. A comparitive overview of the generic FEI performance in MC10 and MC11 using recent 2018 FEI trainings is also provided.

FEI performance is also considered in the context of skimming and analysis. Full .mDST data files must be reduced to the smaller .uDST format to streamline access by multiple analysts. Skims containing events with FEI tags are needed for rare decay analyses, and the development of these skims will be discussed in detail. Early studies of the FEI hadronic tag combined with a semitauonic signal selection are also presented.

4.1   FEI overview with early software

In this Section, we focus on testing and explanations of performance for early generic FEI trainings. The plots shown focus on the performance of the FEI in MC6 and MC7, although trainings from MC5 were also tested. The key uniting feature of each of these early FEI trainings and MC samples is that they were generated and processed with `basf2` releases that used binary PID (see Section 2.5.4). Later in this thesis in Section 4.3 we will look at the efficiency, purity, and retention of more recent FEI trainings.

As briefly discussed in Section 3.3.1, the FEI output includes multiple candidates per event, which can be differentiated by their decay channel (using the decayID or `dmID` variable) and also by their FEI signal probability output (`sigProb`). In the
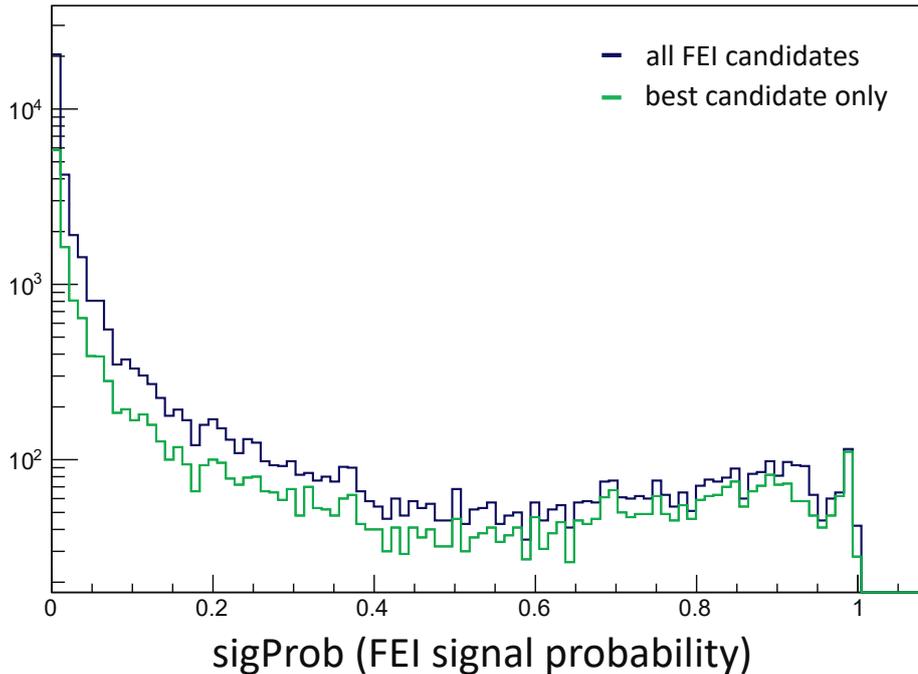
Figure 4.1: Signal probability distribution for all candidates (navy) and best candidates (green) using the B^0 Hadronic FEI trained for MC5 BGx0

context of an analysis, we are often interested more in the number of events that are captured by the FEI than by each of the multiple candidates. For this reason, unless otherwise stated, all plots in this thesis will employ some kind of best candidate selection: typically, we will use the candidate with the highest FEI signal probability in the event. Without a best candidate selection, we give more visual weight to the influence of events that contain multiple candidates—often similar candidates with interchanges of individual pions—and thus favour these combinatorial background effects. Basic cuts must also be applied to ensure a minimum quality level for the B_{tag} candidates. These cuts are

- `sigProb` > 0.001

- $|\Delta E| < 0.2$

- $M_{bc} > 5.24$

and these will be applied to all MC samples studied before the best candidate selection is performed.

A comparison of the `sigProb` distribution with one candidate per event or with multiple per event is shown in Figure 4.1.

With one candidate per event, we are now able to create histograms that show us the distribution of physics properties of the B_{tag} across the sample of events collected. When the FEI is used to reconstruct B^0 tag candidates on a $B^0\bar{B}^0$ sample, we expect the M_{bc} distribution to peak at the B^0 mass, with an ARGUS/Crystal

ball distribution for combinatorial background. In Figure 4.2, we plot the beam constrained mass of the B^0 reconstructed from FEI tags. The FEI decay channel information is also included.

The top plot in the Figure allows us to see that the $D^-\pi^+\pi^+\pi^-\pi^0$ channel (yellow) is considered the best candidate in the largest number of events. However, we can also see that the 'height' of the yellow band is very consistent in size at all values of M_{bc}, indicating that this channel might consist mostly of combinatorial background. In contrast, the $D^-\pi^+$ channel (red) and the $D^-\pi^+\pi^0$ channel (blue) are clearly peaked at the B^0 mass. For more of an overview, grouping the channels into categories (see bottom part of Figure 4.2) allows us to observe that the contribution of DDK and J/Ψ events to the overall shape of the M_{bc} distribution is small, with most reconstructed modes being D or D^* with pions. This is expected from the branching fractions of these modes.

While the M_{bc} distribution (and other physics distributions) of the FEI are important, especially for comparisons to data, in Monte Carlo studies the retention, efficiency, and purity of FEI tagging is often more valuable to analysts. These measures also provide a more numerical way of benchmarking the FEI performance over time as the software develops.

In these early studies, we track the performance of the FEI relative to several different measures of "correct" reconstruction. For tagging studies where the signal B is also fully reconstructed and MC Matched, `isExtendedSignal` may be a preferred measure of tag quality as it is more inclusive. However, in the context of FEI benchmarking studies where we specifically wish to collect fully correctly-reconstructed FEI modes, `isSignal` is preferred. Figures 4.3 and 4.4 show efficiency vs purity curves for each of our different measures of truth, calculated for the MC7 `B0:Hadronic` FEI. Each curve is collected by making successive cuts on the minimum value of the FEI signal probablity. Further discussion of the calculation of error in these quantities is provided in Appendix A.

In Figure 4.3, we compare the performance of the FEI trained and run on MC6 and MC7 generic samples. We can see that the purity and efficiency have decreased at high values of signalProbability as we change from MC6 to MC7. Since the classifier has been re-trained for the new MC, it's hard to tell if this is an impact of the new MC itself, or the training process. Regardless, the error regions overlap for `isExtendedSignal` and `isSignal` for the sample size tested, indicating that there is no major problem with this change in MC.

In general though, tight selections on the signal probability are not used in analyses. The FEI classifier itself assigns high values of signal probability to events that appear correctly reconstructed, whether this is true or not. Thus when tighter selections are used, the combinatorial background is sculpted to have the same
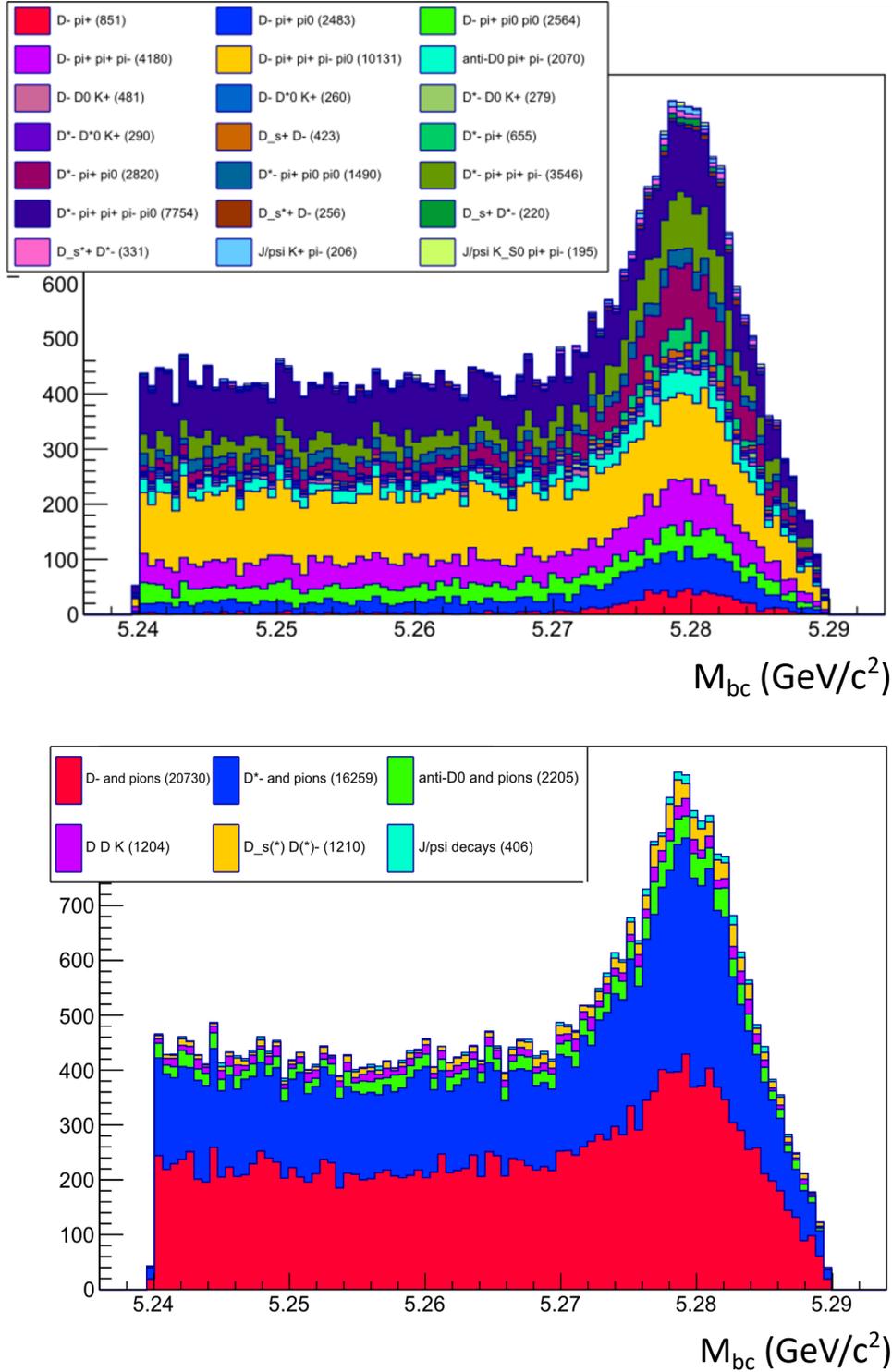
Figure 4.2: M_{bc} distribution for best candidates of B^0 Hadronic FEI, trained for MC6 BGx0, shown with colour blocks indicating individual decay channels (top) or categories of channels (bottom). In the legend, the number in parentheses indicates the number of events with the best candidate in that channel or category. This sample uses the recommended basic cuts $M_{bc} > 5.24, |\Delta E| < 0.2, \text{sigProb} > 0.001$
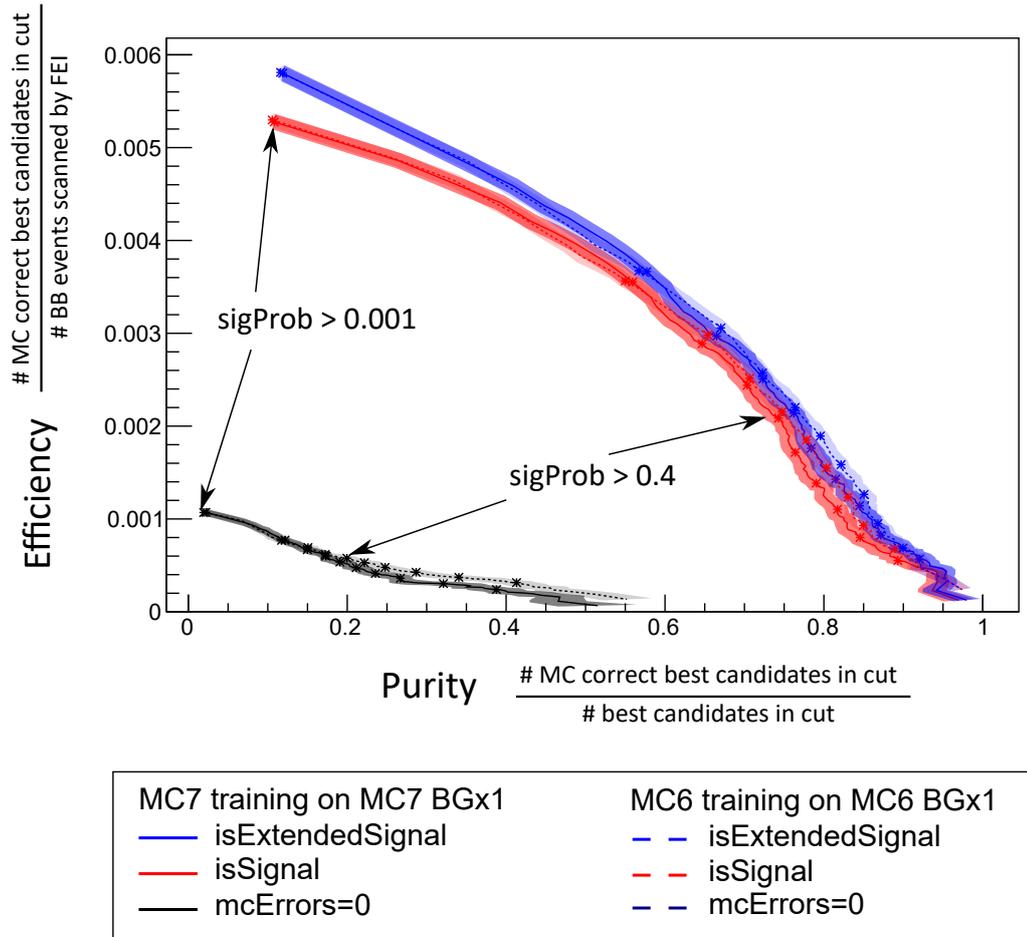
Figure 4.3: Calculated efficiency vs purity curves comparing the FEI performance between the MC7 BGx1 training (solid curves) and the MC6 BGx1 training (dashed curves). Each × marker represents an additional 0.1 increase in the cut on the minimum signal probability.
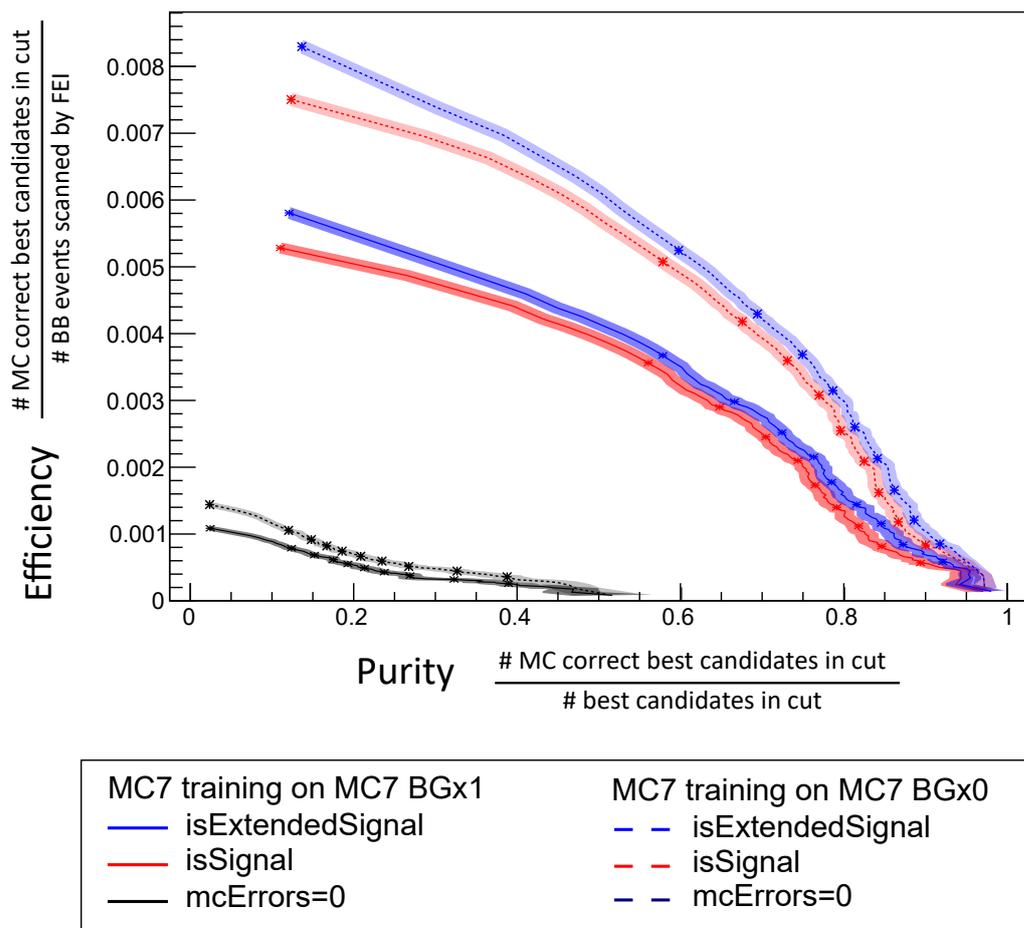
Figure 4.4: Calculated efficiency vs purity curves comparing the FEI performance between MC7 BGx1 training applied to BGx1 MC (solid curves) and applied to BGx0 MC (dashed curves). Each × marker represents an additional 0.1 increase in the cut on the minimum signal probability.

distribution of physics variables as correctly-reconstructed events. This will prevent us from making fits to signal and background on the M_{bc} distribution to calculate the true yield of correct tags.

In Figure 4.4, we compare the MC7-trained FEI performance between MC samples with and without beam background. While the efficiency is lower with beam background, by keeping track of the markers at each step of 0.1 in signalProbability, we observe that the purity at each FEI signal probability cut level is approximately the same with and without background. This is a good indication of the consistency of the FEI classifier values at identifying correctly reconstructed events. We note that in these early studies, the generic FEI was trained on BGx0 data.

For interest, we can also look at the retention vs purity curve of each individual FEI channel, or each channel category. Channel categories for MC6 are shown in Figure 4.5.
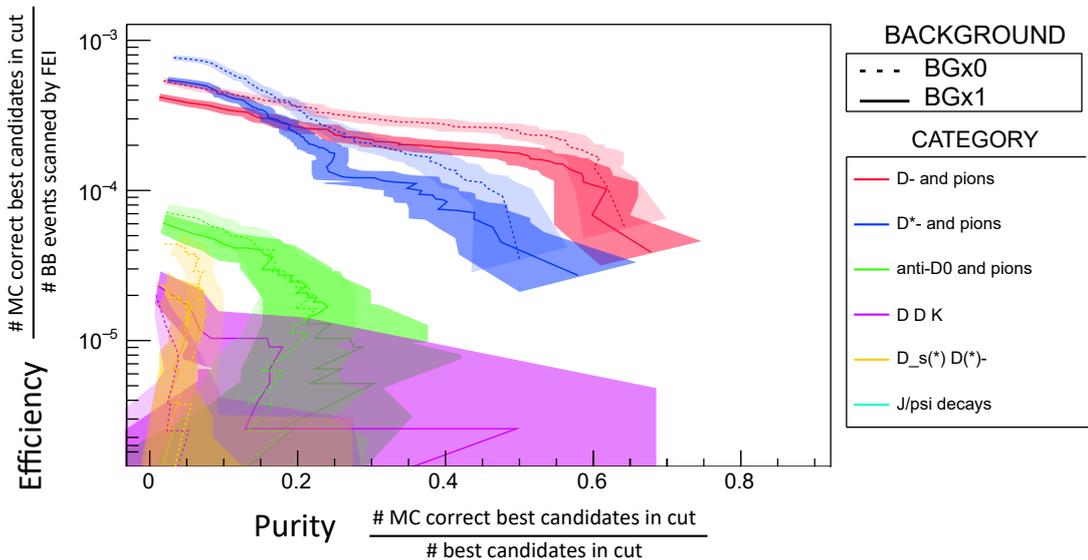
Figure 4.5: Efficiency vs Purity curves, grouping FEI channels together by category. The MC7 FEI is applied to both BGx0 MC (dashed curves) and BGx1 MC (solid curves).

We observe that most of the overall efficiency of the FEI comes from the $D^{(*)}+m\pi$ channels, as we might have expected from the distribution of channel groups in the M_{bc} distribution in Figure 4.2. From this plot, we notice that an expected high-purity group of channels—the J/Ψ decays—are missing, which warrants further investigation. This issue is discussed in detail in the next Section.

4.2 Channel effects in signal probability

A multivariate classifier system like the FEI is somewhat of a black box, and determining whether a change in output is due to changes in training or is instead a

result of changes to other parts of the Belle II software ecosystem requires ongoing study of the FEI outputs over time, in as much detail as we can manage.

As hinted at in the previous Section, some channels and channel groups are present in some FEI trainings but not others. In this section we will investigate the sources of this issue and develop tools to validate trainings before they are released for general analysis. These studies begin with the FEIv4 trainings, where the user interface and training interface of the FEI have been updated.

4.2.1 `sigProb` is NaN

The FEIv4 training for MC7 was one of the first FEIv4 trainings released to analysts. In studying the semileptonic channels with this training, we noticed that the D^-e^+ channel in the `B0:semileptonic` FEI was not available after making the recommended cut on the FEI output (sigProb>0.001). This is a major semileptonic channel in data, and thus further investigation was required. Plots of the M_{bc} distribution for the semileptonic channel best candidates in this FEI version are shown in Figure 4.6, with the $\cos\theta_{BY}$ distribution in Figure 4.7.
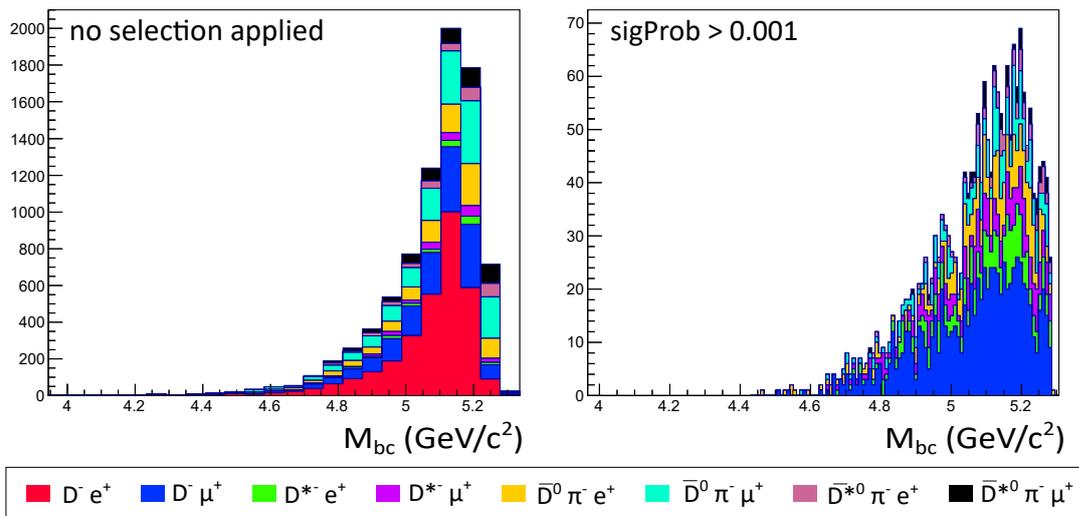
Figure 4.6: Comparison of M_{bc} for `B0:semileptonic` FEI results with FEIv4 MC7 training, with no cuts on signal probability (left) or with the default `sigProb`> 0.001 selection (right). In the right plot, the number of D^-e^+ candidates is exactly zero.

In Figure 4.6 we see that the channel is present in the FEI output lists when no cuts are made on the signal probability. The number of events categorised in this channel, however, is twice as large as the number of events in $D^-\mu^+$ channel, and this warrants further investigation.

More of the issue is revealed in Figure 4.7. We find that candidates in this channel have $\cos\theta_{BY}$ distribution peaking close to -10, when well-reconstructed semileptonic B decays are expected to have $\cos\theta_{BY}$ between -1 and 1. The source
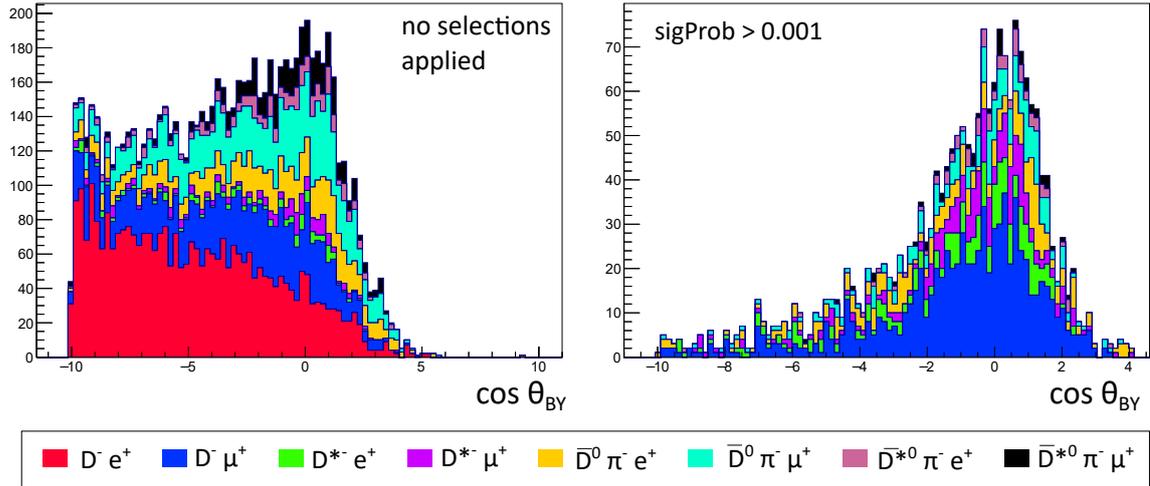
Figure 4.7: Comparison of $\cos\theta_{BY}$ for `B0:semileptonic` FEI results with FEIv4 MC7 training, with no cuts on signal probability (left) or with the default `sigProb>` 0.001 selection (right).

of this issue was found to be twofold: the FEI was assigning a signal probability of NaN (not a number) to all candidates in this channel, and an additional bug in `basf2` causing NaN values to be ranked higher than all other signal probability values in the `rankByHighest()` function used to assign the best candidate. As a result, this training could not be used for analysis of semileptonic tags if candidate selection was performed by `basf2`.

After this result, FEIv4 was further modified to ensure that the FEI would output `sigProb`=0 rather than NaN for channels with missing information. These channels are now removed from consideration in analysis by cuts on the signalProbability, and cannot be marked as best candidate unless no other candidates are present in the event.

4.2.2 `sigProb` is 0

Candidates with `sigProb`=0 are removed from skims with a selection on the minimum signal probability. Determining the channels in which these candidates occur, however, is a key part of monitoring the FEI development between software versions, as `sigProb` values of exactly one or exactly zero indicate edge cases in the classification.

One way to check for channels that are missing from the FEI output, even when loose signal probability selections are applied, is to create a 2D histogram "tag map," shown for MC11 `B0:generic` FEI in Figure 4.8. This tag map allows us to see at a glance which channels are never selected as the best candidate in an event, and with sufficiently high statistics, blank columns indicate channels that have either `sigProb`=0 for all candidates, or no candidates produced in that channel at all.

The other advantage of this approach is that it improves our understanding of the way the signal probability distribution behaves for each individual channel. Higher purity channels can be identified by eye, as these have larger numbers of candidates at the top of the plot, corresponding to more well-reconstructed events.

Figure 4.8: 2D histogram of $\log_{10}(\text{sigProb})$ and FEI decay channel for the `B0:generic` MC11 FEI training. A best candidate selection is applied.

We observe that there are four channels with no candidates shown in the map, and thus four channels that do not have any candidates with `sigProb`> 0. To specifically check whether these channels contain candidates with `sigProb`$=0$ or do not contain any candidates, we produce an additional Figure, Figure 4.9, from an output dataset without a selection on the signal probability. In this case, each column consists of only a single bin.

For this FEI training, all FEI channels are represented in the output. This is not necessarily the case for all trainings (see Section 4.2.4). Interestingly, we observe a very large number of candidates with `sigProb`$=0$ in the $D^{*-}\pi^+\pi^+\pi^-\pi^0$ channel, which also has some candidates with nonzero values.

4.2.3 `sigProb` **is 1**

Similarly to classifier failures where the signal probability may be set to exactly zero, classifier failures where `sigProb`$=1$ also occur in some trainings, often in the semileptonic FEI. Unlike `sigProb`$=0$, however, these classifier failures have an impact on best candidate selection if the candidates marked with `sigProb`$=1$ are not actually well reconstructed. A histogram of signal probability values for the semileptonic B^+ best candidates in the first MC9 FEI training is shown in Figure 4.10.

Figure 4.9: Channels containing candidates with `sigProb`=0. The 2D Histogram format is used to show clear bands to identify channels and allow comparison with Figure 4.8

Figure 4.10: Signal probability distribution for best candidates produced by the first MC9 FEI training for release-02. We notice that no candidates with `sigProb`=1 are marked as `isSignal`.

The number of best candidates with signal probability equal to 1 is somewhat large, but we see that none of these candidates are marked `isSignal` by the MC Matching criteria. These candidates with `sigProb=1` may also have strange physics distributions, as seen in Figure 4.11.

Figure 4.11: ΔE distibution for B_{tag}, where some candidates in this $D^- \pi^+ e^+$ channel of the `B:semileptonic` FEI have `sigProb=1`. This result is from an MC9 training for release-02. The peak around $\Delta E = -3$ is dominated by candidates with `sigProb=1`.

The underlying cause of this problem must be identified and addressed before the FEI signal probability is used to rank semileptonic candidates.

4.2.4 Preventing bugs in the signal probability

While this thesis focuses on testing the generic FEI trainings designed for use with any signal analysis, complementary studies of the performance of FEI trainings targeted to specific signal modes were performed by Nadia Toutounji at the University of Sydney [60]. These studies of the specific FEI training configuration—where a signal mode is reconstructed first and then the FEI is run to build a B_{tag} in the rest of the event—reveal that the source of many channel issues in the FEI are from under-training. On Belle II MC, the weight files from FEI trained in the generic configuration outperform those trained in the specific configuration in efficiency, purity,

and retention, even when both sets of files are used in the specific configuration [60].

Since this feedback was presented, the generic FEI has been trained with larger samples and the channel performance has improved. Unfortunately, these continuing issues are not yet well captured by the FEI software's inbuilt classifier training report. Data from the report for the MC11 FEI training is compared to the results of our validation process in Table 4.1.

Table 4.1: Comparison of the number of FEI channels meeting specific criteria, from the MC11 training

	# possible FEI channels	Channels successful in training report	Channels present in output	Channels with some `sigProb` $\neq 0$
B^0 Hadronic	26	26	26	22
B^+ Hadronic	29	29	29	25
B^0 semileptonic	39	39	34	34
B^+ semileptonic	42	42	39	34

At present, these issues remain open, but further validation of the FEI with MC and with data is currently in progress.

4.3 Recent FEI performance

As discussed above, while the FEI training report is produced for each training, this doesn't tell the whole story about the physics behaviour of the FEI. Instead, each FEI training must be tested before being used for skimming and before being released to analysts. In general, we expect to have approximately consistent behaviour between campaigns even though the tracking and reconstruction software are continuing to evolve, and major differences are cause for further investigation. We are also careful to apply the FEI to MC samples that were not used in the FEI training sample to avoid bias in the results.

In order to check the physics behaviour of the FEI, we check M_{bc}, ΔE, and $\cos \theta_{BY}$ (where relevant) for each FEI classifier, both in aggregate and also for each individual channel. The purity, efficiency, and retention of each channel are also recorded using `isSignal` to mark correct reconstructions.

Unfortunately, we cannot check for FEI missing channels and test the skim runtime and other parameters simultaneously. Instead, we run the FEI without any selections on the signal probability in our script and produce an output dataset including all `sigProb` values. The selections can then be applied at the plotting level to test the skim retention. In the sections to follow, the skimming selections are

applied in each Figure unless otherwise stated. For both the B^0 and B^+ Hadronic FEI, these selections are summarised:

- R2EventLevel < 0.4 ,

- nTracks ≥ 4 ,

- sigProb > 0.001 ,

- $|\Delta E| < 0.2$,

- $M_{bc} > 5.24$.

Further information about the skimming selections and the skimming process is available in Section 4.4.

The total number of diagnostic plots for the FEI is very large in order to check the behaviour of each B decay channel, and as a result, only an overview of the channel-by-channel distributions will be provided. Efficiency, purity, and retention results are also provided for the B^0 and B^+ hadronic FEI.

4.3.1 Channel-by-channel physics distributions

Comparisons of the M_{bc} and ΔE distributions between MC10 and MC11 are displayed in Figure 4.12 for a high purity channel, and Figure 4.13 for a low purity channel. In both the overall distribution and the best candidate distribution with skimming selections applied, we see very little change in shape when the MC version is changed. The MC11 distributions for the lower purity $D^{*-}\pi^+\pi^+\pi^-\pi^0$ channel (without skimming selections) are shown in Figure 4.14, and we recall that this channel has a number of candidates with sigProb$=0$. In particular, we observe that the best candidate distribution for M_{bc} doesn't have a distinct peak at the B mass, again an indication of the low purity in this channel.
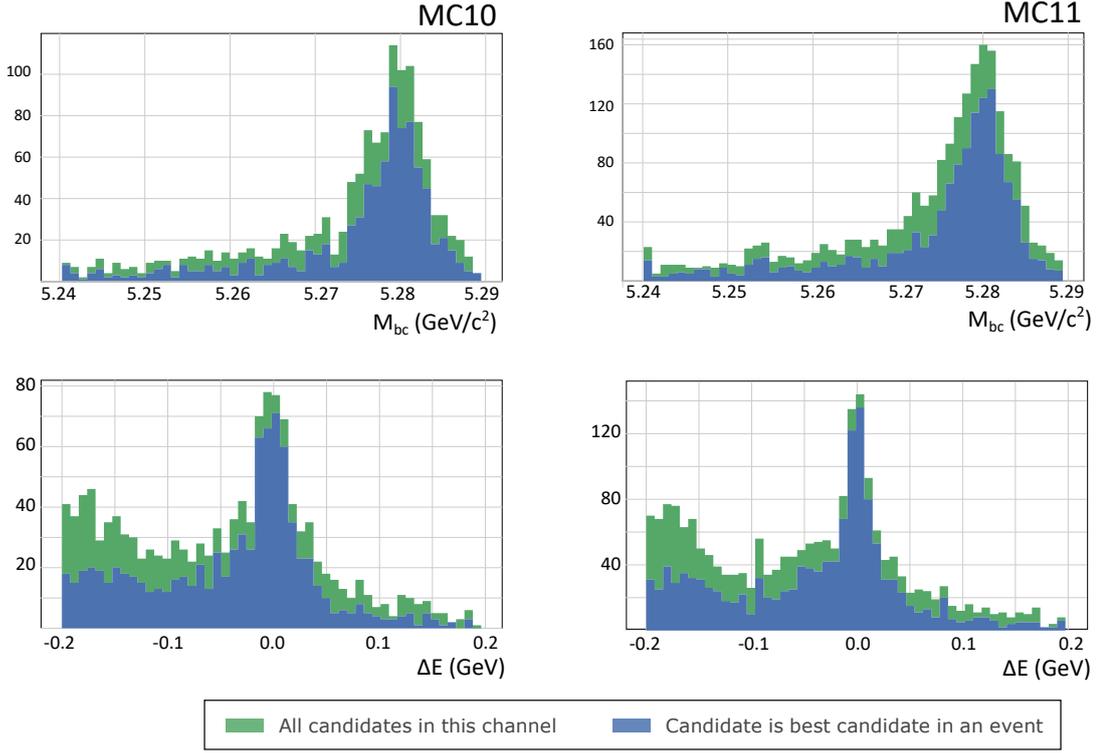
Figure 4.12: Comparison of M_{bc} (top) and ΔE (bottom) distributions between MC10 (left) and MC11 (right) for the high purity $D^-\pi^+$ channel of the B^0 Hadronic FEI. Skimming selections are applied.
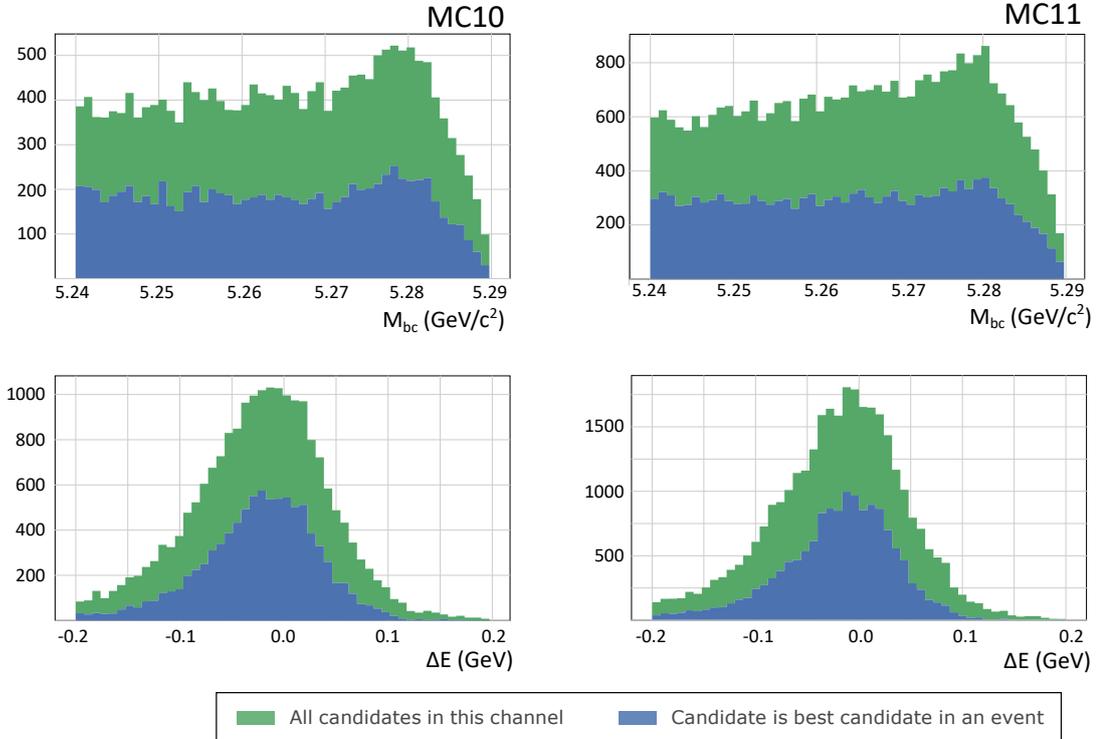
Figure 4.13: Comparison of M_{bc} (top) and ΔE (bottom) distributions between MC10 (left) and MC11 (right) for the lower purity $D^-\pi^+\pi^+\pi^-\pi^0$ channel of the B^0 Hadronic FEI. Skimming selections are applied.
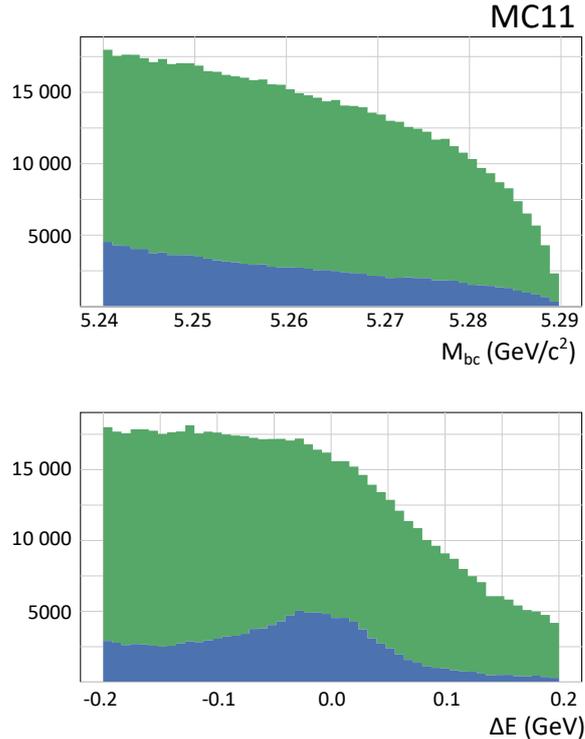
Figure 4.14: MC11 results for $D^-\pi^+\pi^+\pi^-\pi^0$ channel of the B^0 Hadronic FEI as in Figure 4.13, but with no skimming selections applied.

4.3.2 Hadronic B^0 FEI

We compare the channel-by-channel performance of the B0:generic FEI for MC9, MC10, and MC11 in the Figures below. We note that the efficiency and retention marked for each channel are relative to the total number of $B\overline{B}$ events in the input sample, rather than the number of generated events containing a B decaying in each individual channel. Instead, the comparison to the total $B\overline{B}$ allows the total efficiency and retention to be read from the Figure as the sum of the values for each channel.

If we initially focus on the purity in Figure 4.15, we see that MC10 and MC11 have improved purity at the skimming level compared to MC9. This could be for any number of reasons, but given that MC9 was somewhat notorious for having a lower reconstruction efficiency of neutrals than earlier MC, this is probably not a concern. MC9 was also the last Monte Carlo sample produced before the first official software version release-01.

Although purity has also decreased a little from MC10 to MC11 in many channels, from Figure 4.16 and 4.17 we see that we generally have more candidates and more correct candidates passing the skimming selections overall. Much of this improvement between MC11 and MC10, and indeed from MC9, is an increase in the retention rate of rarer channels an channels with larger numbers of pions. We hypothesise that these changes are due to improved reconstruction in the FEI in
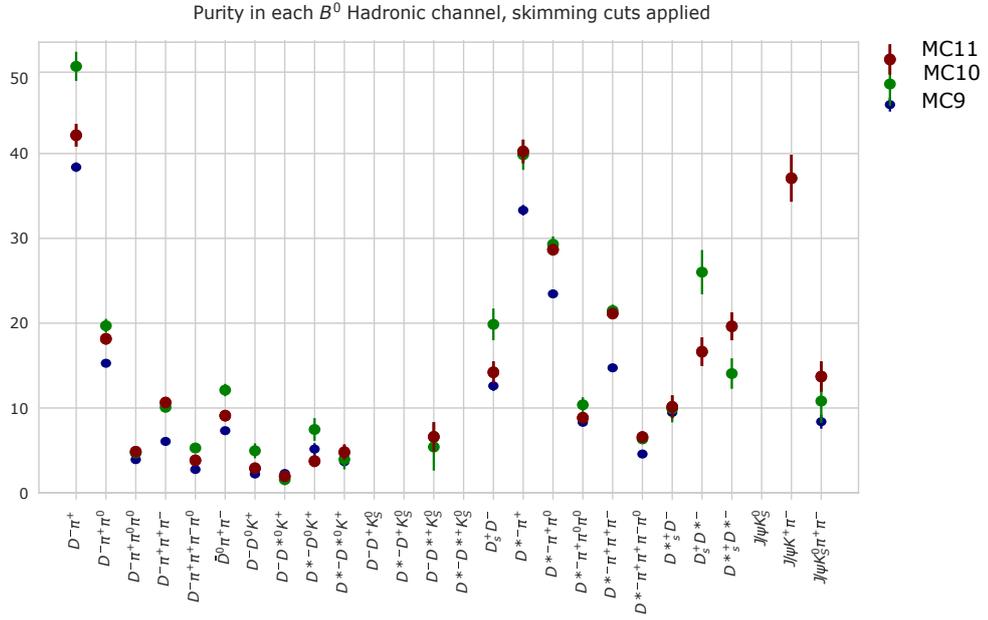
Figure 4.15: Comparison of purity of FEI B^0 tags in different modern FEI trainings
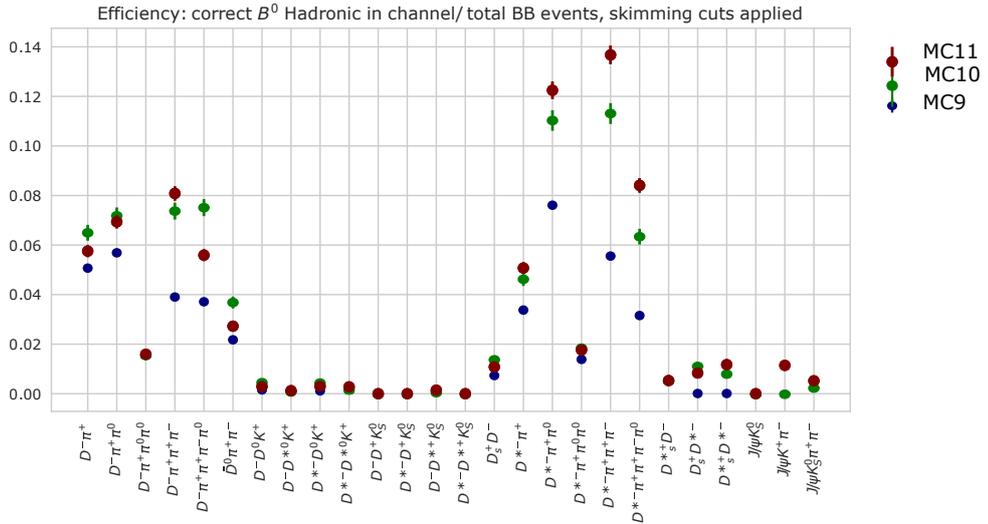
Figure 4.16: Comparison of efficiency of FEI B^0 tags in different modern FEI trainings
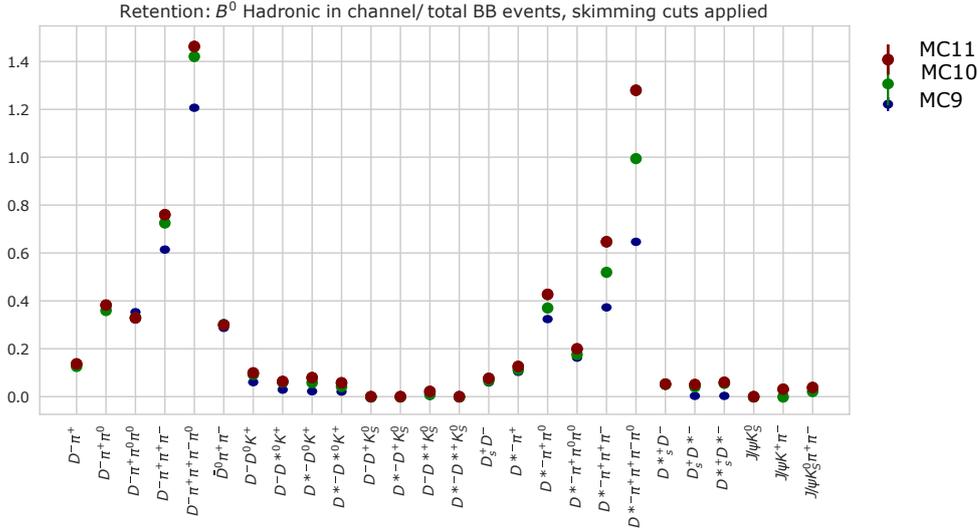
Figure 4.17: Comparison of retention of FEI B^0 tags in different modern FEI trainings

release-03 compared to earlier versions, and also the larger samples used to train the FEI. That is, we might consider the FEI as being "more confident" about its reconstructions and assigning higher `sigProb` values more often when trained on a larger MC sample. This matches well with evidence from the specific FEI studies in [60] discussed in Section 4.2.4.

4.3.3 Hadronic B^+ FEI

Similarly, comparisons of the FEI `B+:generic` tag performance in MC9, MC10, and MC11 with skimming selections are presented in Figures 4.18, 4.19, and 4.20.
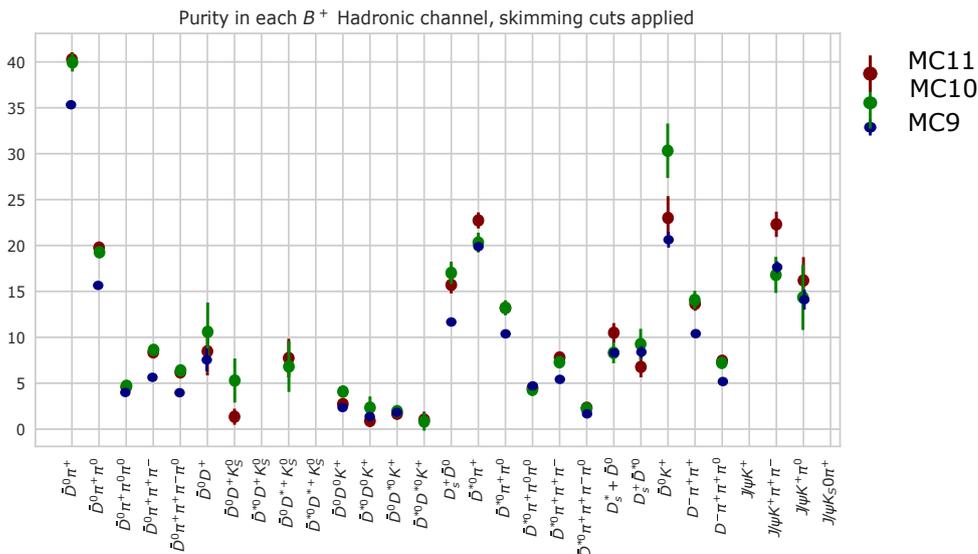
Figure 4.18: Comparison of purity of FEI B^+ tags in different modern FEI trainings
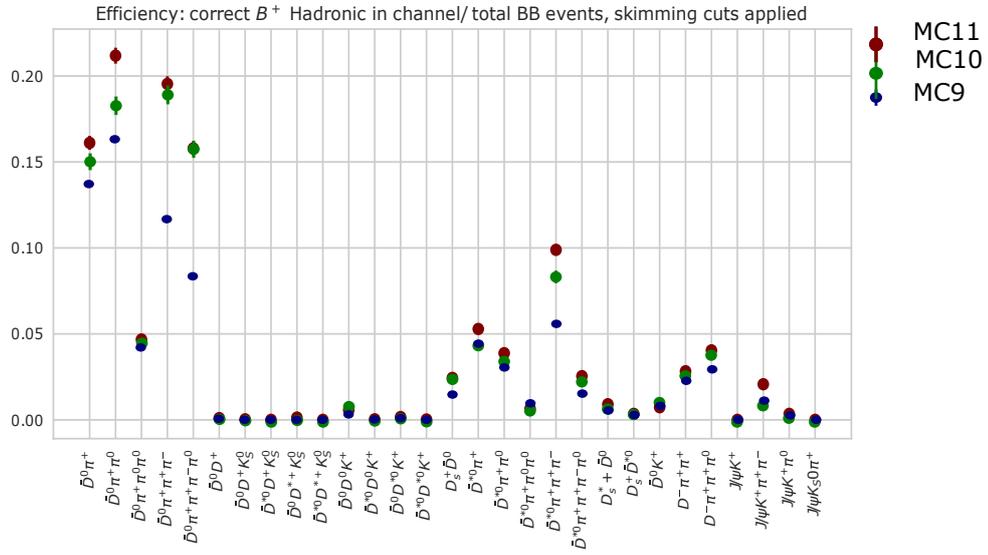
Figure 4.19: Comparison of efficiency of FEI B^+ tags in different modern FEI trainings
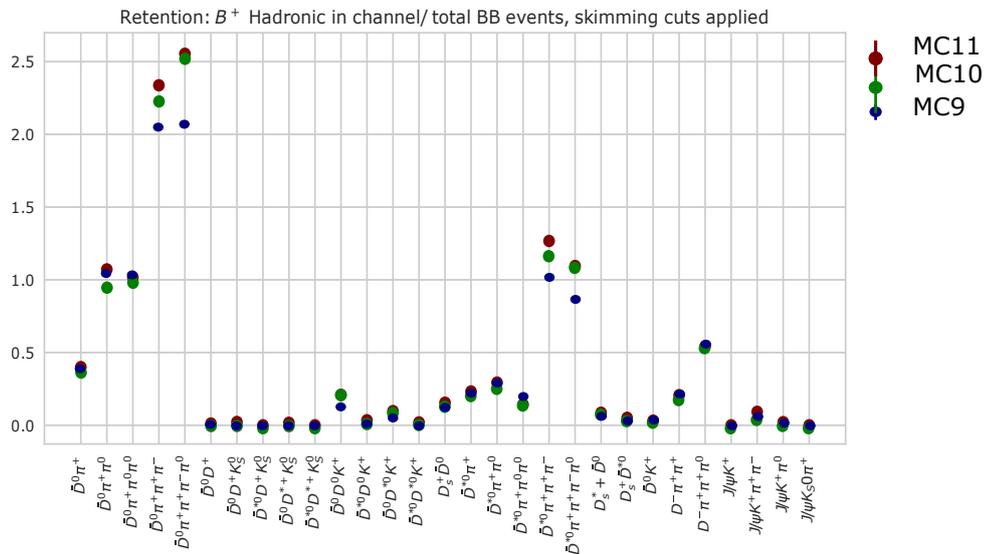
Figure 4.20: Comparison of retention of FEI B^+ tags in different modern FEI trainings

Qualitatively, the efficiency and retention of the B+:generic list is more consistent between MC releases than for the B0:generic list. We might argue that this is related the overall higher retention rate with the B^+ hadronic FEI. Nonetheless, we still see improvements in efficiency and retention (and often also in purity) as the size of the training sample increases between MC9 and MC11.

4.3.4 Performance summary

The overall purity, efficiency, and retention of the FEI tags combined across all channels are presented in Table 4.2.

Table 4.2: Table of FEI performance with skimming selections. Each FEI is applied to its maching MC sample. We note that while the skimming selections are the same for each of the hadronic FEI, the sigProb cut for B0:semileptonic and B+:semileptonic is different (see Section 4.4.2)

		MC11	MC10	MC9
B0:generic	purity	12	12	9.3
	efficiency	0.78	0.71	0.45
	retention	6.7	6.0	4.8
B+:generic	purity	9.5	9.1	7.6
	efficiency	1.1	1.0	0.80
	retention	12	11	10.5
B0:semileptonic	purity	14	14	
	efficiency	2.2	2.1	
	retention	16	16	
B+:semileptonic	purity	12	12	
	efficiency	1.5	1.5	
	retention	13	13	
Number events in test sample		980 000	640 000	5 720 000

As the FEI is separately trained on each MC campaign, the signal probability distribution will be slightly different in each case. In newer trainings of the FEI more events pass the sigProb selectionss originally developed for the MC9 skim, and so new skim selections may need to be chosen to keep the skim retention within goal parameters.

4.4   Skimming studies

Belle II generates large amounts of Monte Carlo simulations for analysis and validation, and will also store large amounts of data during production. To reduce the amount of times the full data sample must be accessed, we produce skims: smaller

data (or MC) samples consisting of only the events of interest to particular analyses. In most cases, skims are designed much like a signal B reconstruction, and collect specific physics modes with medium-to-tight selections. Conversely, the FEI skim is designed to collect hadronic or semileptonic tags to be used by a large number of different rare decay mode analyses.

The skimming process is designed to take .mDST files consisting of tracks, clusters, and vertex hits as input, to produce a micro-dst or .uDST file. These still contain all tracks and clusters, but for significantly fewer events. Additionally, particle lists reconstructed as part of the skim are stored in the .uDST, to prevent these having to be reconstructed again by analysts.

The skimming group has to balance the needs of physics analysts against computational constraints on the system. Physical constraints on run time, memory usage, and disk storage lead to physics-level constraints on retention and number of candidates per event. The FEI in particular is very memory-hungry and produces a large number of candidates of varying quality. To be able to run FEI skims effectively, it's important to reduce this footprint. An overview of the GRID computing limits and skimming targets are provided in Table 4.3.

Table 4.3: Limits and targets for early skim production. GRID limits are direct constraints in the distributed production system, while the combined skim goals convert these into limits that can be tested for smaller input samples on local hardware.

	GRID targets	Combined skim goal
total processing time	43200 sec	
runtime	0.19-0.29 sec/event	
CPU time		< 2 HEPsec/event
log file size	30MB/job	
memory total	2GB/job	
total size of all output .uDST		< 10% of input .mDST
event retention		< 10% per skim

The combined skim goals are designed for running multiple skims in a single script file and as a single GRID job. The main issue for the hadronic FEI skims is that the initial runtime and memory usage are very high compared to all other skims, and in fact the B^0 Hadronic FEI alone had > 7 HEPsec/event at the beginning of this work. The semileptonic FEI skim has additional issues with high retention. While we want to reduce the runtime and memory usage of each skim individually, our goal is also to produce suitable combined skims.

4.4.1 Reducing memory impact with event-level selections

Our main issue to address is that the FEI scans tracks and clusters on every event in order to look for particle candidates. Wherever possible, we want to try to remove events that will fail our final B meson selections before the FEI actually runs in order to reduce memory usage and runtime. Ideally, this is achieved by making cuts on variables that are available before any particles have been reconstructed.

We select on the minimum number of tracks, nTracks$>= 4$. At the lower limit nTracks$=4$, this allows reconstruction of $B \to l\nu$ or $B \to \tau\nu$ with a one-prong tau decay, paired with a low-track-count B_{tag}. Events with nTracks< 4 are typically $q\bar{q}$ events where some final state particles are outside of the detector acceptance, or events with $e^+e^- \to \tau^+\tau^-$. Although the leptonic B meson decay channels are always charged, this selection is applied to both B^0 and B^+ tag FEI for consistency and to allow for shared pre-selections in a combined skim.

An equivalent selection on the maximum number of tracks cannot be made. The FEI tag skims are intended to be used by many analyses, and without a clear idea of the maximum number of tracks in the reconstructed signal mode, we are unable to make a selection that will preserve all possible signals.

Having made our track number selections, we must now consider other variables to further reduce the number of events before the FEI is run. Because of the large number of FEI candidates returned from $c\bar{c}$ continuum, and because we know that the FEI itself does not perform continuum suppression, we consider using continuum suppression variables to help remove events that will fail the FEI selection.

The `R2EventLevel` variable is used to calculate the Fox-Wolfram $R2$ from the tracks and clusters in the event, using pion hypotheses for all tracks. Histograms of the `R2EventLevel` for events with reconstructed FEI tags are shown in Figure 4.21. We observe that well-reconstructed tags (e.g. with `isExtendedSignal`$=1$) mostly have `R2EventLevel`< 0.4.

We test the FEI skim with the `nTracks` and `R2EventLevel` selections individually, as well as together. The statistics for the B^0 Hadronic and B^+ Hadronic skim are measured individually, and also as part of a combined FEI hadronic skim. The results of these studies are shown in Tables 4.4 and 4.5.

The effect of the pre-selection is more obvious on the non-$B\overline{B}$ samples, where we see that our preselection has made a clear impact in reducing the runtime of the skim. On the $B^0\bar{B}^0$ sample, we find that the runtime of the original FEI B^0 skim and the combined skim are not particularly different. This is because the FEI configuration set-up is only loaded once and shared between skims. Similarly, the preselection has little effect on the retention or time, as our selection limits were selected to preserve candidates from $B\bar{B}$ events.
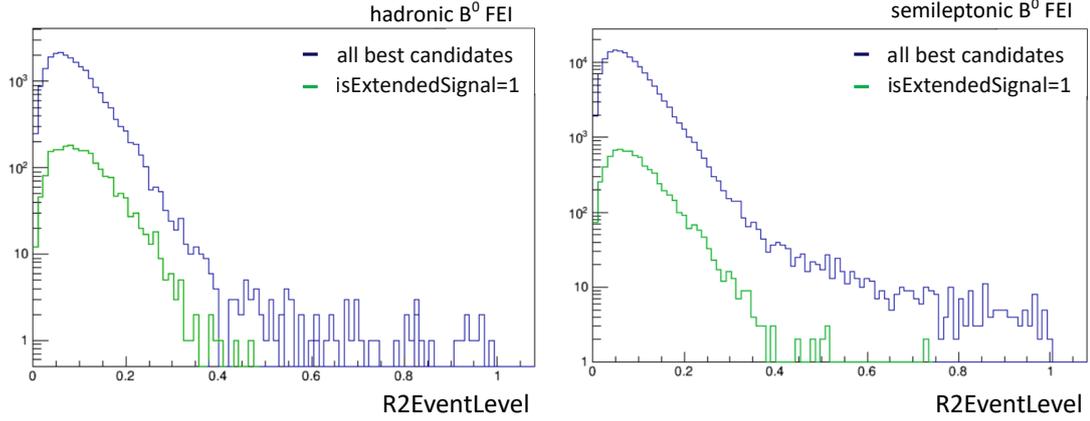
Figure 4.21: `R2EventLevel` variable distributions for the B^0 hadronic FEI (left) and B^0 semileptonic FEI (right). The distribution is shown for all best candidates (blue) and events where the best candidate has `isExtendedSignal`=1 (green).

Table 4.4: Comparison of skimming output for different scripts used to run the FEI Hadronic B^0 skim on $B^0\bar{B}^0$ ('mixed') BGx1 MC.

	Original skim	Combined	Combined + preselection	Combined + preselection + reduced FEI
Retention of B0	0.0389	0.0389	0.0389	0.0388
Time (hepsec) for skim	7.311	7.41	7.41	2.914
Average memory per event (KB)	55.8	56.5	56.5	49.7
B0 skim size over all samples (GB)	261.6	261.6	261.6	261.6
Full skim log size (GB)	0.972	0.98	0.98	0.529

Table 4.5: Comparison of skimming output for different scripts used to run the FEI Hadronic B^0 skim on $u\bar{u}$ BGx1 MC.

	Original skim	Combined	Combined + preselection	Combined + preselection + reduced FEI
Retention of B0	0.0311	0.0311	0.03	0.03
Time (hepsec) for skim	3.332	3.162	2.712	1.159
Average memory per event (KB)	52.46	53.31	53.66	46.3
B0 skim size over all samples (GB)	556.1	556.1	556.1	556.1
Full skim log size (GB)	1.664	1.678	1.684	0.904

The largest improvement is when using the 'reduced' FEI: the original skim was loading and producing the particle lists for semileptonic FEI candidates, but not using them or storing them to file. This contributes significantly to the run time of the FEI skim, as the semileptonic FEI has a naturally high retention rate, and without any selections on the signal probability, $> 95\%$ of $B\overline{B}$ events scanned are retained. By avoiding loading the configuration files for the semileptonic modes into memory, we cut the runtime of the Hadronic FEI combined skim by more than half compared to each individual skim. While the hadronic combined skim with preselection and reduced FEI was approved to run on the GRID, the semileptonic FEI requires further study to reduce retention and runtime.

4.4.2 Reducing runtime of the semileptonic tag skim

As briefly discussed in the previous section, the naturally high retention rate of the semileptonic FEI means that almost every event in a $B\overline{B}$ sample will be assigned at least one FEI semileptonic candidate. Although the retention and number of candidates per event in the output can be reduced by making selections on the signal probability, the time per event remains high due to the number of candidates reconstructed and then ultimately discarded (see Table 4.6). Pre-selections are thus crucial for the viability of an FEI semileptonic skim, as is the ability to combine these preselections for the B^0 and B^+ case.

Table 4.6: Comparison of skimming output for the FEI semileptonic B^0 skim. Signal probability selections are applied to the FEI output to reduce retention before saving events to file. ACMPE stands for Average Candidate Multiplicity Per (retained) Event.

	No selection	`sigProb` > 0.001	`sigProb` > 0.005
Retention	0.8205	0.1937	0.0881
Time	7.593	7.553	7.507
uDST Size / Event (kB)	17.18	15.09	16.3
total uDST Size (GB)	1691	350.7	172.3
ACMPE	8.327	2.943	2.351

Much like for the hadronic FEI, we wish to use selections on the `R2EventLevel` variable and `nTracks` to reduce the number of events where the FEI is run. We also wish to avoid loading FEI channels into memory when their output candidates will ultimately not be used. Before FEIv4, the semileptonic FEI consisted of only eight channels: $Dl\nu$, $D^*l\nu$, $D\pi l\nu$, and $D^*\pi l\nu$ with $l = e, \mu$. New channels were introduced where the B meson decays hadronically into a state including a D meson, but that D then decays semileptonically (see Table 3.1 for further details). These channels are loaded in memory during the semileptonic skim and candidates are produced,

but the number of candidates from these channels is low overall (approximately 10 % of candidates), as seen in Figure 4.22.
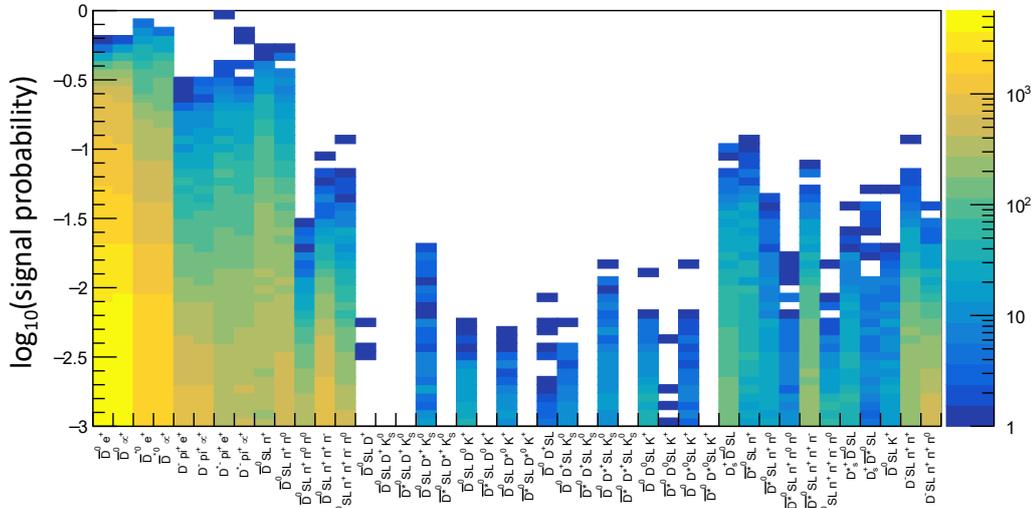
Figure 4.22: 2D histrogram of $\log_{10}$(sigProb) and FEI decay channel for the B^+ semileptonic FEI.

To prevent these low retention channels being loaded into memory, we must add a new flag to the FEI helper function `get_default_channels()` which loads the channel lists. By default, all semileptonic channels including the semileptonic D channels are loaded. By adding a new flag, we leave the default FEI behaviour unmodified for other analysts, but allow the extra channels to be removed for the skim. This is much more memory efficient than selecting on the channels after the FEI has run. We can see this effect by comparing the runtimes of combined skims on different MC data samples, shown in Table 4.7.

Table 4.7: Total runtime in seconds for skimming tests on 10,000 MC9 events from various $B\bar{B}$ samples or background types. Different versions of the skimming scripts are tested in otherwise identical conditions.

MC Sample	Hadronic Combined	Semileptonic Combined	FEI All Combined	FEI Combined, no SLD
mixedBGx1	14836	33933	36743	16083
chargedBGx1	15095	33653	35017	16735
ccbarBGx1	13384	29472	29650	13753
ssbarBGx1	9447	19818	20581	9454
uubarBGx1	10326	21543	20911	10888
ddbarBGx1	8905	17038	20345	8737
taupairBGx1	936	1381	1185	811

Removing the semileptonic D channels from memory significantly improves the performance of the combined skim, with the combined skim runtime only being

slightly larger than the runtime of the combined hadronic skim. With these memory issues addressed, the large retention rate (and thus the output skim size) for the semileptonic skim must now be considered.

4.4.3   Reducing retention in the semileptonic skim

While the retention rate may be reduced by making tighter selections on the FEI signal probability, in general analysts prefer that only loose selections be applied at the skim level to allow for sideband control studies. As a result, we seek physics-based selections that may be able to remove unsuitable semileptonic tag candidates.

Although adding an additional requirement to reconstruct a signal-side lepton will limit the generality of the semileptonic skim, we nonetheless explore this possibility in order to keep the signal probability selections loose. We seek to balance the requirements of leptonic, semileptonic, and semitauonic signal analyses in terms of limitations on the signal lepton.

Before discussing the addition of new selections, we summarise the starting semileptonic FEI skim selections below:

- `R2EventLevel` < 0.4

- nTracks ≥ 4

- `sigProb` > 0.001

- $-10 < \cos(\theta_{BY}) < 10$

We notice that this default cut on $\cos(\theta_{BY})$ is extremely loose when we expect $-1 < \cos(\theta_{BY}) < 1$ for semileptonic decays from theory. This inspires us to test a set of tighter selections on $\cos(\theta_{BY})$, provided we do not tighten these too far to prevent analyst fits to the $\cos(\theta_{BY})$ distribution. We choose $-5 < \cos(\theta_{BY}) < 3$ as our new selection.

The signal-side lepton selection is also intended to be loose to avoid overly restricting signal-side analyses. In software, this selection involves making a reconstruction of a dummy $B \to \ell$ decay with no restriction on the mass or momentum of the B. This B can then be combined with the reconstructed B_{tag} to form a dummy $\Upsilon(4S)$. Events without a signal-side lepton are then eliminated by requiring that the event contain one or more dummy $\Upsilon(4S)$ candidates. The signal-side lepton is chosen from pre-filled lists of electron or muon candidates provided by `basf2`. We select either all candidates, or choose candidates from the 95 % reconstruction efficiency list.

The performance of the skim when these selections are applied is shown in Table 4.8. We note that the runtime for these skim tests is quite large, as these first tests

were performed before the channels with semileptonic D mesons were removed from memory.

Table 4.8: Comparison of skimming output for the FEI semileptonic B^0 skim on a $B^0\bar{B}^0$ (mixed) BGx1 sample. The selections that include the signal side lepton also include the tightened $\cos(\theta_{BY})$ selection.

	Original selection	Tighten $\cos(\theta_{BY})$	Add lepton (any)	Add lepton (95 % eff.)
Retention of B0	0.2242	0.2073	0.1401	0.139
Time (hepsec) for skim	6.666	6.804	6.645	6.665
Average memory per event (KB)	60.6	60.6	60.2	60.3
B0 skim size over all samples (GB)	1450	1333	957.9	957.6
Full skim log size (GB)	0.767	0.758	0.771	0.775

The combination of the signal side lepton selection and the $\cos(\theta_{BY})$ selection is somewhat effective in reducing the retention, but we have still not reached our target of less than 10 percent. In order to produce a skim with these limits, a tighter FEI signal probability selection would be required.

Without any analyst demand for a semileptonic skim, however, further development of this skim has been low priority compared to skim production, FEI validation, and rare decay studies using the hadronic FEI tags. Now that Belle II Phase 3 data production has begun, studies of the viability of a semileptonic FEI skim (or semileptonic FEI tag + signal lepton skim) are gaining more interest and this work will be extended for more recent software and MC versions. Future studies of a semileptonic + lepton skim are suggested to use a minimum momentum selection on the signal-side lepton.

4.5 Semitauonic decay studies

Although many rare decay studies at Belle II are expected to use the FEI skims for tagging, we focus on semitauonic decays relevant to the $R(D)$ and $R(D^*)$ analysis. The $R(D)$ and $R(D^*)$ analysis with the FEI hadronic tag is considered one of Belle II's "Golden Modes", and the responsibilities of preparing for this analysis are shared among a team of developers. In this Section, we continue our focus on FEI performance in the context of a signal analysis.

The studies were performed using MC7 and MC9 Monte Carlo, using FEI trainings with binary PID.

4.5.1 Early studies with FEI SL tag representing signal

We begin with a toy analysis in order to develop intuition for further studies. As we already have some understanding of the properties of the FEI semileptonic tag, we use this as a stand-in for a semitauonic signal mode.

We reconstruct an $\Upsilon(4S)$ from one FEI hadronic B_{tag} and one FEI semileptonic B_{tag}. The semileptonic tags are chosen without any of the semileptonic D modes. For this analysis, we produce samples for each of the four FEI tags (B^0 and B^+, hadronic and semileptonic), and an additional sample of $\Upsilon(4S)$ candidates in both $B\bar{B}$ and $B^0\bar{B}^0$ channels. This analysis is also performed on both charged and mixed $B\overline{B}$ MC samples in order to help produce a wrong-B combinatorial background in our final plots.

A key aspect of this study is an investigation of using the FEI for best candidate selection compared to using $\Upsilon(4S)$ variables or semileptonic physics variables. We seek to test whether different methods of best candidate selection will result in similar purity of the output.

Best Candidate Selection

In this study, ranking of candidates is performed independently of the core `basf2` software. Information from the hadronic B_{tag}, semileptonic B_{tag}, and the reconstructed $\Upsilon(4S)$ is collected and candidates for all three particle types are stored together indexed by event. This allows us to test how the best candidate in an event may change when different selections are applied, while still keeping track of which candidate is which in a consistent way. Similarly, we are able to test how often the best hadronic tag candidate chosen via FEI signal probability is part of our final best $\Upsilon(4S)$.

We test the effectiveness of four kinds of best candidate selection:

- Choosing $\Upsilon(4S)$ with the best hadronic tag according to signal probability. If multiple $\Upsilon(4S)$ candidates remain, choosing the one with the best semileptonic tag.

- Choosing $\Upsilon(4S)$ with the best semileptonic tag according to signal probability. If multiple $\Upsilon(4S)$ candidates remain, choosing the one with the best hadronic tag.

- Selecting $\Upsilon(4S)$ where the sum of the signal probability for the hadronic tag and the semileptonic tag is greatest

- Selecting the $\Upsilon(4S)$ candidate with the minimum value of E_{extra}.

In practice, the best candidate selection using sum of the FEI signal probability produces the highest purity of $\Upsilon(4S)$ where both the hadronic and semileptonic tags are marked as `isSignalAcceptMissingNeutrino=1`. This is somewhat expected, as the FEI signal probability is trained using `isSignal` variables. Nonetheless, the selections are similar for all methods tested.

When the default selections are applied to the FEI tags, we calculate a purity of 6.88 % using the signal probability sum to select best candidates, or a purity of 4.75% when choosing the best candidate using E_{extra}. We choose the minimum E_{extra} as our method for selecting best candidates in the next Section as this best takes into account the quality of both B in a way that extends well to analysis with a different signal mode.

Results

Having made a best candidate selection, we can now study histograms of signal distributions at an event-by-event level. We reconstruct $\Upsilon(4S) \rightarrow B^0_{\text{hadronic}} \bar{B}^0_{\text{semileptonic}}$ on both mixed and charged Monte Carlo samples. Stacked histograms for $\cos(\theta_{BY})$ and M_{missing} are shown in Figure 4.23.
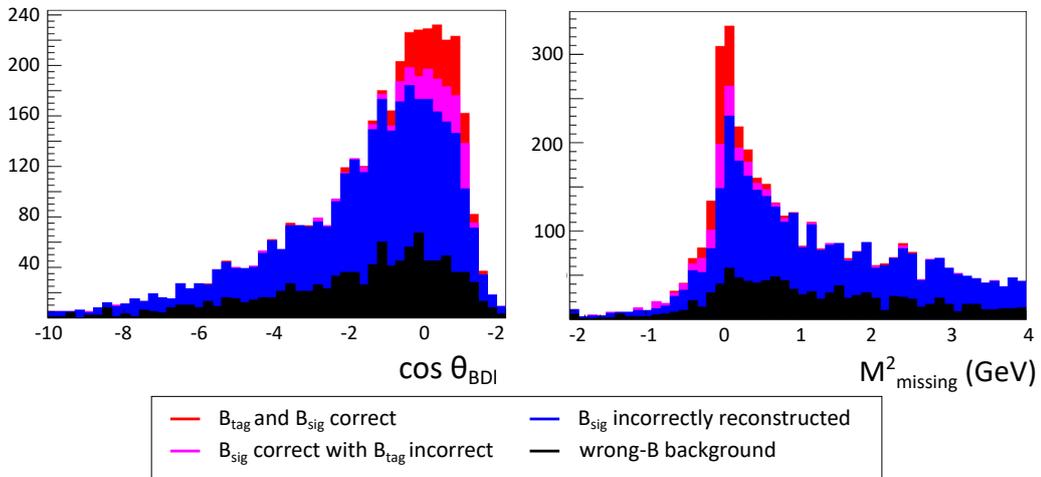
Figure 4.23: Stacked histograms of $\cos(\theta_{BY})$ (left) and M^2_{missing} (right) for $\Upsilon(4S) \rightarrow B^0_{\text{hadronic}} \bar{B}^0_{\text{semileptonic}}$ reconstructions. The wrong-B background shown in black is reconstructed on a B^+B^- sample.

We observe that almost all of the $B_{\text{semileptonic}}$ marked as `isSignalAcceptMissingNeutrino` have $\cos(\theta_{BY})$ between -1 and 1. Similarly, the distribution of correctly-reconstructed M^2_{missing} peaks at zero. In both cases, however, the background distribution also has a similar shape. Overall, about one fifth of the $\Upsilon(4S)$ we have reconstructed arose from B^+B^- events.

We can also consider the E_{extra} distribution in Figure 4.24. This particular implementation of E_{extra} is fairly naive, and will include contributions from beam back-

ground photons. This is easily seen from the way that the correctly reconstructed $\Upsilon(4S)$ still have an E_{extra} distribution peaking away from zero.
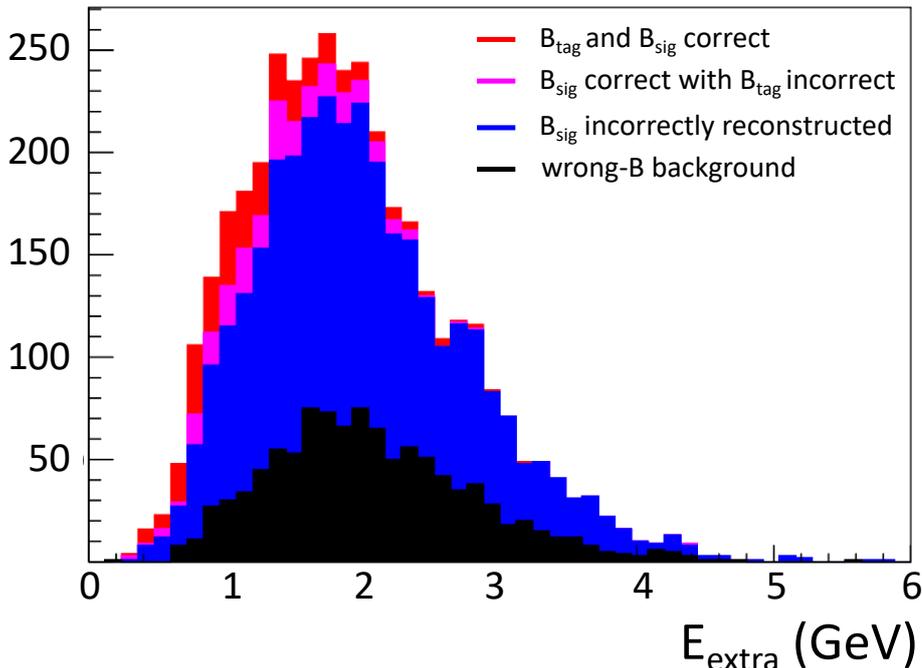
Figure 4.24: Stacked histogram of E_{extra} for $\Upsilon(4S) \to B^0_{\text{hadronic}} \bar{B}^0_{\text{semileptonic}}$ recon-structions. The wrong-B background shown in black is reconstructed on a B^+B^- sample.

In order to improve E_{extra}, timing cuts can be added so that clusters in the calorimeter originating from background rather than the e^+e^- collision may be re-moved. Selections can also be applied to ensure that no charged tracks remain after the $\Upsilon(4S)$ has been reconstructed. Unfortunately, in early software, the count of remaining tracks in the event often contained reversed clones of tracks used in the particle recostruction, limiting our ability to clean up the sample in this way.

Nonetheless, this toy study provided a good overview of the techniques needed to reconstruct the $\Upsilon(4S)$. In the next Section, we move on to using a signal recon-struction of the semitauonic decay instead of the FEI semileptonic tag.

4.5.2 Using early semitauonic signal selections

These studies were performed with early semitauonic group signal selections designed in release-01, applied to MC9 BGx0 FEI-skimmed data. $\Upsilon(4S)$ candidates with neutral B are reconstructed in $D^{(*)+}\tau\nu$ signal modes, with only the $\tau \to \mu$ lepton mode in this early stage. These early signal selections also use the MC matching information to produce additional constraints, for example ensuring that the D meson reconstructed in the $B \to D\tau\nu$ channel did not come from a D^* decay, and that the lepton is a true τ daughter. The best $\Upsilon(4S)$ is selected using the minimum

E_{extra}. We note that some timing cuts are applied for this measure of E_{extra}.

An overview of the tag and $\Upsilon(4S)$ retention on two different neutral MC samples is given in Table 4.9. Both of these samples are generated with one B^0 or $\bar{B}^0$ decaying generically, and the other B^0 or $\bar{B}^0$ decaying in either a $D\tau\nu$, $D^*\tau\nu$, or $D^{**}\tau\nu$ channel. $B^0\bar{B}^0$ mixing is simulated on these samples. The decay channels and branching ratios of these decays are chosen to match the PDG.

Table 4.9: Statistics for reconstructing $\Upsilon(4S) \rightarrow B^0_{\text{tag}}\bar{B}^0_{\text{semitauonic}}$ on different Monte Carlo samples. Quoted error for retention and purity is statistical.

MC sample type	$B^0 \rightarrow (D, D^*, \text{ or } D^{**})\tau\nu$	$B^0 \rightarrow D^{(*)}\tau\nu$
Events scanned	2.72×10^6	3.35×10^6
Events with FEI hadronic tags	8.97×10^4	1.08×10^5
Retention of FEI hadronic tags	3.30 ± 0.01 %	3.21 ± 0.01 %
Events with $\Upsilon(4S)$ reconstructed	1047	1109
Retention of $\Upsilon(4S)$ from tagged events	1.17 ± 0.04 %	1.03 ± 0.03 %
Retention of $\Upsilon(4S)$ from total events	0.038 ± 0.001 %	0.033 ± 0.001 %
Purity of $\Upsilon(4S)$	3.343 ± 0.005 %	5.410 ± 0.007 %
Purity of B_{signal} in $\Upsilon(4S)$	5.922 ± 0.007 %	8.656 ± 0.008 %

In total, we retain $0.03 - 0.04\%$ of input events on either sample. If we recall that we are only reconstructing $\tau \rightarrow \mu$ decays and take into account this branching fraction $\mathcal{B}(\tau \rightarrow \mu) = 17.39\%$ [3], we have retained 0.19% of $B \rightarrow D^{(*)}\tau(\rightarrow \mu\nu)\nu$ in the $B^0 \rightarrow D^{(*)}\tau\nu$ sample. Similarly in the $B^0 \rightarrow (D, D^*, \text{ or } D^{**})\tau\nu$ sample, we have retained 0.22% of events with $\tau \rightarrow \mu$ decays.

We are also interested in the relative fractions of D and D^* mesons reconstructed by our script. This additional information is provided in Table 4.10.

Table 4.10: Additional statistics for the semitauonic signal decay mode in our $\Upsilon(4S) \rightarrow B^0_{\text{tag}}\bar{B}^0_{semitauonic}$ reconstruction on different Monte Carlo samples.

MC sample type	$B^0 \rightarrow (D, D^*, \text{ or } D^{**})\tau\nu$		$B^0 \rightarrow D^{(*)}\tau\nu$	
Decay channel in signal candidate	D	D^*	D	D^*
Best candidates in channel	951	96	922	187
As % of $\Upsilon(4S)$ best candidates	91%	9%	83%	17%
% in MC sample	31%	49%	66 %	34%

Both MC samples are produced with more D^* than D, but our reconstruction efficiency of D^* modes is fairly low with early software.

As a tag study specifically, we wish to know if further tag selections can be used to improve the $\Upsilon(4S)$ or if the additional constraints from reconstructing a signal mode are sufficient. Figure 4.25 compares the M_{bc} and ΔE distributions for FEI B_{tag} before and after the signal selections are applied.
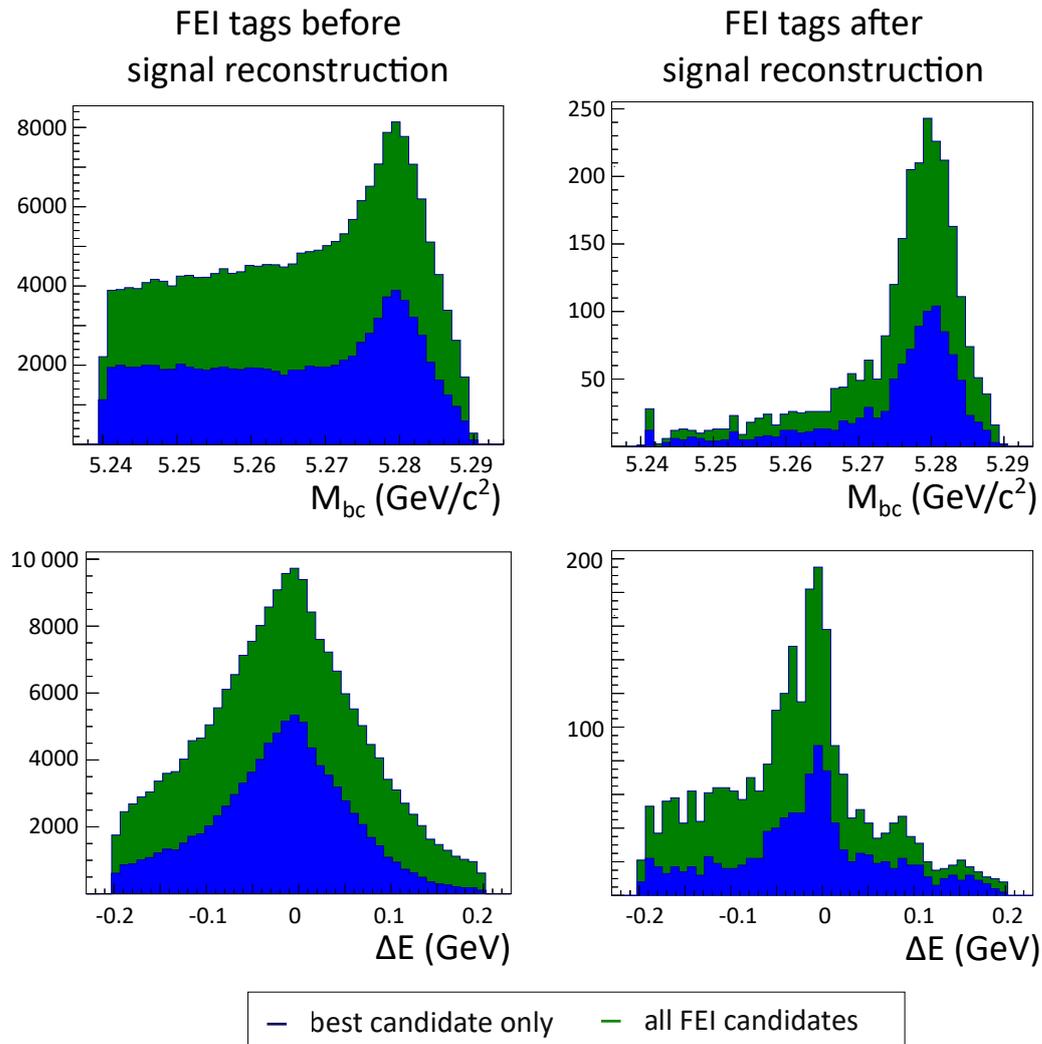
Figure 4.25: M_{bc} (top) and ΔE (bottom) distributions for FEI tags. In the two left plots, the distributions are shown for the FEI tag when no signal B is reconstructed. On the right, FEI tags must be part of a reconstructed $\Upsilon(4S)$.

We observe that the reconstruction of signal removes most of the large 'tail' of combinatorial background in the FEI tag distribution, without applying hard limits on the signal probability or the number of remaining tracks in the event. However, from our toy study with FEI semileptonic tags representing signal, we suspect that this restriction is due to the MCTruth selections applied during the signal reconstruction.

Overall, we are only able to reconstruct a small number of $\Upsilon(4S)$ to study due to the low retention rate of the FEI Hadronic skim for this MC version. As a result, we also consider a "cocktail" MC sample designed to boost tag retention for study. This sample is produced with a $B^0 \to D^-(\to K^+\pi^-\pi^-)\pi^+$ tag mode only, and with $B_{\text{signal}} \to D^{(*)}\tau(\to l\nu)\nu$ as the signal mode. The signal D decays generically rather than to any specific modes. This will give some qualitative indication of how often the FEI collects particles from the incorrect B meson when producing tag candidates, and how well these can be removed with signal selection. The comparison of the M_{bc} distribution for the B_{tag} before and after signal selection is shown in Figure 4.26. The comparison of the tag maps is shown in Figure 4.27.
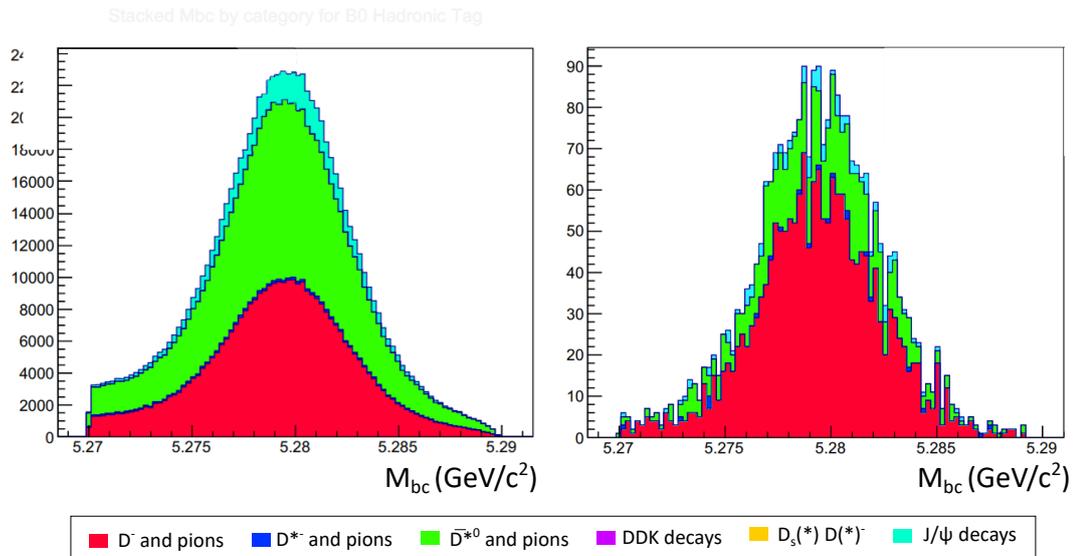
Figure 4.26: Comparison of M_{bc} distribution for FEI hadronic tags on the cocktail sample, either with only tags reconstructed (left) or with the entire $\Upsilon(4S)$ reconstructed.

Even without the signal selection, the M_{bc} distribution shows minimal combinatorial background as a result of the single tag mode chosen in the generation of the cocktail sample. We see that many of the tags are already reconstructed matching the generated mode, and in fact, the main other channel reconstructed in the sample is $\bar{D}^0\pi^+\pi^-$, which with the simplest reconstruction $\bar{D}^0 \to K^+\pi^-$, has the same final state particle content as our generated tag mode.

After applying the signal selection, the fraction of tags reconstructed in the

Figure 4.27: Comparison of the tag map for the B^0 Hadronic FEI on the cocktail sample, either with only tags reconstructed (top) or with the entire $\Upsilon(4S)$ reconstructed (bottom).

correct $B \to D\pi$ channel increases to 68%.

In the time since this study was performed, the `basf2` software has been updated significantly from release-01 to release-03, and the semitauonic signal selections have also been greatly improved and updated to be compatible with global PID. As a result, a repeat of these studies may be of interest within the coming year as part of the shared preparations for $R(D)$ and $R(D^*)$ studies with data.

4.6 Belle II rare decay prospects

The following prospects for Belle II are predicted in the Belle II physics book. These predictions use early Belle II MC and software (MC7), predicted to have approximately 0.5 % hadronic tag reconstruction efficiency. This efficiency is chosen to correspond to 10 % purity in order to compare to Belle performance.

$B \to D^{(*)}\tau\nu$ and $R(D^{(*)})$

Predictions for $R(D)$, $R(D^*)$, and the tau polarisation $P_\tau(D^*)$ in $B \to D^*\tau\nu$ decays are shown in Table 4.11. The Belle II prospect for the $R(D)$ and $R(D^*)$ combination is also shown in comparison to the 2018 World Average in Figure 4.28. The 50 ab^{-1} Belle II projection has a similar size to the world SM prediction.

Table 4.11: Predictions for overall statistical and systematic uncertainty in Belle II results at different data levels. Systematic error includes both theoretical and experimental systematics.

	Observable	Uncertainty stat.	Uncertainty syst.
1 ab^{-1}	$\mathcal{B}(B \to \tau\nu)$	29%	13%
5 ab^{-1}	$\mathcal{B}(B \to \tau\nu)$	13%	7%
	$R(D)$	6.0%	3.9%
	$R(D^*)$	3.0%	2.5%
	$P_\tau(D^*)$	0.18	0.08
50 ab^{-1}	$\mathcal{B}(B \to \tau\nu)$	4%	5%
	$R(D)$	2.0%	2.5%
	$R(D^*)$	1.0%	2.0%
	$P_\tau(D^*)$	0.06	0.04

For context, the Belle II results for $R(D)$ and $R(D^*)$ at 5 ab^{-1} were expected to have smaller uncertainty than the 2018 world average. Theory uncertainties in $|V_{cb}|$ and the $B \to D^{(*)}l\nu$ form factors largely cancel in $R(D^{(*)})$. The remaining predicted systematic uncertainty in these measurements is dominated by uncertainty in decays involving excited D resonances. Studies of both $B \to D^{**}\tau\nu$ and $B \to D^{**}l\nu$ decays are planned to help further reduce these uncertainties in the long term.
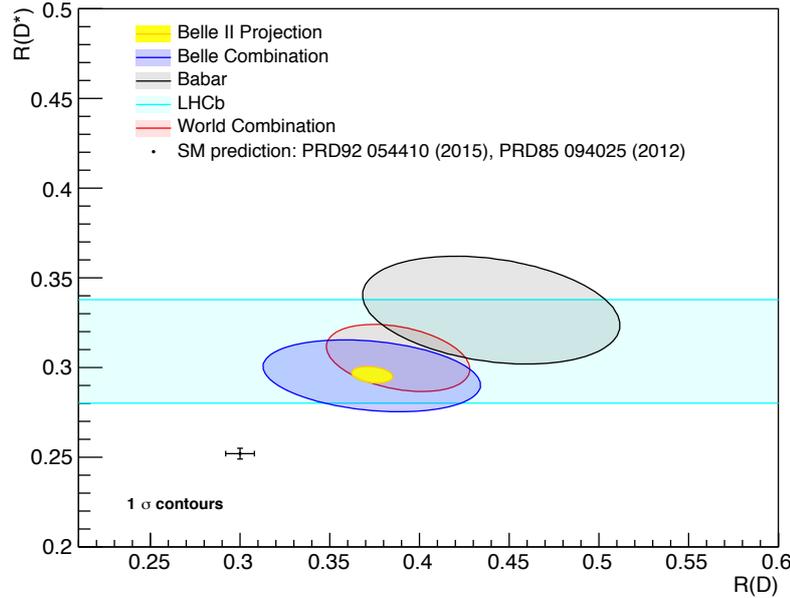
Figure 4.28:   Projection of $R(D^*)$ and $R(D)$ uncertainties for the full Belle II dataset (yellow), compared to 2018 experimental status and world averages. Figure also seen in [28].

With the full Belle II dataset, measurement of additional observables in $B \to D^{(*)}\tau\nu$ will also be possible. Both the tau polarisation $P_\tau(D^*)$ and q^2 distribution will be available, and if the $R(D^{(*)})$ anomalies continue, this polarisation may be able to give insight into the type of new physics effects causing the differences from the Standard Model. In particular, at 50 ab^{-1} $P_\tau(D^*)$ may be able to discriminate between scalar, vector, or tensor New Physics.

$B^+ \to \tau^+\nu$

Although the $B^+ \to \tau^+\nu$ channel has been observed as a world average, no single experiment has reached 5σ discovery individually. The most recent Belle result has 24% uncertainty [62], and higher luminosity at Belle II will allow improvement on this result. The Belle II measurement of $B \to \tau\nu$ will be made using FEI hadronic B tag and a one-prong tau decay in the signal B. Correctly reconstructed events will have no additional charged tracks and will not have missed any neutral particles arising from B decays.

Prospects for $B \to \tau\nu$ measurements are presented in Table 4.11. A Standard Model $B \to \tau\nu$ 5σ discovery is predicted around 2.6 ab^{-1}.

$B \to K^{(*)}\nu\overline{\nu}$

In the longer term, $B \to K^{(*)}\nu\overline{\nu}$ is a rare decay channel with a lot of potential for SM and new physics studies. Measurements of these modes will be possible at Belle

II, and the exact factorisation of hadrons and leptons in this decay allows precise experimental measurement of the $B \to K^{(*)}$ form factors.

Monte Carlo studies for this decay mode use the FEI Hadronic B tag, event shape constraints, and quality constraints on the K or K^* to determine an appropriate signal region. The current Standard Model estimate of the branching ratio for this decay has 10% uncertainty, which could be matched in experiment with the full 50 ab^{-1} dataset at Belle II. Evidence of a SM branching ratio is expected at 4 ab^{-1} and 5σ discovery is predicted with approximately 18 ab^{-1} of Belle II data.

4.7 Summary

In this Chapter, we have presented multiple tests of the FEI software in analysis. The FEI classifier performance is analysed, and bugs in the training for individual Monte Carlo samples are identified. For the modern MC11 FEI training, representative M_{bc} and ΔE distributions for high- and low-purity channels are presented, and the efficiency, purity, and retention of the FEI is compared to previous trainings.

We also develop our understanding of FEI performance in the context of analysis tasks. Hadronic and semileptonic FEI skims are developed and tested such that bulk skims can be produced for MC9, MC10, and MC11. A model semitauonic reconstruction is also performed using MC7 and MC9 FEI skimmed samples to assess the performance of the FEI in tagging.

Future Belle II results for rare decays depend on the success of FEI tagging, but also on advancements in lattice QCD that can reduce theory error. Predicted targets for future lattice QCD research are outlined in the Belle II Physics Book report [28]. For studies of $B \to \tau\nu$ decays, the B decay constant is of particular interest. In the next Chapter, we introduce the theory of lattice QCD in detail, in order to discuss our calculations of the B decay constants.

Chapter 5

Theory overview of flavour and lattice QCD

While the Chapter 1 focuses on the particles of the Standard Model and their interactions in the context of physical observables in experiment, in this Chapter we investigate the underlying theory in order to support our calculations of B observables in lattice QCD.

We begin with a brief overview of electroweak theory and the Higgs mechanism leading to the CKM matrix, and then discuss the properties of QCD with an emphasis on quark flavour effects. The discretisation of QCD onto the lattice will be discussed briefly in both a computational and theoretical context to provide background for our lattice studies of B mesons. The light and b quark actions used in this thesis are also described.

5.1 The Standard Model Lagrangian

The full Standard Model Lagrangian has form [63]

$$\mathcal{L}_{SM} = \mathcal{L}_{\text{gauge}} + \mathcal{L}_{\text{leptons}} + \mathcal{L}_{\text{quarks}} + \mathcal{L}_{\text{Higgs}} \tag{5.1}$$

where

$$\mathcal{L}_{\text{gauge}} = -\frac{1}{4} B_{\mu\nu} B^{\mu\nu} - \frac{1}{8} \text{Tr}(\mathbf{W}_{\mu\nu} \mathbf{W}^{\mu\nu}) - \frac{1}{2} \text{Tr}(\mathbf{G}_{\mu\nu} \mathbf{G}^{\mu\nu}), \tag{5.2}$$

$$\begin{aligned}
\mathcal{L}_{\text{leptons}} = & \begin{pmatrix} \bar{\nu}_L & \bar{e}_L \end{pmatrix} \bar{\sigma}^\mu i D_\mu \begin{pmatrix} \nu_L \\ e_L \end{pmatrix} + \bar{e}_R \sigma^\mu i D_\mu e_R + \bar{\nu}_R \sigma^\mu i D_\mu \nu_R + \text{H.C.} \\
& - \frac{\sqrt{2}}{v} \left[\begin{pmatrix} \bar{\nu}_L & \bar{e}_L \end{pmatrix} \Phi M^e e_R + \bar{e}_R \bar{M}^e \bar{\Phi} \begin{pmatrix} \nu_L \\ e_L \end{pmatrix} \right] \\
& - \frac{\sqrt{2}}{v} \left[\begin{pmatrix} -\bar{e}_L & \bar{\nu}_L \end{pmatrix} \Phi^* M^\nu \nu_R + \bar{\nu}_R M^\nu \Phi^T \begin{pmatrix} -e_L \\ \nu_L \end{pmatrix} \right],
\end{aligned} \tag{5.3}$$

$$\mathcal{L}_{\text{quarks}} = \begin{pmatrix} \bar{u}_L & \bar{d}_L \end{pmatrix} \tilde{\sigma}^\mu i D_\mu \begin{pmatrix} u_L \\ d_L \end{pmatrix} + \bar{u}_R \sigma^\mu i D_\mu u_R + \bar{d}_R \sigma^\mu i D_\mu d_R + \text{H.C.}$$

$$- \frac{\sqrt{2}}{v} \left[\begin{pmatrix} \bar{u}_L & \bar{d}_L \end{pmatrix} \Phi M^d d_R + \bar{d}_R \bar{M}^d \bar{\Phi} \begin{pmatrix} u_L \\ d_L \end{pmatrix} \right] \tag{5.4}$$

$$- \frac{\sqrt{2}}{v} \left[\begin{pmatrix} -\bar{d}_L & \bar{u}_L \end{pmatrix} \Phi^* M^u u_R + \bar{u}_R M^u \Phi^T \begin{pmatrix} -d_L \\ u_L \end{pmatrix} \right],$$

$$\mathcal{L}_{\text{Higgs}} = \frac{(D_\mu \Phi)^\dagger D^\mu \Phi - m_h^2 [\bar{\Phi}\Phi - v^2/2]^2}{2v^2}. \tag{5.5}$$

The notation H.C. is used to represent the Hermitian conjugate of preceeding terms. The Lagrangians $\mathcal{L}_{\text{leptons}}$ and $\mathcal{L}_{\text{quarks}}$ contain an implicit sum over the three generations of quarks. We note that the fields $\mathbf{W}_\mu$ that make up $\mathbf{W}_{\mu\nu} = \partial_\mu \mathbf{W}_\nu - \partial_\nu \mathbf{W}_\mu + ig/2[\mathbf{W}_\mu, \mathbf{W}_\nu]$ in this notation are matrix-valued as they have absorbed the Pauli matrices for the SU(2) symmetry as part of the field.

In this thesis, we are most concerned with weak decays. In these Lagrangians, there are no explicit weak decay terms, and instead terms dependent on $\mathbf{W}_\mu$ are hidden in the derivative operator D_μ as applied to left-handed fermions and the two-component Higgs field Φ. In the next Section, we discuss the electroweak interaction in some additional detail.

5.1.1 Electroweak theory

The physically observable states of the electroweak theory – $W^\pm$, Z, and the photon – are rotations and combinations of the underlying electroweak boson fields $\mathbf{W}$ and B. The Higgs mechanism is responsible for this rotation and for the Z and W bosons acquiring mass.

In order to explain the Higgs mechanism, we begin with a toy Lagrangian

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + (D_\mu \phi)^* (D_\mu \phi) - V(\phi^*, \phi) \tag{5.6}$$

of a single scalar Higgs field ϕ, where

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu, \tag{5.7}$$

$$D_\mu = \partial_\mu + iq A_\mu \tag{5.8}$$

for a vector field A and coupling constant q, and the potential V specifies the form of the self-interactions of ϕ. We consider

$$V(\phi^*, \phi) = \lambda(\phi^* \phi - v^2)^2, \quad \lambda > 0 \tag{5.9}$$

which is gauge invariant and has a 'Mexican hat' shape when plotted in $\mathbf{Re}(\phi), \mathbf{Im}(\phi)$

space. We also notice that

$$\frac{\partial V}{\partial |\phi|} = 2\lambda(|\phi|^2 - v^2) \times 2|\phi| \tag{5.10}$$

which has its minimum value when $|\phi|^2 = v^2$. In the quantum theory, ϕ fluctuates around the classical vacuum state $\phi_{\min} = v$. We expand the field ϕ around this minimum such that $\phi = v + H$ for some new scalar field H, and then the Lagrangian can be written

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + (\partial H)^2 + q^2 v^2 |A_\mu|^2 - \lambda |H|^2 + \text{additional interaction terms.} \tag{5.11}$$

We observe that the vector field A_μ has acquired a mass-like term as a result of this expansion around the minimum.

Upgrading the toy model, in an electroweak $SU(2) \times U(1)$ theory we require multiple components for the Higgs field and choose

$$\Phi = \begin{pmatrix} \phi_+ \\ \phi_0 \end{pmatrix} \tag{5.12}$$

where ϕ_+ has charge 1 and ϕ_0 has charge 0. Then the vacuum minimum can be chosen to have the form

$$\Phi_{\min} = \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{5.13}$$

resulting in a general Higgs field of the form

$$\Phi(x) = \begin{pmatrix} 0 \\ v + \frac{H(x)}{\sqrt{2}} \end{pmatrix} \tag{5.14}$$

in the expansion, represented in unitary gauge with real field H. In this theory, the covariant derivative is

$$D_\mu = \partial_\mu + \frac{ig_2}{2}\mathbf{W}_\mu - \frac{ig_1}{2}B_\mu \tag{5.15}$$

$$= \partial_\mu - ig_2 A_\mu^j \frac{\sigma^j}{2} - \frac{ig_1}{2}B_\mu \tag{5.16}$$

where the second line includes a change of notation to split $\mathbf{W}$ to specifically include the Pauli matrices σ^j and vector fields A_μ^j. Then by using the explicit forms of the Pauli matrices, we can rewrite the covariant derivative as

$$D_\mu = \partial_\mu - \frac{i}{2}\mathcal{A}_\mu, \tag{5.17}$$

with

$$\mathcal{A}_\mu = \begin{pmatrix} g_1 B_\mu + g_2 A_\mu^3 & \sqrt{2} g_1 W^+ \\ \sqrt{2} g_2 W^- & g_1 B_\mu - g_2 A_\mu^3 \end{pmatrix} \tag{5.18}$$

and

$$W_\mu^\pm = \frac{1}{\sqrt{2}} (A_\mu^1 \mp i A_\mu^2) \tag{5.19}$$

so that the Lagrangian term $(D_\mu \Phi)^\dagger (D^\mu \Phi)$ contains multiple mass terms

$$(D_\mu \Phi)^\dagger (D^\mu \Phi) = \frac{1}{2} (\partial H)^2 + \frac{1}{2} v^2 g_2^2 |W_\mu|^2 + \frac{1}{4} v^2 (g_1 B_\mu - g_2 A_\mu^3)^2 + \dots \tag{5.20}$$

leading to the definition

$$Z_\mu = \frac{-g_1 B_\mu + g_2 A_\mu^3}{\sqrt{g_1^2 + g_2^2}} \tag{5.21}$$

for the Z boson.

To continue our discussion of weak mixing relevant to this thesis, we consider the part of $\mathcal{L}_{\text{quarks}}$ corresponding to the interactions between the W bosons and the quarks. Written explicitly and simplified, for the W^+ boson these terms have the form

$$\mathcal{L}_{Wq} = \frac{g}{\sqrt{2}} W_\mu^+ \bar{U}_L \gamma^\mu D_L \tag{5.22}$$

where U and D are vectors of all up-type and down-type quark weak eigenstates. As discussed in Chapter 1, the weak eigenstates are different from the quark mass eigenstates. The masses of the quarks are generated via the Yukawa couplings λ in the interaction with the Higgs Φ, written

$$\mathcal{L}_{\Phi q} = -v \left(\bar{U}_L \lambda_u U_R + \bar{U}_R \lambda_u^\dagger U_L + \bar{D}_L \lambda_d D_R + \bar{D}_R \lambda_d^\dagger D_L \right) \tag{5.23}$$

when we consider $H(x) = 0$. Again, as U and D are in the weak eigenstate basis, these λ are not necessarily diagonal. Instead, we diagonalise λ via the introduction of four unitary matrices $V_{L/R}^{u/d}$, such that we can write

$$\frac{1}{v} \bar{U}_L \lambda_u U_R = \bar{U}_L V_L^{u\dagger} V_L^u \left(\frac{1}{v} \lambda_u \right) V_R^{u\dagger} V_R^u U_R \tag{5.24}$$

$$= \bar{\tilde{U}}_L \tilde{\lambda}_u \tilde{U}_R \tag{5.25}$$

for the diagonalised matrix

$$\tilde{\lambda}_u = \frac{1}{v} \begin{pmatrix} m_u & & \\ & m_c & \\ & & m_t \end{pmatrix} \tag{5.26}$$

corresponding to the mass eigenstate basis. The weak interaction can then be written

$$\mathcal{L}_{Wq} = \frac{g}{\sqrt{2}} W_\mu^+ \tilde{\bar{U}}_L \gamma^\mu V_L^u V_L^{d\dagger} \tilde{D}_L \tag{5.27}$$

$$= \frac{g}{\sqrt{2}} W_\mu^+ \tilde{\bar{U}}_L \gamma^\mu V_{\text{CKM}} \tilde{D}_L \tag{5.28}$$

5.1.2 QCD theory

When we consider only QCD, we represent all fields in the mass eigenstate basis and simplify the quark masses to exclude reference to the Yukawa couplings. In pure QCD without electromagnetic, weak, or Higgs interactions, quark flavour is represented entirely by the mass.

The QCD lagrangian has form

$$\mathcal{L}_{\text{QCD}} = \sum_{\text{quarks}} \left(\bar{\psi}_j^a(x) \left(i \slashed{D}_{jk}^{ab}(x) - m\delta_{jk}\delta^{ab} \right) \psi_k^b(x) \right) - \frac{1}{2} G_{\mu\nu}^{ab} G_{ba}^{\mu\nu}(x) \tag{5.29}$$

where colour is represented by roman indices a, b, Dirac spin indices are represented j, k and the covariant derivative has explicit form

$$\slashed{D}_{jk}^{ab}(x) = \gamma_{jk}^\mu D_\mu^{ab}(x) \tag{5.30}$$

$$= \delta^{ab} \slashed{\partial}_{jk} + ig\slashed{A}_{jk}^{ab}(x). \tag{5.31}$$

The quark fields transform under the fundamental representation of the gauge group SU(3), where gauge transformations Ω can be written as the implicit sum $\Omega = \exp(i\theta^r \frac{\lambda_r}{2})$ for each λ_r one of the eight Gell-Mann matrices and θ a free parameter. It is the introduction of γ^μ and the gluon fields A into the covariant derivative that allow the quark fields to be invariant under both global and local gauge transformations. The gluon fields transform under the adjoint representation of SU(3), and can be expressed as

$$\slashed{A}_{jk}^{ab}(x) = \gamma_{jk}^\mu A_\mu^r(x) \frac{\lambda_r^{ab}}{2}. \tag{5.32}$$

This Lagrangian gives rise to three types of QCD interaction vertices, shown in Figure 5.1.

This self-interaction between gluons gives QCD a non-trivial vacuum structure, and helps to make the dynamics of QCD highly nonlinear. In electroweak theory, we are able to consider perturbative approaches where cross-sections and decay rates are calculated by considering simple tree-level Feynman diagrams and then adding additional diagrams as corrections, because the coupling constant is small enough that each additional vertex decreases the contribution of the Feynman diagram to the final result. For low-energy QCD related to the formation of hadrons, however,
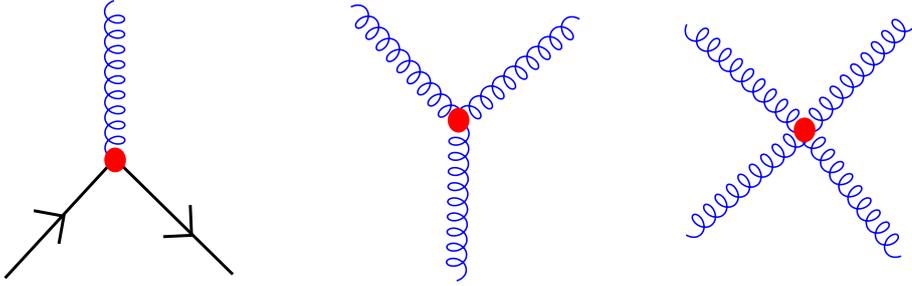
Figure 5.1: QCD interaction vertices. Quarks (solid lines) cannot interact directly with quarks of other flavours. Multiple gluons (spiral lines) may interact.

this is not the case. The effective QCD coupling constant α_S is shown against the energy scale Q in Figure 5.2.
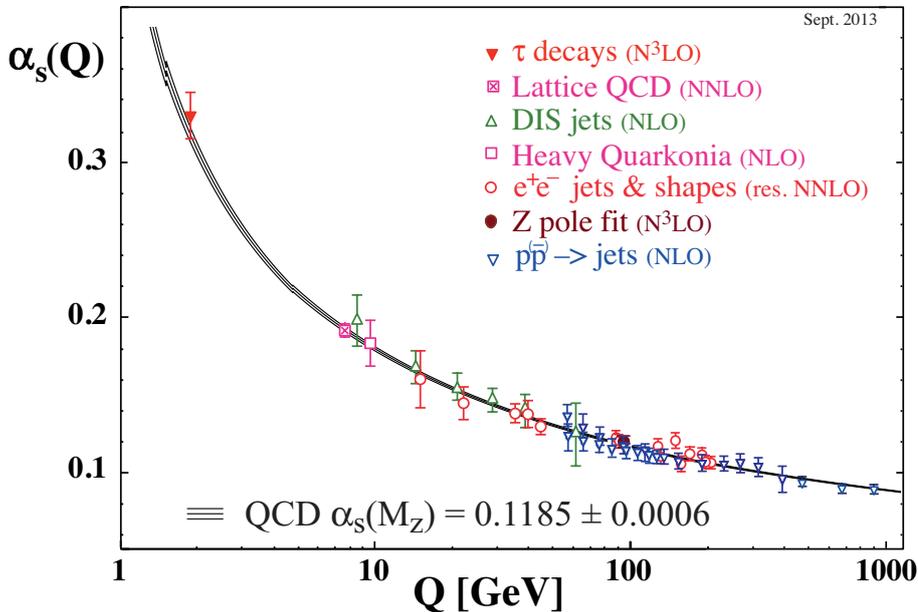
Figure 5.2: The running of the effective strong coupling α_S as a function of the energy scale Q. Image from [64].

While QCD is perturbative at high energies, when $Q^2 < 1$ the effective coupling constant becomes greater than one, and higher-order diagrams and lower-order diagrams may contribute similarly to the final result. Instead, a non-perturbative calculation—lattice QCD—is required in order to probe the internal structure of hadrons. The formulation of QCD on the lattice will be discussed in Section 5.2.

Returning to the symmetry properties of QCD, we find that while the QCD Lagrangian has an explicit SU(3) colour symmetry, additional symmetries are present when we consider the masses of the quarks. If we consider for a moment all flavours of quarks having equal masses, we notice an SU(N_f) symmetry and can write

$$\delta\psi^f = -i\delta\theta^r (T^r)^f_g \psi^g \tag{5.33}$$

where we are using the indices f and g to represent flavour, and again $1 < r < N_f^2 - 1$ is used to cycle through the generating matrices T^r of $SU(N_f)$. Using Noether's theorem, we have $(N_f^2 - 1)$ conserved currents

$$j_\mu^r(x) = -i \frac{\partial \mathcal{L}_{\text{QCD}}}{\partial(\partial^\mu \psi^f)} (T^r)_g^f \psi^g = \bar{\psi}_f \gamma_\mu (T^r)_g^f \psi^g \tag{5.34}$$

which can be used to construct charges

$$Q^r = \int d^3x \, j_0^r(x) \tag{5.35}$$

that commute with the Hamiltonian of the system and satisfy the $SU(N_f)$ algebra.

Now we consider a situation in which the quark masses vary. The mass term in the Lagrangian is no longer invariant under $SU(N_f)$ and the Noether currents are not conserved. However, if the mass differences are much smaller than the interaction scale Λ_{QCD}, we can treat the differences in the quark masses as perturbations with respect to the $SU(N_f)$ symmetry. As the physical Λ_{QCD} is approximately 300 MeV, we can create both $SU(2)$ and $SU(3)$ approximate symmetries with the lightest quarks. The $SU(2)$ symmetry of the u and d quarks, called isospin symmetry, is related to the similar masses of the proton and neutron.

The $SU(3)$ quark symmetry represents the u, d, s quarks in the representation $\mathbf{3}$ and the antiquarks in the representation $\bar{\mathbf{3}}$. Then mesons – combinations of one quark and one antiquark – can be written as a direct sum of an octet representation and a singlet according to

$$\mathbf{3} \otimes \bar{\mathbf{3}} = \mathbf{8} \oplus \mathbf{1}. \tag{5.36}$$

This allows us to represent the pseudoscalar mesons as shown in Figure 5.3. A similar octet can be produced for the vector mesons.
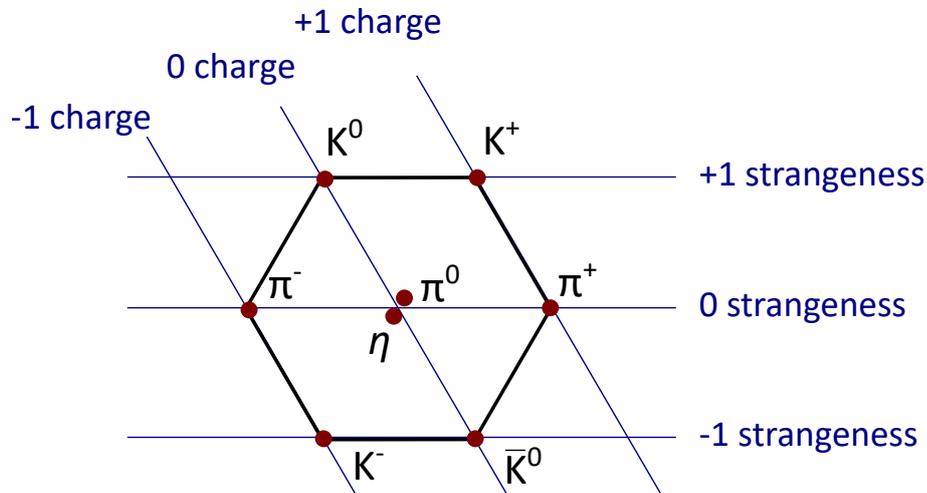
Figure 5.3: Light psuedoscalar mesons in the octet representation of SU(3)

In a similar fashion, the chiral symmetry of the Lagrangian is constructed by considering γ_5 alongside the generating matrices T^r of the flavour symmetry (together, this is often called an axial flavour transformation). This chiral symmetry will only hold in the limit $m \to 0$ due to the presence of terms

$$\delta\mathcal{L}_{\text{QCD}} = 2im\delta\theta^r \bar{\psi}_f (T^r)^f_g \gamma_5 \psi^g \tag{5.37}$$

when the Lagrangian is transformed. The axial current

$$j^r_{5\mu}(x) = \bar{\psi}_f \gamma_\mu \gamma_5 (T^r)^f_g \psi^g \tag{5.38}$$

can be used to form axial charges Q^r_5 which can be combined with the flavour charges Q^r to form a combined chiral algebra. In fact, this algebra can be divided into separate left-handed and right-handed algebras that do not interact, allowing the chiral group to be written $SU(N_f)_L \otimes SU(N_f)_R$ with a subgroup $SU(N_f)_V$ of the vector flavour currents. Then returning to the circumstances of physical QCD with three lighter flavours, the approximate chiral symmetry can be written $SU(3)_L \otimes SU(3)_R$ with an SU(3) flavour subgroup.

Spontaneous symmetry breaking from vacuum state of the axial charges breaks $SU(3)_L \otimes SU(3)_R$ and leaves just the SU(3) flavour subgroup, suggesting eight massless pseudoscalar mesons called Nambu-Goldstone bosons. The additional explicit symmetry breaking from the non-equal quark masses then causes these eight mesons to acquire mass. In the chiral limit, the SU(3) octet of light pseudoscalar mesons match the eight Nambu-Goldstone bosons. As a result, the SU(3) flavour symmetries in the meson sector are important to consider in QCD calculations, especially when the quark masses are varied.

5.2 Formulating QCD on the lattice

Lattice QCD considers spacetime as a finite, discretized, Euclidean lattice. The non-trivial vacuum structure of QCD is then interpreted as a number of different individual vacuum configurations, which can be weighted and averaged over using similar techniques to statistical mechanics. The construction of this interpretation begins with the Feynman path integral formalism, in which QCD observables can be calculated from vacuum expectation values of any general operator $\mathcal{O}[\psi, \bar{\psi}, A_\mu]$ according to

$$\langle \Omega | \mathcal{O}[\psi, \bar{\psi}, A_\mu] | \Omega \rangle = \frac{1}{\mathcal{Z}} \int \mathcal{D}\psi \mathcal{D}\bar{\psi} \mathcal{D}A_\mu \mathcal{O}[\psi, \bar{\psi}, A] \exp\left(i\mathcal{S}_{\text{QCD}}[\psi, \bar{\psi}, A_\mu]\right), \tag{5.39}$$

where Ω is used to represent the QCD vacuum, and we integrate over all possible field values for the quark, anti-quark, and gluon fields. The generating functional $\mathcal{Z}$ is written

$$\mathcal{Z} = \int \mathcal{D}\psi \mathcal{D}\bar{\psi} \mathcal{D}A_\mu \exp\left(i\mathcal{S}_{\mathrm{QCD}}[\psi, \bar{\psi}, A_\mu]\right). \tag{5.40}$$

While this has the same form as the partition function in statistical mechanics, the exponent in $\mathcal{Z}$ is imaginary here rather than real. We can formulate the vacuum expectation value with a real exponent by performing a Wick rotation from Minkowski spacetime into Euclidean spacetime, changing the coordinates such that

$$x^0 \to -ix_0, \tag{5.41}$$

$$A^0 \to +iA_0, \tag{5.42}$$

$$\mathcal{S}_{\mathrm{QCD}} \to +i\mathcal{S}_{\mathrm{QCD}}^{\mathrm{Eucl.}}. \tag{5.43}$$

This Euclidean spacetime does not distinguish between covariant and contravariant indices.

The calculation of the vacuum expectation value as a weighted average across all possible configurations of the quark and gluon fields can be approximated by an average across a finite set of field configurations selected according to their expected weight. If an ensemble of N field configurations is chosen so that each configuration j has a probability P with

$$P[\psi^j, \bar{\psi}^j, A_\mu^j] = \exp\left(-\mathcal{S}_{\mathrm{QCD}}^{\mathrm{Eucl.}}[\psi^j, \bar{\psi}^j, A_\mu^j]\right), \tag{5.44}$$

the operator expectation value becomes

$$\langle\Omega|\mathcal{O}[\psi^j, \bar{\psi}^j, A_\mu^j]|\Omega\rangle \approx \frac{1}{N}\sum_{j=1}^{N} \mathcal{O}[\psi^j, \bar{\psi}^j, A_\mu^j]. \tag{5.45}$$

As a brief aside before explaining discretisation and gauge invariance in lattice QCD, we wish to discuss the generation of ensembles of these field configurations distributed according to the action. This generation is performed using Hybrid Monte Carlo methods, where new gauge field configurations are proposed based on a previous configuration, and then this new configuration may be accepted or rejected with a probability P dependent on the action (see Equation 5.44) to produce an appropriately distributed sample. Typically, multiple updates to the configuration are performed before the accept/reject step. By repeating the process to produce new configurations sufficiently many times, we obtain a configuration that is independent of the initial conditions. This process—called 'thermalisation'—ensures the configurations we generate have the correct physics. By further updating and

accepting/rejecting configurations after thermalisation, we can collect many configurations to form an ensemble, and these configurations will be independent if the update step is performed sufficiently many times between each accept/reject stage.

5.2.1 Discretisation and gauge invariance

The discretisation from an infinite Euclidean space to a finite four-dimensional lattice is performed by transforming from continuous coordinates x^μ to discrete coordinates an^μ, where n^μ are integer coordinates for the vertices of the lattice in 4D space and a is called the lattice spacing. The total spacetime volume of the lattice V is often written as $V = (aN_s)^3 \times (aN_t)$, for N_s and N_t the number of lattice sites in the spatial and temporal extents of the lattice respectively.

Finite-volume errors in calculated lattice results are possible if the number of sites in the lattice is too small. The current rule of thumb in the field is that the product of the pion mass used (m_π) and the span of the lattice in the spatial dimension (L) should be greater than four [32,65]. Typically, pion masses larger than the physical pion mass are used in lattice calculations in order to reduce the computational expense, although recent computational techniques now allow calculations at the physical point with $m_\pi L > 3$.

When formulating QCD on the lattice, the fermion fields representing quarks and anti-quarks take values on the lattice sites, while gluons are defined on the links between sites in order to maintain gauge invariance. A diagram of the formulation of quark fields and gluon fields on the lattice is shown in Figure 5.4. The discretisation of spacetime requires that derivatives in QCD be considered as finite differences and integrals as finite sums:

$$\partial_\mu f(x) \rightarrow \frac{f(x + a\hat{\mu}) - f(x - a\hat{\mu})}{2a} \tag{5.46}$$

$$\int d^4x f(x) \rightarrow a^4 \sum_x f(x) \tag{5.47}$$

for $\hat{\mu}$ the unit vector along the axis of the lattice corresponding to the index μ, and x (now discrete) corresponding to a site on the lattice.

The dependence of derivatives upon nearest-neighbour differences means that the boundary conditions of the lattice must also be specified to ensure that derivatives can be defined everywhere. Most groups will select periodic boundary conditions for the lattice, such that $x + aN_\mu\hat{\mu} = x$ for N_μ the extent of the lattice along the μ axis.

Returning to our explanation of discretisation in lattice QCD, we must also check the behaviour of these finite-difference derivatives under a gauge transformation. When the (naive) derivative is applied to the fermion field, we find that under a
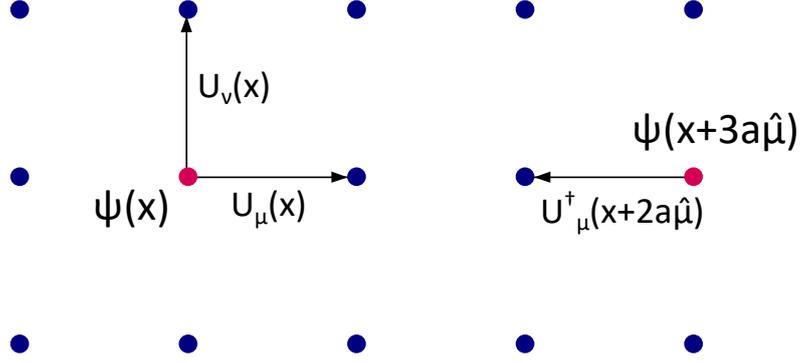
Figure 5.4: Visual representation of quarks (ψ) and gluon links (U) on the lattice. Arrows are used to clarify the source position associated with each gluon link.

gauge transformation $\Omega(x)$,

$$\delta_\mu \psi_j^a(x) \overset{\Omega(x)}{\to} \frac{1}{2a} \left(\Omega_{jk}^{ab}(x + a\hat{\mu})\psi_k^b(x + a\hat{\mu}) - \Omega_{jk}^{ab}(x - a\hat{\mu})\psi_k^b(x - a\hat{\mu}) \right). \tag{5.48}$$

This does not have the correct form to maintain gauge invariance. To restore the invariance, gluon fields are represented by the link variables

$$U_\mu(x) = \mathcal{P} \exp \left(ig \int_x^{x+a\hat{\mu}} dz \quad A_\mu(z) \right) \tag{5.49}$$

which take values on the links between adjacent lattice sites (see again Figure 5.4). The gluon link variables transform according to

$$U_\mu^{ab}(x) \overset{\Omega(x)}{\to} \Omega^{ac}(x)U_\mu^{cd}(x)\Omega^{\dagger db}(x + a\hat{\mu}) \tag{5.50}$$

$$U_\mu^{\dagger ab}(x) \overset{\Omega(x)}{\to} \Omega^{ac}(x + a\hat{\mu})U_\mu^{\dagger cd}(x)\Omega^{\dagger db}(x) \tag{5.51}$$

under local gauge transformations. Now that we have an appropriate definition for the gluons, we are able to define the covariant finite difference operator

$$\nabla_\mu \psi_j^a(x) = \frac{1}{2a} \left(U_\mu^{ab}(x)\psi_j^b(x + a\hat{\mu}) - U_\mu^{\dagger ab}(x - a\hat{\mu})\psi_j^b(x - a\hat{\mu}) \right) \tag{5.52}$$

which has the correct transformation

$$\nabla_\mu \psi_j^a(x) \overset{\Omega(x)}{\to} \Omega^{ab}(x)\nabla_\mu \psi_j^b(x) \tag{5.53}$$

and satisfies $\nabla_\mu \to D_\mu$ in the continuum limit $a \to 0$.

5.3 Quark actions

We seek a discretised quark action that takes the form of the Euclidean fermion action

$$\int d^4x \; \bar{\psi}_j^a(x) \left(\slashed{D}_{jk}^{ab}(x) + \delta^{ab}\delta_{jk}m \right) \psi_k^b(x). \tag{5.54}$$

Making a naive replacement of the covariant derivative with the finite difference operator ∇_μ yields

$$\mathcal{S}_F[\psi, \bar{\psi}, U_\mu] = a^4 \sum_x \bar{\psi}_j^a(x) \left(\slashed{\nabla}_{jk}^{ab}(x) + \delta^{ab}\delta_{jk}m \right) \psi_k^b(x) \tag{5.55}$$

which can also be written

$$\mathcal{S}_F[\psi, \bar{\psi}, U_\mu] = a^4 \sum_x \bar{\psi}_j^a(x) \left(\delta^{ab}\slashed{\partial}_{jk} + ig\slashed{A}_{jk}^{ab}(x) + \delta^{ab}\delta_{jk}m + \mathcal{O}(a^2) \right) \psi_k^b(x). \tag{5.56}$$

This naive action has errors of $\mathcal{O}(a^2)$ and preserves chiral symmetry, but unfortunately also produces fifteen additional unphysical quark species – called fermion doublers – as a result of the sinusoidal form for the finite difference operator in momentum space.

This problem can be handled in a number of different ways. One such approach is to use staggered fermions [66–68], which reduces the total number of fermions with a given flavour from sixteen to four. These four fermions can be identified by a new quantum number, called 'taste', which has additional symmetries connected to translations on the lattice. The benefit of staggered fermion actions is that they are often much faster to generate computationally, but the new taste symmetries can be violated by high momentum gluons, ultimately leading to large splittings between pion masses [69]. In more modern formulations, staggered fermions can be produced with only one taste by using a fourth-root method, and $\mathcal{O}(a^2)$ improvement at tree level is possible using the 'asqtad' action for lighter quarks (which also removes taste violations at high momentum transfer) or a HISQ action for heavier quarks [70].

The other method of removing these fermion doublers involves removing all the extra zeroes of the finite difference operator in momentum space so that only one pole remains in the $m \to 0$ limit of the propagator. This can be performed by including the Wilson term

$$\Delta\psi_j^a(x) = \frac{1}{a^2} \sum_{\mu=1}^{4} \Big(2\psi_j^a(x) - U_\mu^{ab}(x)\psi_j^b(x + a\hat{\mu}) \tag{5.57}$$
$$- U^{\dagger}{}_\mu^{ab}(x - a\hat{\mu})\psi_j^b(x - a\hat{\mu}) \Big)$$

so that the Wilson quark action can be written

$$\mathcal{S}_F^W[\psi, \bar{\psi}, U_\mu] = a^4 \sum_x \bar{\psi}_j^a(x) \left(\overset{\leftrightarrow}{\nabla}{}_{jk}^{ab}(x) + \delta^{ab}\delta_{jk}\frac{ra}{2}\Delta + \delta^{ab}\delta_{jk}m \right) \psi_k^b(x) \tag{5.58}$$

where the Wilson parameter r is usually set to 1. The Wilson fermion action can be further simplified to a form

$$\mathcal{S}_F^W[\psi, \bar{\psi}, U_\mu] = a^4 \sum_x \bar{\psi}_j^a M_W{}_{jk}^{ab}(x, y)\psi_k^b(x) \tag{5.59}$$

by rescaling the fermion fields using the hopping parameter κ as follows:

$$\psi_j^a(x) \rightarrow \frac{\psi_j^a(x)}{\sqrt{2\kappa}}, \; \kappa = \frac{1}{2ma + 8r}. \tag{5.60}$$

Then, the Wilson fermion matrix is written

$$M_W{}_{jk}^{ab}(x, y) = \delta_{xy} - \kappa \sum_\mu [(r - \gamma^\mu)U_\mu(x)\delta_{x+a\hat{\mu},y} + (r + \gamma^\mu)U_\mu^\dagger(x - a\hat{\mu})\delta_{x-a\hat{\mu},y}] \tag{5.61}$$

Despite successfully removing the fermion doublers, the Wilson action also introduces new discretisation errors at $\mathcal{O}(a)$ and breaks chiral symmetry. The new $\mathcal{O}(a)$ errors can be removed with the introduction of the clover term [71]

$$-\frac{igaC_{SW}r}{4}\sigma_{jk}^{\mu\nu}F_{\mu\nu}^{ab}(x) \tag{5.62}$$

for $\sigma_{\mu\nu}$ the Pauli matrices, C_{SW} the clover coefficient (which must be tuned to properly remove the $\mathcal{O}(a)$ errors), and

$$F_{\mu\nu}(x) = \frac{-i}{8a^2g} \left[(C_{\mu\nu}(x) - C_{\mu\nu}^\dagger(x)) - \frac{1}{3}\text{Tr}(C_{\mu\nu}(x) - C_{\mu\nu}^\dagger(x)) \right] \tag{5.63}$$

for clover loops $C_{\mu\nu}$, illustrated in Figure 5.5.

5.3.1 Light quark action in this thesis

In this thesis, we use the clover-Wilson action for all light quarks. The clover coefficient has already been tuned for the QCDSF lattice configurations used in this work [72, 73], and these parameters are shown in Table 5.1. The same clover coefficient is used for both light and strange quarks at each lattice spacing. Further information about the lattice ensembles used in this thesis is provided in Section 6.1.1.

Additional parameters for the lattice ensembles are also required. The β value
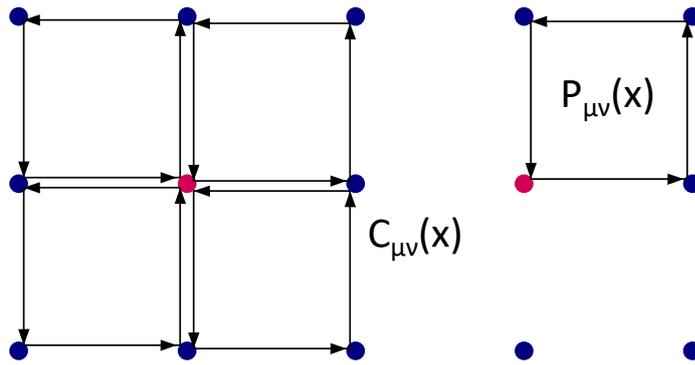
Figure 5.5: The clover loop $C_{\mu\nu}$ is constructed from 16 link variables, arranged as a sum of four plaquettes, defined $P_{\mu\nu} = U_\mu(x)U_\nu(x + a\hat{\mu})U_\mu^\dagger(x + a\hat{\nu})U_\nu^\dagger(x)$.

Table 5.1: Light quark parameters common to all QCDSF ensembles

β	Spacing a (fm)	Clover coefficient C_{SW}
5.4	0.082	2.79
5.5	0.074	2.65
5.65	0.068	2.48
5.8	0.059	2.34

is related to the coupling via the relation $\beta = 10/g^2$ for the improved gluon action employed here. The quark action used in generating the lattices also includes a single iteration of stout-link smearing of the gluon links in some terms (see [72]). Stout-smeared gluon links (or "fat links") include additional contributions from neighbouring links, such that the new stout link $\tilde{U}_\mu(x)$ is written

$$\tilde{U}_\mu(x) = \exp(iQ_\mu(x))U_\mu(x) \tag{5.64}$$

where

$$Q_\mu = \frac{\alpha}{2i}\left[V_\mu U_\mu^\dagger - U_\mu V_\mu^\dagger - \frac{1}{3}\text{Tr}(V_\mu U_\mu^\dagger - U_\mu V_\mu^\dagger)\right] \tag{5.65}$$

and V_μ is the sum of all 'staples' around U_μ, and each staple is a construction made of nearby gluon links (shown in Figure 5.6). The smearing value $\alpha = 0.1$ is used for these ensembles [72].

5.3.2 Heavy quark action

The b quark is significantly heavier than the u, d, s quarks typically calculated using the Clover-Wilson action. Unless the chosen lattice spacing is very fine, the Compton wavelength of the b is less than or similar in size to the lattice spacing, resulting in significant discretisation errors if a conventional light quark action is used for the b.
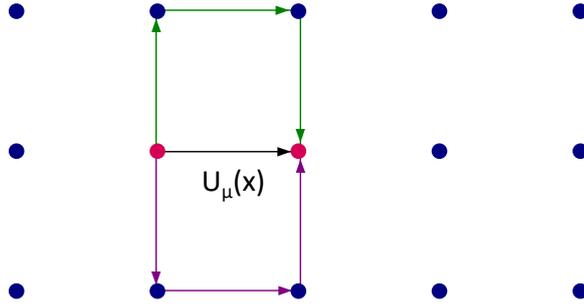
Figure 5.6: Visual representation of staples used in smearing. In this 2D projection, there are two staples for $U_\mu(x)$. In 4D, each gluon link $U_\mu(x)$ has a total of six staples.

Instead, we use the Fermilab action or RHQ action [74–76]:

$$
\mathcal{S}_{\text{heavy}} = a^4 \sum_{x,x'} \overline{\psi}(x') \bigg(m_0 + \gamma_0 \nabla_0 + \zeta \vec{\gamma} \cdot \vec{\nabla} - \frac{a}{2}(\nabla^0)^2
$$
$$
- \frac{a}{2}\zeta(\vec{\nabla})^2 + \sum_{\mu,\nu} \frac{ia}{4} c_P \sigma_{\mu\nu} F_{\mu\nu} \bigg)_{x,x'} \psi(x)
$$

(5.66)

where $\vec{\nabla}$ is the spatial three-vector component of ∇_μ, and similarly for $\vec{\gamma}$. This is a clover-improved Wilson action, with the addition of the ζ parameter to allow for anisotropy between the temporal and spatial axes on the lattice.

Multiple different methods may be used to tune the free parameters in this action. In the Fermilab/MILC collaboration, ζ is set to 1 and c_P is chosen based on a tree-level mean-field-improved lattice perturbation theory result [77], leaving only m_0 to be tuned. Instead, we follow the procedure outlined in [78], where the three free parameters for the b quark are tuned to physical observables of the B mesons. For this procedure, properties of the B and B^* mesons are combined, and the spin-average of the masses and the hyperfine splitting between them are chosen for tuning.

Aside from the physical B properties, the b quark tuning also depends on the dispersion relation of the calculated mesons. Due to lattice artefacts, the dispersion relation is written

$$
E^2 = M^2 + Ap^2 + \mathcal{O}(p^4)
$$

(5.67)

considering the energy E, mass M, and momentum p of the lattice states in physical units. The rest mass M is the energy of the state measured at $p = 0$, and the

dispersion coefficient A is calculated

$$A = \left. \frac{\partial(E^2)}{\partial(p^2)} \right|_{p^2=0} \tag{5.68}$$

This coefficient is also sometimes written $A = M_1/M_2$ for M_1 the rest mass and M_2 the kinetic mass of the state.

In this method of tuning, seven b quarks are generated in a tuning 'star' as shown in Figure 5.7. This star has a central estimate of the parameters, with some independent variation in each of m_0, c_P, and ζ, labelled $\Delta m_0, \Delta c_P, \Delta \zeta$ respectively in this work. The parameters are iteratively changed until a set (m_0, c_P, ζ) is found that matches the physical properties required for the B. The tuning algorithm will be discussed in further detail in Section 6.3 and beyond.
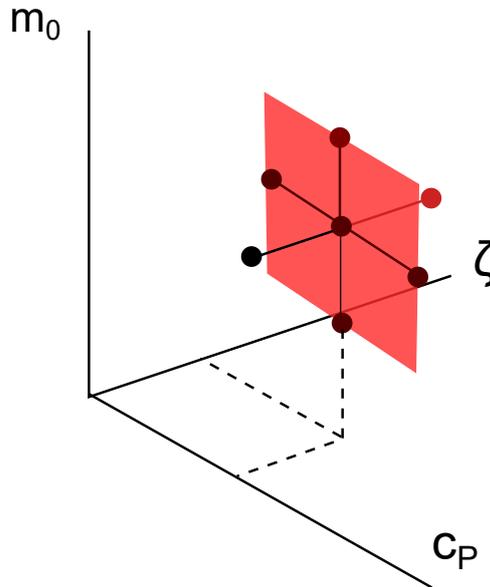
Figure 5.7: A diagram of the tuning 'star' of seven different sets of b quark properties. A plane is shown as a guide for the eye.

5.4 Measuring observables

5.4.1 Correlation functions

For hadronic observables, we need operators that can produce our states of interest on the lattice. Typically, we refer to the creation and annihilation operators as $\tilde{\chi}$ and χ respectively, such that a hadronic state created at x and annihilated at x' is written

$$\langle \Omega | \chi(x') \tilde{\chi}(x) | \Omega \rangle \, . \tag{5.69}$$

The Euclidean two point correlation function is then defined

$$C(x', x) = \langle \chi(x') \tilde{\chi}(x) \rangle . \tag{5.70}$$

These operators χ and $\tilde{\chi}$ couple to QCD eigenstates with specific quantum numbers, producing a 'tower' of states with these same quantum numbers but different energies. In the finite volume of lattice QCD, this tower is actually discrete rather than continuous, and the complete set of states is finite. If we label each state by $|X(\mathbf{p})\rangle$ for the 3-momentum $\mathbf{p}$, these eigenstates have properties

$$H |X(\mathbf{p})\rangle = E_X(\mathbf{p}) |X(\mathbf{p})\rangle \tag{5.71}$$

$$\langle X(\mathbf{p})|Y(\mathbf{q})\rangle = 2E_X(\mathbf{p})(2\pi)^3 \delta_{XY} \delta^3(\mathbf{p} - \mathbf{q}) \tag{5.72}$$

$$\sum_{X,\mathbf{p}} \frac{\Delta^3 p}{(2\pi)^3} \frac{1}{2E_X(\mathbf{p})} |X(\mathbf{p})\rangle \langle X(\mathbf{p})| = 1 \tag{5.73}$$

using the Hamiltonian operator H and where the last property, the completeness relation, uses $\Delta^3 p$ to represent the discrete equivalent of the continuous infinitesimals $d^3 p$. In the finite volume, these elements $\Delta^3 p = (2\pi)^3/V$.

The Euclidean spacetime of the lattice means that a general, translationally-invariant operator $\mathcal{O}$ has the property

$$\mathcal{O}(x) = e^{-i\mathbf{P}\cdot\mathbf{x}} e^{Ht} \mathcal{O}(0) e^{-Ht} e^{i\mathbf{P}\cdot\mathbf{x}} \tag{5.74}$$

using the 3-momentum operator $\mathbf{P}$. This allows us to write the two-point correlation function $C(x', x)$ in terms of the energy and momentum eigenstates:

$$C(x', x) = \sum_{X,\mathbf{k}} \frac{\Delta^3 k}{(2\pi)^3} \frac{e^{-E_X(\mathbf{k})(t'-t)}}{2E_X(\mathbf{k})} e^{i\mathbf{k}\cdot(\mathbf{x}'-\mathbf{x})} \langle \Omega|\chi(0)|X(\mathbf{k})\rangle \langle X(\mathbf{k})|\tilde{\chi}(0)|\Omega\rangle . \tag{5.75}$$

We now consider the Fourier projection of the correlation function, defined

$$\mathcal{C}_2(\mathbf{p}; t', t) = \sum_{\mathbf{x}'} \Delta^3 x' e^{-i\mathbf{p}\cdot(\mathbf{x}'-\mathbf{x})} C(x', x) \tag{5.76}$$

$$= \sum_{\mathbf{x}'} \Delta^3 x' e^{-i\mathbf{p}\cdot(\mathbf{x}')} C(x', 0) \tag{5.77}$$

when we use translational invariance to shift the source x to 0 and relabel x'. Then we can insert the form of the two point correlation function written in terms of the momentum eigenstates (Equation 5.75) and use

$$\sum_{\mathbf{x}'} \Delta^3 x' \; e^{i(\mathbf{p}'-\mathbf{p})\cdot(\mathbf{x}')} = (2\pi)^3 \; \delta^3(\mathbf{p}' - \mathbf{p}) \tag{5.78}$$

(where again, the source x is translated to 0) to obtain the result

$$\mathcal{C}_2(\mathbf{p}; \Delta t) = \sum_X \frac{e^{-E_X(\mathbf{p})\Delta t}}{2E_X(\mathbf{p})} \langle \Omega | \chi(0) | X(\mathbf{p}) \rangle \langle X(\mathbf{p}) | \tilde{\chi}(0) | \Omega \rangle \tag{5.79}$$

where Δt is $t' - t$.

This exponential time dependence means that the energies of individual hadronic states on the lattice may be extracted from this two point correlator by making fits to its time evolution. The contribution to the correlation function from higher-energy states exponentially decays more quickly with respect to time, meaning that lower-energy states and indeed the ground state can be extracted using fits at later times. In the limit as Δt becomes large, we expect the exponential of the ground state to dominate the sum, such that

$$\mathcal{C}_2(\mathbf{p}; \Delta t) \to \frac{e^{-E_{X_0}(\mathbf{p})\Delta t}}{2E_{X_0}(\mathbf{p})} \sum_\alpha \langle \Omega | \chi(0) | X_0(\mathbf{p}, \alpha) \rangle \langle X_0(\mathbf{p}, \alpha) | \tilde{\chi}(0) | \Omega \rangle \tag{5.80}$$

where we use X_0 to represent the lowest energy state, with α to label the degenerate eigenstates (if any). If the time is sufficiently large to isolate the ground state, we expect that we will be able to fit a single exponential to $\mathcal{C}_2$ in order to find the ground state's energy, or indeed its mass if $\mathbf{p} = 0$. To ensure that we have captured the right region, we need some way to check that the exponent is behaving like a single exponential with respect to time. We define the effective mass

$$m_{\text{Eff}}(\Delta t + \tfrac{a}{2}) = \frac{1}{a} \ln \left(\frac{\mathcal{C}_2(\vec{0}; \Delta t)}{\mathcal{C}_2(\vec{0}; \Delta t + a)} \right), \tag{5.81}$$

which will reach a plateau in the region where the ground state is dominant. The effective mass is used throughout this thesis to display the two point correlator and show the fit regions used to determine the mass of $B_{(s)}$ and $B^*_{(s)}$ mesons. This discussion of correlator fits used in this thesis begins in Section 6.2.

Our b quark tuning process will also require the value of the hyperfine splitting $M_{B^*} - M_B$. While this can be computed as the difference of the individual mass values for each state, this will often have greater uncertainty given that the expected value of the splitting is so small compared to the size of the masses. Instead, the ratio

$$\frac{\exp(E_2 t)}{\exp(E_1 t)} = \exp((E_2 - E_1)t) \tag{5.82}$$

implies that taking the ratio of two point correlators for B^* and B will allow us to extract $M_{B^*} - M_B$ from $\mathcal{C}_{2,B^*}/\mathcal{C}_{2,B}$ directly.

So far, we have only considered mass and energy observables related to individual

hadronic states. Calculation of more general matrix elements may require probing the hadron with an external current of some kind to allow a change of momentum (or of quark flavour) between the source and sink. To do this, we need to consider a third spacetime location for the insertion of the current. Then, this new three-point correlation function can be defined similarly to the two point function. We begin with an expectation value including the creation and annihilation operators, but the additional (local) current operator is inserted at some additional spacetime point y:

$$C(x', y, x) = \langle \chi(x') \mathcal{O}(y) \tilde{\chi}(x) \rangle \tag{5.83}$$

also written

$$C(x', y, x) = \sum_{\substack{X, \mathbf{p} \\ Y, \mathbf{q}}} \iint \frac{\Delta^3 k}{(2\pi^3)} \frac{\Delta^3 \ell}{(2\pi^3)} \left(\frac{e^{-E_X(\mathbf{k})(t'-\tau)}}{2E_X(\mathbf{k})} \frac{e^{-E_Y(\ell)(\tau-t)}}{2E_Y(\ell)} e^{i\mathbf{k}\cdot(\mathbf{x}'-\mathbf{y})} e^{i\ell\cdot(\mathbf{y}'-\mathbf{x})} \right.$$

$$\left. \times \langle \Omega | \chi(0) | X(\mathbf{k}) \rangle \langle X(\mathbf{k}) | \mathcal{O}(0) | Y(\ell) \rangle \langle Y(\ell) | \tilde{\chi}(0) | \Omega \rangle \right) \tag{5.84}$$

after inserting the full tower of states, and where τ is the time corresponding to the spacetime y. In a similar process to the two point function, we define the Fourier-projected three point correlation function

$$\mathcal{C}_3(\mathbf{p}', \mathbf{p}; t', \tau, t) = \sum_{\mathbf{x}', \mathbf{y}'} \Delta^3 x' \Delta^3 y \, e^{i\mathbf{p}'\cdot(\mathbf{x}'-\mathbf{y})} e^{i\mathbf{p}'\cdot(\mathbf{y}-\mathbf{x})} C(x', y, x) \tag{5.85}$$

$$= \sum_{X,Y} \frac{e^{-E_X(\mathbf{p}')(t'-\tau)}}{2E_X(\mathbf{p}')} \frac{e^{-E_Y(\mathbf{p})(\tau-t)}}{2E_Y(\mathbf{p})}$$

$$\times \langle \Omega | \chi(0) | X(\mathbf{p}') \rangle \langle X(\mathbf{p}') | \mathcal{O}(0) | Y(\mathbf{p}) \rangle \langle Y(\mathbf{p}) | \tilde{\chi}(0) | \Omega \rangle \tag{5.86}$$

using Equation 5.84. Much like Equation 5.79, we can also rewrite this three point correlation function to use new variables, selecting a new t equal to $\tau - t_{\text{source}}$ and choosing t_{fixed} is $t_{\text{sink}} - t_{\text{source}}$ ($t' - t$ in the previous notation). Then,

$$\mathcal{C}_3(\mathbf{p}', \mathbf{p}; t, t_{\text{fixed}}) = \sum_{X,Y} \frac{e^{-E_X(\mathbf{p}')(t_{\text{fixed}}-t)}}{2E_X(\mathbf{p}')} \frac{e^{-E_Y(\mathbf{p})t}}{2E_Y(\mathbf{p})}$$

$$\times \langle \Omega | \chi(0) | X(\mathbf{p}') \rangle \langle X(\mathbf{p}') | \mathcal{O}(0) | Y(\mathbf{p}) \rangle \langle Y(\mathbf{p}) | \tilde{\chi}(0) | \Omega \rangle . \tag{5.87}$$

If we consider the three point function sufficiently far from the source and sink times (that is, at large t and $t_{\text{fixed}} - t$), the lowest energy state for the coupling dominates.

Ratios of the three- and two-point functions can be used to eliminate the time dependence of the lowest-lying state, leaving only the matrix element values which

can be fit with a constant provided we remain sufficiently far from the source and sink times. The general three point function ratio has form

$$R(\mathbf{p}', \mathbf{p}; t, t_{\text{fixed}},) = \frac{\mathcal{C}_3(\mathbf{p}', \mathbf{p}; t, t_{\text{fixed}},)}{\mathcal{C}_2(t_{\text{fixed}}, \mathbf{p})} \times \sqrt{\frac{\mathcal{C}_2(t, \mathbf{p})\mathcal{C}_2(t_{\text{fixed}}, \mathbf{p})\mathcal{C}_2(t_{\text{fixed}} - t, \mathbf{p}')}{\mathcal{C}_2(t, \mathbf{p}')\mathcal{C}_2(t_{\text{fixed}}, \mathbf{p}')\mathcal{C}_2(t_{\text{fixed}} - t, \mathbf{p})}} \quad (5.88)$$

where $\mathbf{p}'$ is the sink momentum and $\mathbf{p}$ is the source momentum.

5.4.2 Choosing operators

In this thesis, we consider B (pseudoscalar) and B^* (vector) mesons. The creation and annihilation operators for these states are

$$\chi_P(x) = \bar{b}(x)\gamma_5 q(x) \qquad\qquad \tilde{\chi}_P(x) = -\bar{q}(x)\gamma_5 b(x) \qquad (5.89)$$

$$[\chi_V]_\mu(x) = \bar{b}(x)\gamma_\mu q(x) \qquad\qquad [\tilde{\chi}_V]_\mu(x) = \bar{q}(x)\gamma_\mu b(x) \qquad (5.90)$$

where $b(x)$ and $q(x)$ represent the fermion fields $\psi(x)$ corresponding to the b and q valence quark flavours respectively. The B^* states are typically calculated using γ_i. We notice that for the meson states without any transitions, $\tilde{\chi} = \chi^\dagger$ and thus $\langle\Omega|\chi(0)|X(\mathbf{p})\rangle$ and $\langle X(\mathbf{p}')|\chi(0)|\Omega\rangle$ are complex conjugates.

In our discussion of correlation functions so far (see Section 5.4.1), we had only considered the forward-propagating meson states. Given that the lattice is periodic, however, backward-propagating states may travel around the lattice and interact with the forward-propagating state. When we include this effect in a lattice of extent T, as Δt and $T - \Delta t$ become sufficiently large we expect the ground state correlation function to have a form

$$\mathcal{C}_2(\mathbf{p}; \Delta t) \to \frac{|\langle\Omega|\chi_0|X_0(\mathbf{p})\rangle|^2}{2E_{X_0}(\mathbf{p})} e^{-E_{X_0}(\mathbf{p})\frac{T}{2}} \cosh\left(E_{X_0}(\mathbf{p})(\Delta t - \tfrac{T}{2})\right), \qquad (5.91)$$

for the mesons of interest.

So far, we have only considered creation and annihilation operators confined to a single lattice site. In practice, extraction of matrix elements relies on having operators with a sufficiently large coupling to the ground state to overcome noise. One way that the coupling can be improved is by adding smearing to the operators so that the hadronic states occupy multiple lattice sites and thus have physical extent. This smearing is usually introduced as part of the inversion of the Dirac or fermion matrix (for example M_W in Equations 5.59 and 5.61) in the process of calculating the quark propagator S_f. Without smearing, the Dirac matrix is inverted in 'columns'

according to the system of equations

$$\sum_z [M_W(x,z)]^{ac}_{\alpha\gamma} [S_f(z,y_0)]^{cb_0}_{\gamma\beta_0} = \delta_{\alpha\beta_0} \delta^{ab_0} \delta(x - y_0) \tag{5.92}$$

where the spin, colour, and spatial indices (β_0, b_0, y_0) are fixed during each calculation. While this matrix M has $N_s^3 \times N_t \times N_c \times N_D$ components for $N_c = 3$ the number of colours and $N_D = 4$ the number of Dirac spin components, by using a single fixed spacetime location y_0 for the source or sink, we only need to invert this equation twelve times: once for each combination of colour and spin indices. Source smearing is then applied by altering the 'source vector' in this inversion from a product of delta functions to a new vector η which has connections between nearby lattice sites and is produced by successive applications of a specialised smearing function.

5.4.3 Building the decay constant

The decay constant f_{B_q} is defined in terms of the matrix element

$$f_{B_q} p_\mu = \langle \Omega | [\chi_A]_\mu | B_q \rangle \tag{5.93}$$

where $[\chi_A]_\mu = \bar{b}(x) \gamma_\mu \gamma_5 q(x)$ is the axial vector current corresponding to the heavy-light change of flavour, and B_q is the ground state. To simplify, we can consider the B_q meson at rest so that

$$f_{B_q} M_{B_q} = \langle \Omega | [\chi_A]_0 | B_q \rangle . \tag{5.94}$$

In practice, the axial current is often represented by an expansion of simpler currents and correction terms to reduce lattice discretisation effects. We can define the decay constant as an expansion

$$f_B = \frac{1}{a} Z_\Phi \left[\Phi_B^0 + c_A \Phi_B^1 \right] \tag{5.95}$$

where the renormalisation Z_Φ for the axial current is calculated [79]

$$Z_\Phi = \rho_A^{bl} \sqrt{Z_V^{bb} Z_V^{ll}} \tag{5.96}$$

and the terms Φ_B^0 and Φ_B^1 are each calculated with different parts of the expansion of the axial vector current. In this work, the perturbative constant ρ_A^{bl} is set to 1 and the higher-order correction coefficient c_A in f_B is set to zero (and thus the decay constant correction Φ_B^1 is not calculated). Z_V^{bb} and Z_V^{ll} are calculated from three point functions of the vector current, and the calculation of these values is discussed further in Section 6.4.1.

The value Φ_B^0 is defined

$$\Phi_B = -\frac{\sqrt{2M_B}\mathcal{C}_{AP}}{\sqrt{\mathcal{C}_{PP}}} \tag{5.97}$$

in terms of the two point correlator amplitudes

$$\mathcal{C}_{AP} = \frac{\langle\Omega|[\chi_A]_0|B_q\rangle\,\langle B_q|\chi_P|\Omega\rangle}{2M_B}, \tag{5.98}$$

$$\mathcal{C}_{PP} = \frac{\langle\Omega|\chi_P|B_q\rangle\,\langle B_q|\chi_P|\Omega\rangle}{2M_B}. \tag{5.99}$$

This allows us to isolate the required matrix element $\langle\Omega|[\chi_A]_0|B_q\rangle$ in the ratio of amplitudes, given that the overlap amplitudes of $\langle\Omega|\chi_P|B_q\rangle$ and $\langle B_q|\chi_P|\Omega\rangle$ have the same magnitude. The advantage of using the ratio with the pseudoscalar matrix elements is that smeared sources and/or sinks can be used in the calculation to increase the overlap of the interpolating operators with the expected physical ground state. Due to the nature of the decay to a weak boson, a point sink must be used for the axial interpolator. As an added bonus of using the pseudoscalar elements, we notice that $\mathcal{C}_{PP}$ is just the amplitude of the two point function used to calculate the matching B_q mass.

The calculation of f_B and f_{B_s}—and the masses of the B and B^* mesons—is discussed in the context of the b quark tuning beginning in Section 6.4.2.

5.5   Summary

In this Chapter, we have discussed the theory of the Standard Model with a focus on weak decay and flavour effects. QCD is discretised onto the lattice, and an overview of the generation of lattice configurations and the production of correlators used to measure hadronic observables is presented. The specific quark actions used in this thesis are also discussed.

Chapter 6

B mesons on the lattice

In order to construct B mesons on the lattice, we must tune the free parameters in the heavy quark action presented in Section 5.3.2. This b quark tuning process forms a foundation on which various lattice studies of the b quark—including calculations of the decay constant—can be made. In this Chapter, we discuss the implementation of the b quark tuning and the construction of B mesons on the lattice.

Details of the entire process of calculating the decay constants are important to give context to our SU(3)-breaking results for the decay constant, presented in Chapter 7. The controlled-SU(3)-breaking lattice ensembles used in this thesis are described, and then the full tuning process—including fitting to meson correlators—is presented. The renormalisation of the decay constant and possible sources of systematic error are briefly discussed, to be expanded upon in further detail in Chapter 7.

6.1   Light quarks and SU(3) symmetry

The QCDSF collaboration focuses on SU(3) symmetry breaking effects in lattice calculations. Unlike most lattice collaborations, both the light and strange quark masses are changed as we approach the physical m_u, m_d, and m_s from an SU(3) symmetric point, instead of the conventional approach where m_s is held constant at its physical value. More specifically, the average quark mass

$$\overline{m} = \frac{1}{3}(2m_l + m_s) \tag{6.1}$$

is tuned to its physical value and kept constant across these ensembles. As the light and strange quark masses move away from the SU(3) symmetric point toward their physical values, a natural SU(3) breaking parameter presents itself in the form of

$$\delta m_q = m_q - \overline{m}. \tag{6.2}$$

Since $\overline{m}$ is held constant as the individual quark masses change,

$$\delta m_u + \delta m_d + \delta m_s = 0 \tag{6.3}$$

and given $m_u = m_d = m_l$ for the mass-degenerate light quarks, we can write

$$\delta m_s = -2\delta m_l. \tag{6.4}$$

on each ensemble. As part of this approach, all quantities containing equal numbers of u, d, s quarks—flavour-singlet quantities—will only be affected by the SU(3) flavour breaking effects at $\mathcal{O}((\delta m_q)^2)$, and have been shown to stay approximately constant from the SU(3) symmetric point to the physical point [72]. A diagram comparing this behaviour to the conventional approach is shown in Figure 6.1. The x-axis is scaled with the flavour-singlet combination of light pseudoscalar meson masses, $X_\pi^2 = \frac{1}{3}(2m_K^2 + m_\pi^2)$. In the context of lattice results, we will continue to denote flavour-averaged quantities using the letter X.
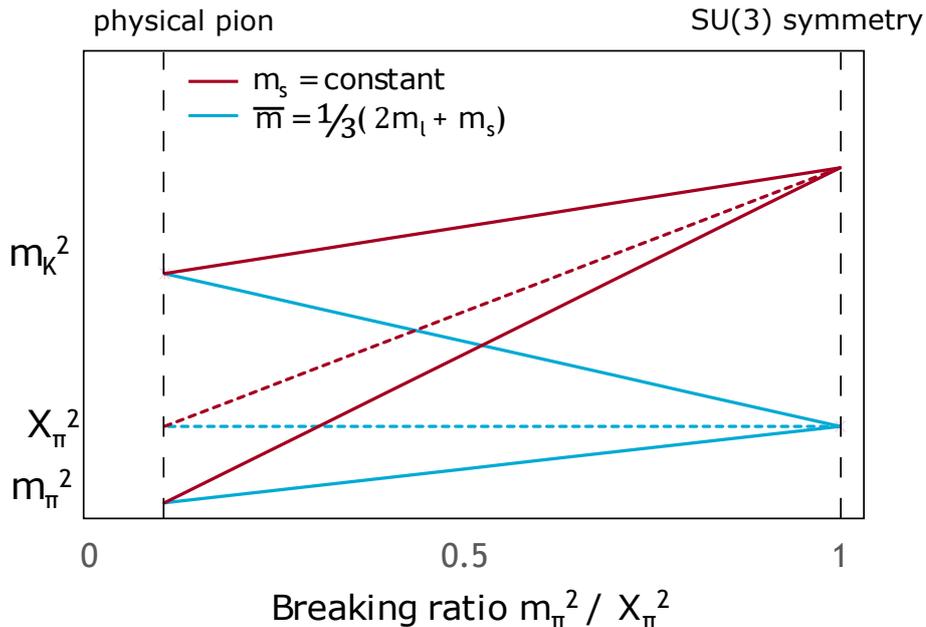
Figure 6.1: Diagram of the evolution of m_π^2 and m_K^2 for constant strange quark mass m_s (red) or constant average quark mass (blue). The coloured dashed lines show the evolution of the flavour singlet $X_\pi^2 = \frac{1}{3}(2m_K^2 + m_\pi^2)$

6.1.1 Ensembles used in this work

A variety of QCDSF $N_f = 2 + 1$ gauge field ensembles are used in this work. The ensemble lattice spacings and pion masses are represented graphically in Figure 6.2 for ease of comparison. Two different colours are used for lattices with spacing $a = 0.074$, as we have two different trajectories toward the physical masses of the

pion and kaon. In particular, the dark blue trajectory "misses" the physical point, and thus we must be careful with its analysis, in particular for quantities sensitive to the singlet quark mass such as the meson masses or decay constants. The trajectories for all sets of ensembles are shown in Figure 6.3. Throughout this thesis, the same colour will be used to represent the same set of ensembles in each graph, unless otherwise specified.
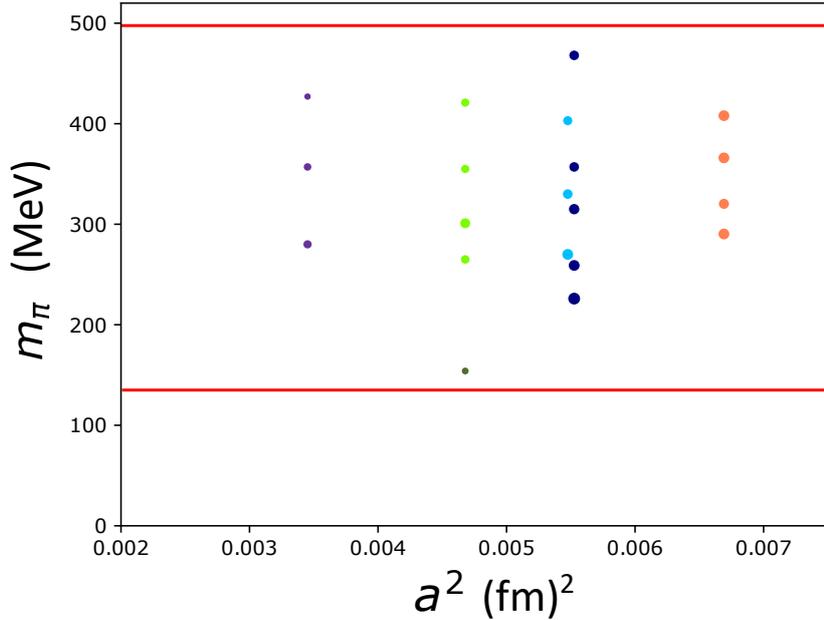
Figure 6.2: Lattice ensembles used in this work, distinguished by lattice spacing a and pion mass m_π. The red lines denote the physical pion and physical kaon masses for comparison, while the continuum limit occurs as $a \to 0$. The size of the spot for each ensemble indicates the number of lattice configurations used, detailed in Table 6.1. The same colours will be used for each set of ensembles throughout this thesis.

Further information about these ensembles is presented in Table 6.1. Valence quark correlators are generated using CHROMA [80]. For most results, the source locations for the calculated mesons are randomised to reduce correlations between neighbouring configurations in the ensemble. Due to the smaller number of configurations in the near-physical-pion 64^3 ensemble, however, multiple sources are used on each lattice to provide enough statistical samples for analysis.

Table 6.1: Table of lattice ensembles used in this work. * indicates ensembles with a different value of $\overline{m}$, further from the physical $\overline{m}$. † indicates ensembles where multiple sources per configuration are used to produce additional samples. Marked ensembles use 2 randomised sources, except for the $64^3 \times 96$ sample with 4 randomised sources used.

β	a (fm)	Lattice volume	# Samples	$(\kappa_{\text{light}}, \kappa_{\text{strange}})$	m_π (MeV)	m_K (MeV)	
5.4	0.082	$32^3 \times 64$	1015	(0.11993 , 0.11993)	408	408	
			1004	(0.119989 , 0.119812)	366	424	
			877	(0.120048 , 0.119695)	320	440	
			1006	(0.120084 , 0.119623)	290	450	
5.5	0.074	$32^3 \times 64$	677†	(0.12095 , 0.12095)	403	403	
			786	(0.12104 , 0.12077)	331	435	
			1021	(0.121099 , 0.120653)	270	454	
		$48^3 \times 96$	778	(0.1209 , 0.1209)	468	468	*
			758	(0.12104 , 0.12062)	357	505	*
			902†	(0.121095 , 0.120512)	315	526	*
			1002	(0.121145 , 0.120413)	258	537	*
			1251†	(0.121166 , 0.120371)	226	539	*
5.65	0.068	$48^3 \times 96$	500	(0.122005 , 0.122005)	412	412	
			500	(0.122078 , 0.121859)	355	441	
			845†	(0.12213 , 0.121756)	302	457	
			576	(0.122167 , 0.121682)	265	474	
		$64^3 \times 96$	320†	(0.122227 , 0.121563)	155	480	
5.8	0.059	$64^3 \times 96$	298	(0.12281 , 0.12281)	427	427	
		$48^3 \times 96$	415	(0.12288 , 0.12267)	357	456	
			525	(0.12294 , 0.122551)	280	477	
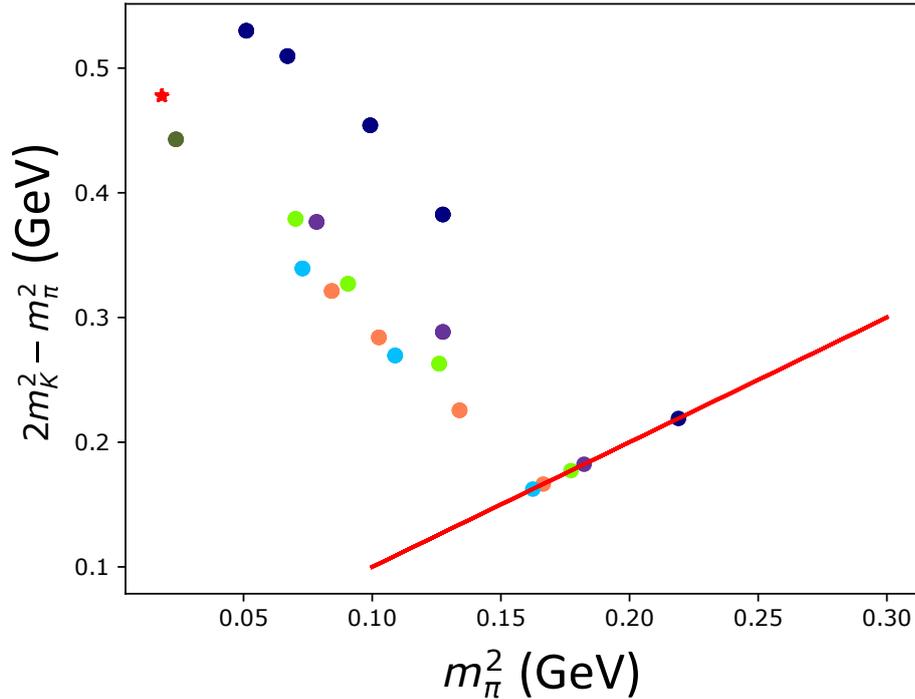
Figure 6.3: SU(3) breaking "trajectories" for each set of lattice ensembles, shown in the quark mass plane. m_π^2 is proportional to m_l, while $2m_K^2 - m_\pi^2$ is proportional to m_s. The SU(3) symmetric line is shown in red, with the physical pion and kaon masses represented by the red star. The same colours will be used for each set of ensembles throughout this thesis.

6.2 Correlator fits for tuning

The tuning of the best b using our heavy quark action (see Section 5.3.2), and indeed the selection of the b quark tuning 'star' (see Figure 5.7) for each set of lattice ensembles is dependent on the correlator fits used. In this Section, we discuss the fitting methods used to produce the results seen in this thesis and suggest fitting approaches that could be used in future work. The tuning method is discussed in detail in Section 6.3.

6.2.1 Default fits for tuning

The tuning process for the b quark requires multiple correlator fits, and in fact we require fits for 6 different correlator types, seven b quarks, and up to 2 light quarks for a single ensemble. Despite the known cosh functional form for two-point meson correlators (see Section 5.4.2), by default we use single exponential fits to avoid the noisy region in the centre of the lattice.

The χ^2 per degree of freedom can be used in assessment of the fit. Fits can be calculated using either the correlated or uncorrelated χ^2 as part of the minimisation.
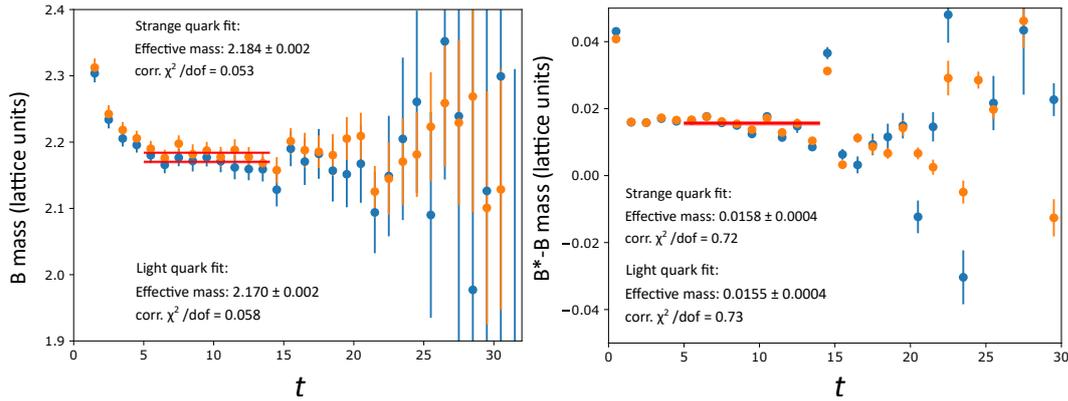
Figure 6.4: Left: An example correlator fit to the *B* meson mass, from the central *b* quark of the $\beta = 5.4$, (κ_l, κ_s)=(0.120048,0.119695) ensemble. Right: An example correlator fit to the ratio of the B^* and B correlators to produce the hyperfine splitting, shown for the $\beta = 5.65$, (κ_l, κ_s)=(0.122130,0.121756) ensemble.

The uncorrelated χ^2 per degree of freedom is defined

$$\chi^2 = \sum_i \frac{(Y_i - f(X_i))^2}{\nu \sigma_i^2} \tag{6.5}$$

for f our fit function, and each (X, Y) representing $(t, \mathcal{C})$ pairs in the case of our correlator fits. The parameter ν is the number of degrees of freedom, defined as the number of (X, Y) pairs minus the number of free parameters of f. In the case of the correlated χ^2, the inverse of the correlation matrix is inserted and χ^2 may be written $\chi^2 = r^T W r$ for the vector $r = (Y - f(X))$. While the correlation matrix can also be applied to the uncorrelated χ^2 minimised by a fit, in general this will not match the minimum correlated χ^2 calculated from a correlated fit.

Systematic errors in the final calculated values as a result of the fit windows chosen for the correlators are inevitable in any lattice work, but given that multiple *B* mesons must be fit for each ensemble, we expect this effect to be somewhat magnified. To keep our fitting methodology consistent in the default study, we use the same fit window for each of the seven *b* quarks. Similarly, we wish to use the same fit function for as many ensembles as possible.

The tuning process requires that the *B* and B^* masses and the hyperfine splitting are calculated from correlators. The energy of the *B* meson at different momenta is also required. As an example, effective mass plots for the *B* and for the ratio of correlators corresponding to $B^* - B$ are shown in Figure 6.4. We notice that an earlier fit window for the $B^* - B$ splitting could also be considered, the impact of this change in fit window is discussed in Section 7.3.1 and beyond.

A 3×3 grid of the correlator fits for the B_l and B_s meson masses for one $\beta = 5.5$ ensemble is shown in Figure 6.5. Each individual subplot shows the light and strange *B* meson correlators corresponding to a single *b* quark in the tuning star.

In practice, the central fit value and fit error are calculated from a minimisation of the uncorrelated χ^2 value, but the reported χ^2 is from a different fit using a minimisation of the correlated χ^2 on the same fit region. This is due to low fit stability using the correlated χ^2 in the minimisation, as the correlation matrix can be difficult to constrain. When measured, the correlation between neighbouring points in Euclidean time is usually around 70 %.
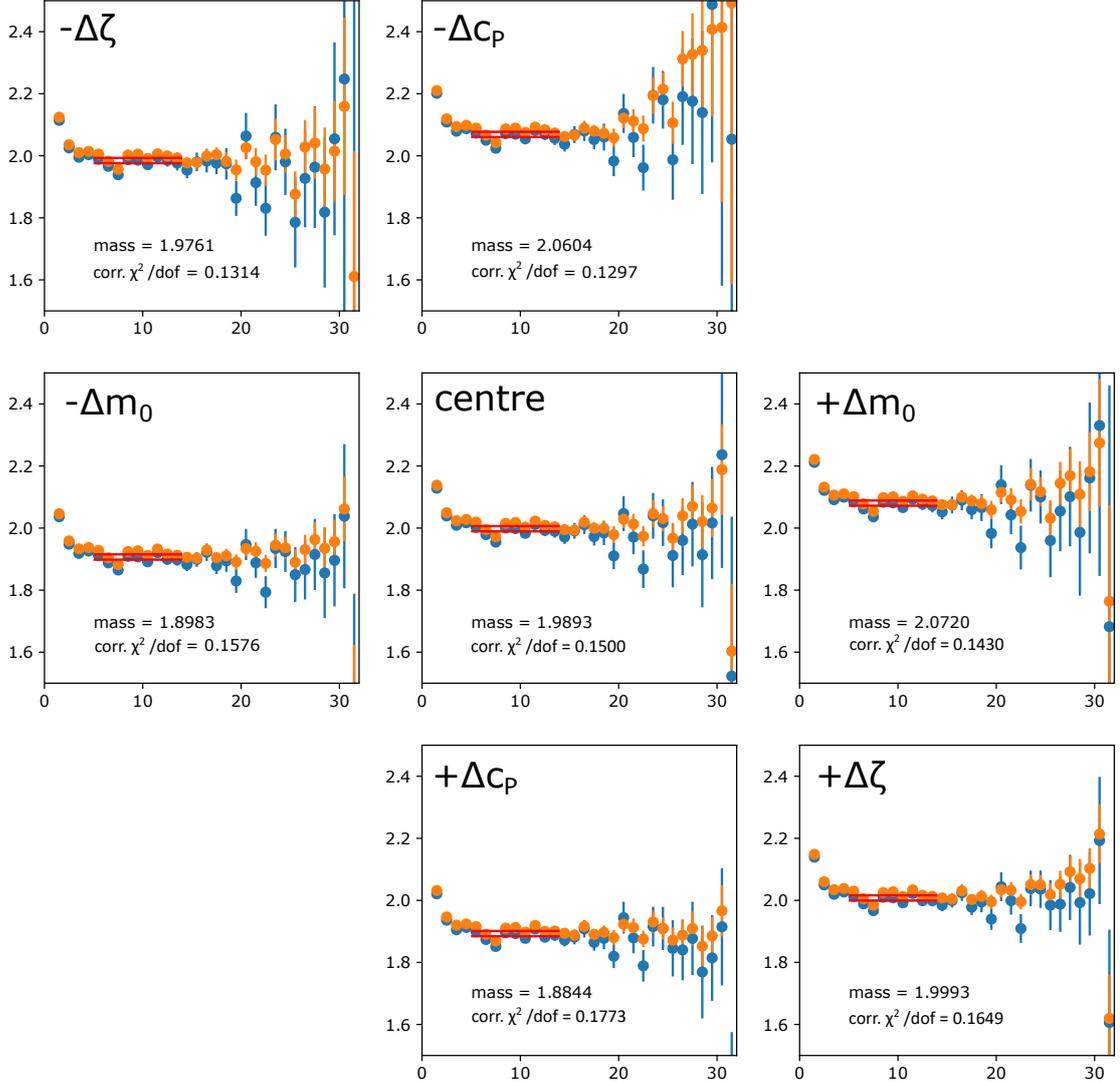
Figure 6.5: Correlator fits for the effective masses of the B_l (blue) and B_s (orange) mesons on the $(\kappa_l, \kappa_s)=(0.12104, 0.12062)$ ensemble, shown for all 7 b quark parameters in the tuning 'star'. The three free parameters for the tuning 'star' are discussed in Section 5.3.2.

We observe that the shape of the jitter, or vertical movement between neighbouring points in Euclidean time, is the same across all seven b quarks, until approximately $t = 25$, where the effective mass becomes too noisy. Some early excited-state contamination is also observed, but in general, the plateau region of the B states is very large, especially on ensembles close to the SU(3) symmetric point.

These fit windows for the single exponential fit were tested on the $\beta = 5.5$ ensembles, and found to be acceptable at all light and strange quark values. As a result, when we refer to the default fitting configuration we are referring to this same set of fit window parameters. Due to large size of the plateau region in these correlators, the same fit windows are also acceptable at other lattice spacings. The fit windows for each correlator type are listed in Table 6.2. We notice that the chosen plateau in the decay constant fit is longer than for the other fits, this is in part due to a fit failure of the $t = 5 - 14$ window for some bootstraps for the decay fits on some ensembles, making the $t = 5 - 14$ window unacceptable if we wish to use a consistent window on all samples. The impact of fit window choices on the decay constant results is explored in some detail in Section 7.3.1, while the default fit configuration for the decay constant is presented in Section 6.4.2.

Table 6.2: Default fit windows for correlators on all ensembles

Correlator type	Default fit window
B mass	5-14
B^* mass	5-14
B^*/B ratio	5-14
B at $p^2 = 1$	5-14
B at $p^2 = 2$	5-14
B at $p^2 = 3$	5-14
decay constant $\mathcal{C}_{AP}$	5-20
Z_V	5-14

6.2.2 Searching for new fits

While Figure 6.5 allows us to visualise the selected fit across all 7 b quarks, additional tools may be needed to assess the choice of different fits in a more systematic way.

The χ^2 and output fit value can be compared for multiple fit windows simultaneously by using colour to represent the value of a 2D distribution. An example of these plots is shown in Figure 6.6, where we present the B mass value and the χ^2/dof together for each fit. These fits are calculated using a minimisation of the uncorrelated χ^2.

We see that the extracted mass value is approximately constant around the 5-14 fit window we have used for our default fit. In order to keep the fit window consistent across all b quarks, we would need to check these plots for all 7 quarks simultaneously. Visualising the output fit value across multiple b quarks is not especially effective with a shared colour map because the differences in the masses of the B mesons with different b quarks are larger than the variations in fit output for each individual b. We could, however, compare the χ^2 using the same colour
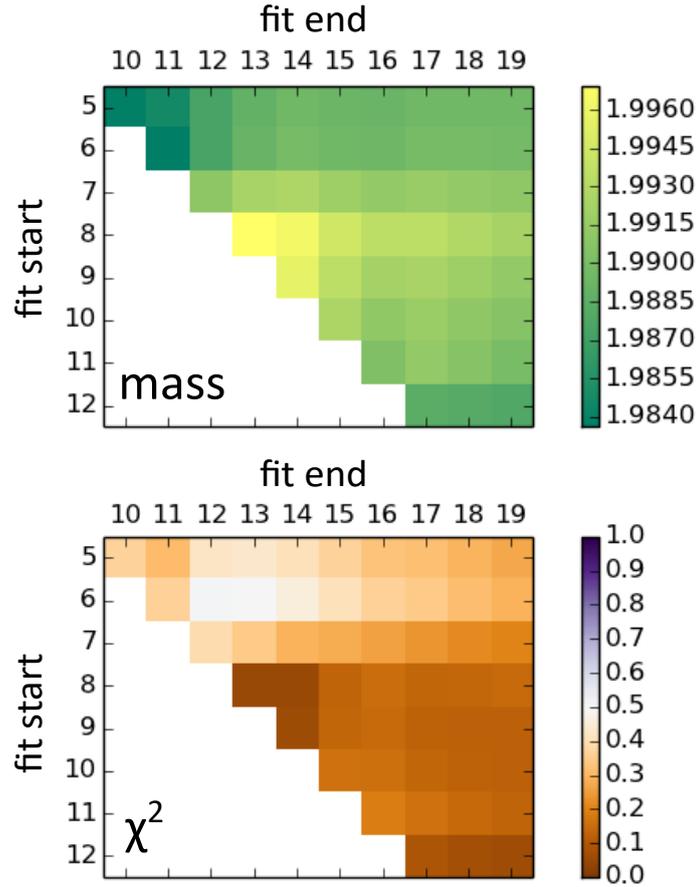
Figure 6.6: 2D grids of B_l meson mass (top) and χ^2/d.o.f (bottom) from single exponential fits. The fit start point is shown on the y-axis, with the fit end on the x-axis. These plots are produced for the central b quark on the (κ_l, κ_s)=(0.12104,0.12062) ensemble.

scale for all 7 *b* quarks, in a similar fashion to the effective mass visualisations in Figure 6.5. This is shown in Figure 6.7, again using the uncorrelated χ^2/d.o.f of each fit.
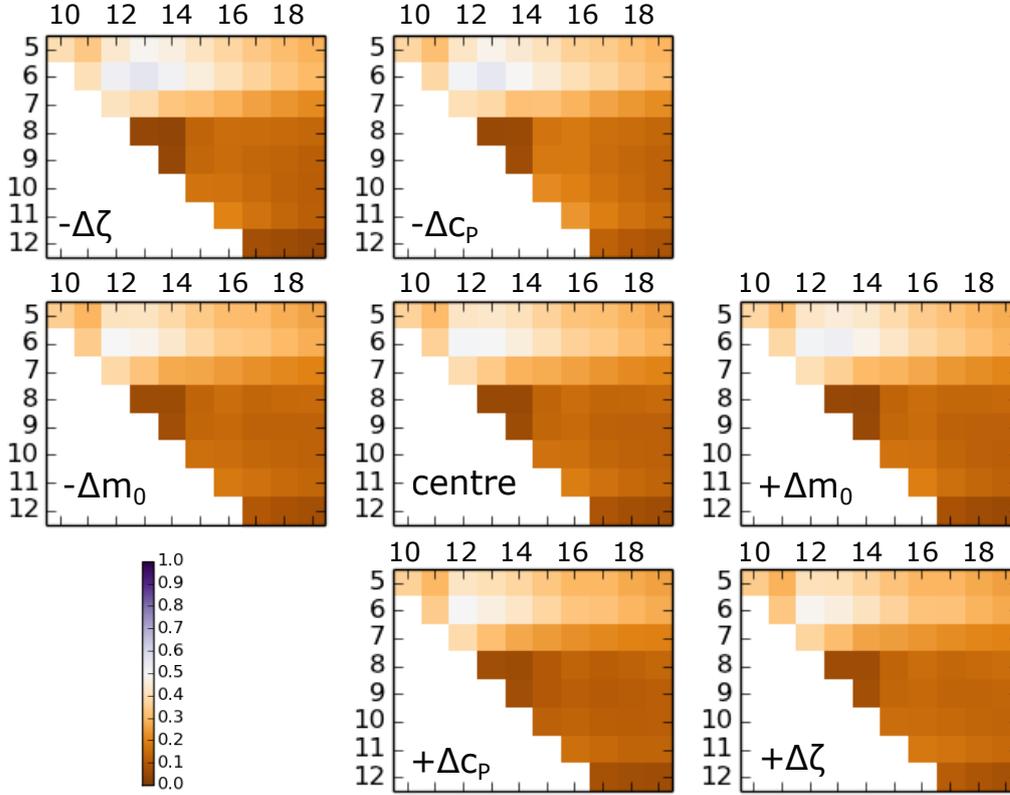
Figure 6.7: χ^2 values for B_l mass fits for each of the 7 *b* quarks on the (κ_l, κ_s)=(0.12104,0.12062) ensemble. On each subplot, the *y*-axis has the fit start point, and the *x*-axis the fit end point.

We notice that the pattern of χ^2 values is very similar for each *b* quark, though there is some variation in the χ^2 values. As an example, we see that the square representing the t=6–13 fit window is white for some *b* quarks (indicating a value close to 0.5), while more purple in others (indicating a value greater than 0.5).

Further development of visualisations to assist with fitting multiple *b* quarks simultaneously is likely to speed up the process of trialling different fit methods for these samples. Despite the possibility of improved understanding of the correlators and fit windows, human systematic error in the choice of fit windows will still occur unless an automated method of selecting the fit windows across multiple *b* quarks can be developed. These systematics, however, can be estimated by varying the fits and checking the impact on the tuning, or on the decay constant. This study is presented in Section 7.3.1 alongside our SU(3) breaking results.

6.3 Tuning B mesons

While a brief overview of the b quark action was provided in Chapter 5, the mechanics of the b tuning process are described here. The original tuning process outlined in [78] uses the physical properties of the B_s and B_s^* mesons to tune all ensembles, because the strange quark mass is held constant across ensembles for the configurations used in that work. As our strange quark mass is not constant, we instead use the properties of the flavour-average B mesons X_B and X_{B^*}. This means that we will need to incorporate flavour averaging of the B_l and B_s meson masses (and other properties) into the tuning process.

In brief, our tuning process involves combining calculated properties of $B_l^{(*)}$ and $B_s^{(*)}$ mesons and then comparing to the physical equivalents in PDG2015 [3]. Statistical errors in the fits and in the tuning are calculated using the bootstrap method, presented as an overview here but described in some further detail in Appendix B. Each ensemble of lattice configurations is resampled with replacement to produce 200 bootstrapped ensembles the same size as the original ensemble, and drawn from the same pool of configurations. This mimics the process of collecting multiple statistical samples from a population, so that we can calculate (say) the sample mean from a distribution of means. Similarly, by performing a fit to the correlator for each bootstrap ensemble, we produce 200 estimates of the fit value for which we can calculate the mean fit value and standard error. In this study, we perform as many calculations at the bootstrap level as possible in order to propagate correlations in the error to the final results.

The tuning process starts with the calculation of the required parameters (spin-averaged mass, hyperfine splitting, and kinetic mass coefficient) for each b quark and each bootstrap sample. The process of combining correlators is outlined below.

Spin-averaged mass

1. We calculate fits to individual correlators to obtain the masses for B_l, B_s, B_l^*, and B_s^* separately.

2. The spin-averaged mass is calculated from the fitted masses for each light or strange quark q:

$$\overline{M}_{B_q} = \frac{1}{4}\left(M_{B_q} + 3M_{B_q^*}\right) \tag{6.6}$$

3. These are then further combined into the flavour average:

$$\overline{M}_{X_B} = \frac{2}{3}\overline{M}_{B_l} + \frac{1}{3}\overline{M}_{B_s} \tag{6.7}$$

Hyperfine splitting

1. We take the correlator ratio B_q^*/B_q to allow the mass difference $\Delta M_{B_q} = M_{B_q^*} - M_{B_q}$ to be obtained directly from a single fit.

2. Splittings for the light and strange quarks are combined to form a flavour average $X_{B^*} - X_B$, similarly to the process in Equation 6.7.

Kinetic mass coefficient/dispersion coefficient

1. The energy of B_l and B_s is calculated from fits to correlators for each of $(a\vec{p})^2 = 0, 1, 2, 3$ in units of $(2\pi/L)^2$.

2. Energy for X_B at each momentum is calculated, labelled E_p.

3. We rewrite the dispersion equation as $(E_p - E_0)(E_p + E_0) = Ap^2$, converting to physical units for momentum p and the energies of each state. This formulation allows us to use a one-parameter weighted linear fit through (0,0) to determine A. From our correlator fits, we construct the sum $(E_p + E_0)$ and difference $(E_p - E_0)$ of energy values for each bootstrap. The standard deviation of all bootstraps for each momentum is used to provide energy errors input to the fit.

We now have 200 bootstrap samples of the spin-averaged mass, hyperfine splitting, and dispersion coefficient for each of the 7 b quarks in our tuning star. For each bootstrap, we follow the method in [78] to interpolate to the best tuning, assuming we are in a well-behaved region of the parameter space where the dependence of our output parameters is linear and could be described by the matrix equation

$$
\begin{bmatrix} \overline{M}_{X_B} \\ X_{B^*} - X_B \\ A \end{bmatrix} = J \cdot \begin{bmatrix} m_0 \\ c_P \\ \zeta \end{bmatrix} + C
\tag{6.8}
$$

for J a 3x3 matrix of linear coefficients and C the 3x1 column vector "intercept". The values of J and C can be computed from the points in our tuning star using simple finite difference methods. The first column of J is computed

$$
\begin{bmatrix} J_{11} \\ J_{21} \\ J_{31} \end{bmatrix} = \frac{1}{2\Delta m_0} \left(\begin{bmatrix} \overline{M}_{X_B} \\ X_{B^*} - X_B \\ A \end{bmatrix} \Bigg|_{m_0 + \Delta m_0} - \begin{bmatrix} \overline{M}_{X_B} \\ X_{B^*} - X_B \\ A \end{bmatrix} \Bigg|_{m_0 - \Delta m_0} \right)
\tag{6.9}
$$

with, for example, evaluation at $m_0 + \Delta m_0$ representing the b quark at $\{m_0 + \Delta m_0, c_P, \zeta\}$ on the tuning star. The other two columns of J are computed using c_P and ζ respectively. Having determined J, the intercept C is simply

$$
\left. \begin{bmatrix} \overline{M}_{X_B} \\ X_{B^*} - X_B \\ A \end{bmatrix} \right|_{\text{centre}} - J \cdot \left. \begin{bmatrix} m_0 \\ c_P \\ \zeta \end{bmatrix} \right|_{\text{centre}} .
\tag{6.10}
$$

Finally, the tuned parameters are given by

$$
\begin{bmatrix} m_0 \\ c_P \\ \zeta \end{bmatrix}^{\text{best}} = J^{-1} \cdot \left(\begin{bmatrix} \overline{M}_{X_B} \\ X_{B^*} - X_B \\ A \end{bmatrix}^{\text{PDG}} - C \right)
\tag{6.11}
$$

for each bootstrap.

An example of the plots of spin-averaged mass, hyperfine splitting, and kinetic mass coefficient against the tuned parameters m_0, c_P, and ζ for the $\beta = 5.5, (\kappa_l, \kappa_s) = (0.12104, 0.12062)$ ensemble are shown in Figure 6.8. Figure 6.9 shows a zoomed-in version of the upper-left plot in the tuning array, the spin-averaged mass with respect to m_0.

In practice, the tuning star must be re-calculated and new correlators generated several times for each lattice spacing/lattice size in order to find a region of the parameter space where an appropriate b can be linearly interpolated. Fortunately, this only needs to be performed once for each value of β rather than for each individual ensemble. We also apply an additional constraint that only ζ values ≥ 1 are used in the tuning stars, as the $\zeta < 1$ region displayed nonlinear and unstable behaviour in early tuning studies.

The upper-right plot of Figure 6.8 displays spin- and flavour-averaged B meson masses against ζ. In practice, this plot is the most likely to diverge from linear behaviour. Typically at some parts of the parameter space, this plot will show evidence of quadratic behaviour opening downward. We must then choose whether to tune to a high-ζ or low-ζ space. Occasionally, this quadratic behaviour means that a low-ζ space will have an extrapolated 'best' tuning in the high-ζ space, and a star centred on the suggested high-ζ tuning will have an extrapolated 'best' tuning matching the original low-ζ region. This means that a best tuning fit cannot be determined using only the linear method in Equation 6.8, as the assumption that the tuning is linear in all variables within a region surrounding the tuning star has been violated.

From early studies, we find that $24^3 \times 48$ lattice volumes suffer from more systematic errors and require a different tuning star configuration to $32^3 \times 64$ lattices
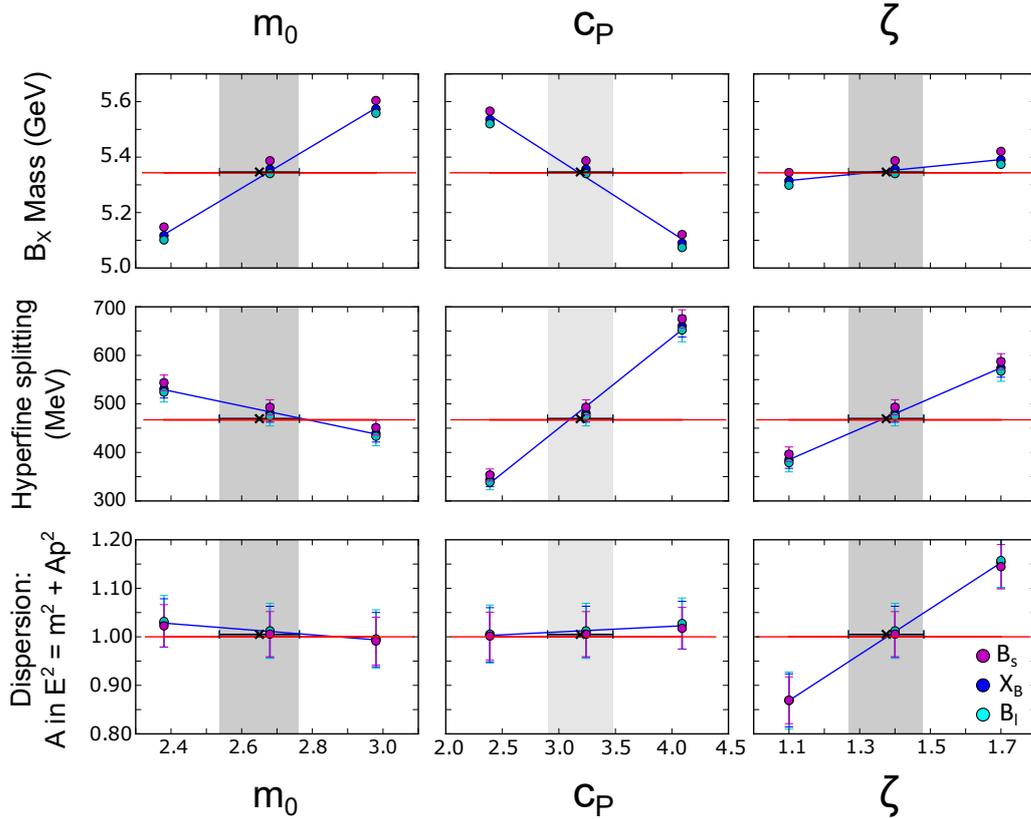
Figure 6.8: Plots of each physical quantity against each of the tuning parameters $\{m_0, c_P, \zeta\}$, for the $\beta = 5.5, (\kappa_l, \kappa_s) = (0.12104, 0.12062)$ ensemble. The 1σ error bands for the best combination of m_0, c_P, ζ are shown in grey. with the central value marked by an x. The dark blue circles show the properties of the calculated X_B meson for each b quark, with the lines indicating the linear fit to each parameter.
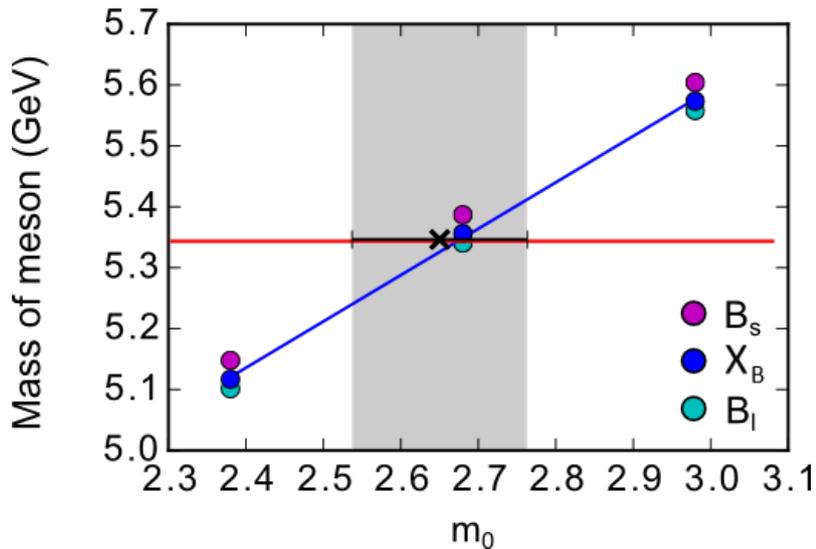
Figure 6.9: Zoomed-in view of the measured meson mass vs m_0 plot in Figure 6.8. The individual calculated B_s and B_l meson properties are shown alongside the calculated X_B

Table 6.3: The calculated 'best' tuning parameters and error margins for each of the ensembles used. * denotes ensembles with a different value of $\overline{m}$, further from the physical $\overline{m}$, represented in dark blue in all Figures. $\dagger$ denotes the near-physical $64^3 x96$ ensemble which has extrapolated parameters

β	κ_l	m_0	c_P	ζ
5.4	0.11993	3.56 ± 0.14	3.73 ± 0.36	1.59 ± 0.12
	0.119989	3.62 ± 0.13	3.88 ± 0.35	1.60 ± 0.12
	0.120048	3.58 ± 0.15	3.73 ± 0.40	1.57 ± 0.14
	0.120084	3.76 ± 0.16	4.27 ± 0.41	1.53 ± 0.14
5.5	0.12095	2.92 ± 0.13	3.86 ± 0.34	1.23 ± 0.12
	0.12104	2.82 ± 0.13	3.59 ± 0.34	1.38 ± 0.10
	0.121099	2.83 ± 0.12	3.61 ± 0.31	1.26 ± 0.11
5.5*	0.1209	2.80 ± 0.13	3.60 ± 0.34	1.30 ± 0.11
	0.12104	2.65 ± 0.11	3.19 ± 0.29	1.37 ± 0.11
	0.121095	2.86 ± 0.11	3.70 ± 0.29	1.21 ± 0.09
	0.121145	2.92 ± 0.14	3.86 ± 0.35	1.11 ± 0.14
	0.121166	2.75 ± 0.10	3.42 ± 0.25	1.34 ± 0.08
5.65	0.122005	2.67 ± 0.14	4.18 ± 0.38	1.07 ± 0.10
	0.122078	2.48 ± 0.15	3.72 ± 0.39	1.12 ± 0.11
	0.12213	2.52 ± 0.09	3.78 ± 0.24	1.16 ± 0.08
	$0.122167^{\dagger}$	2.49 ± 0.13	3.67 ± 0.34	1.25 ± 0.10
5.8	0.122227	3.18 ± 0.20	5.42 ± 0.52	0.96 ± 0.13
	0.12281	3.03 ± 0.09	5.30 ± 0.24	1.21 ± 0.07
	0.12288	3.28 ± 0.09	6.06 ± 0.27	1.14 ± 0.06
	0.12294	3.00 ± 0.08	5.25 ± 0.22	1.30 ± 0.06

with matching parameters. Lattice volume effects on $24^3 \times 48$ lattices can also be observed in the ratio of two- and three-point correlators (see Section 6.4.1 and Figure 6.15). In general however, the same tuning star may be used appropriately for both $32^3 \times 64$ and $48^3 \times 96$ ensembles and no volume effects are observed.

Using our default fitting method (see Section 6.2.1), we calculate the parameters corresponding to the 'best' B meson for each ensemble. These parameters are presented in Table 6.3 and displayed graphically in Figure 6.10. On the coarsest lattice ensemble ($\beta = 5.4$, orange), we are tuned to a higher-ζ region to account for stronger discretisation effects. As we approach finer lattice spacings and thus also the continuum limit, we favour values of ζ closer to 1.

During the tuning process, a smaller number of configurations per ensemble are used (typically 200-400) to test candidate b tuning star options more quickly. Throughout this thesis, the tuning star for each lattice spacing a has been selected via tuning on the ensemble with exact SU(3) symmetry, $m_u = m_d = m_s$. Due to time and computing constraints, tuning for the $64^3 \times 96$ near-physical-pion-mass ensemble (olive marker in figures) could not be completed before completion of this thesis, and as a result the best tuning is extrapolated for this ensemble.
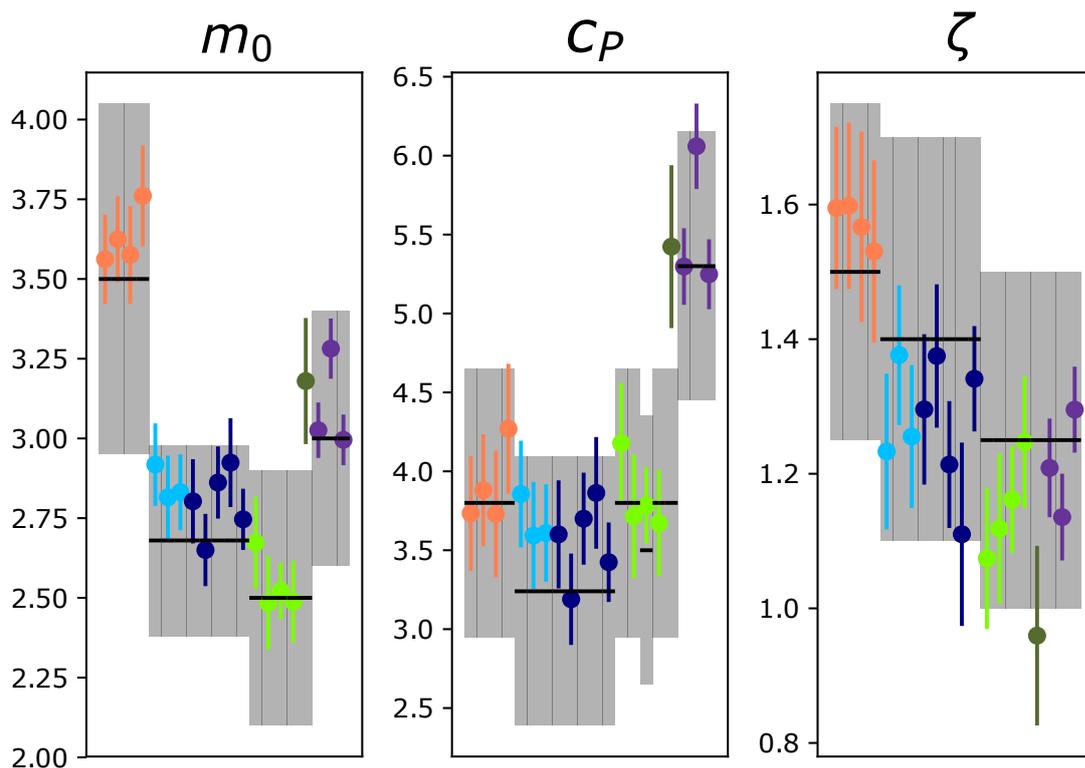
Figure 6.10: The outcome of the tuning process on all ensembles. On each subfigure, colour is used to represent the β value, with the leftmost point in each colour being the ensemble at the SU(3) point, and continuing toward the physical splitting on the right of each block. The colours are assigned to ensembles matching the colour scheme used in Figure 6.2. The black horizontal lines show the central value in the tuning "star", while the edges of the grey bands are at $\pm\Delta m_0$, c_P, or ζ for the star.

In future work, the systematic error in the tuning might be somewhat reduced by combining several correlators to calculate the spin-averaged mass with a single fit. Additionally, the flavour average could be applied at correlator level to reduce the error in combining the B_l and B_s results. Unfortunately, this does increase the overall fitting burden, as individual mass fits for B_l and B_s mesons still are required for the decay constants f_B and f_{B_s}.

6.4 Decay constants on the lattice

The decay constant is calculated from two point correlation functions, and the form of this equation is defined in Section 5.4.3. This calculation requires new fits to correlation functions, including to $\mathcal{C}_{AP}$ for the decay constant itself, and to the three point functions for the renormalisation of the axial current. In this Section, the overall renormalisation Z_Φ will be discussed first.

6.4.1 Renormalisation constants Z_V

The overall renormalisation Z_Φ requires the calculation of the Z_V^{ll}, Z_V^{ss}, and Z_V^{bb} components on the lattice. The first constant, Z_V^{bb} is calculated with a current insertion on the bottom quark in a meson, with the other quark as a spectator. In practice, rather than generating two- and three-point correlators for a $b\bar{b}$ meson, we calculate Z_V^{bb} using a lighter spectator quark so that the two-point correlators from the main B meson study can be used alongside the new three-point functions in the ratio (see Equation 5.88). A diagram of the three point function is shown in Figure 6.11. The three point functions are generated using a sequential-source method, such that the momentum at the sink is fixed. In this method, the entire light or strange quark propagator from source to sink is considered as a source for the remaining b propagator.

As the current insertion is on the heavy quark, the light and strange quark propagators can be recycled for each b quark in the tuning star and also for the calculation of Z_V^{ll} and Z_V^{ss}. We find that there is no significant difference between Z_V^{bb} calculated with a light vs strange spectator quark on the lattices tested.

The actual Z_V values are calculated as $1/R$ for R the ratio of two- and three point correlators, as seen in Equation 5.88. In this particular case, the momenta at the source and sink are both zero and the large square root term simplifies to 1. It is crucial to select a sink time late enough that we have a sufficiently long plateau to fit, but also not too late, to ensure that the correlators at the fixed sink are not too noisy. In these results, we have used a source-sink separation of 16 lattice sites.

An example fit to the ratio of the two and three point correlators is shown in
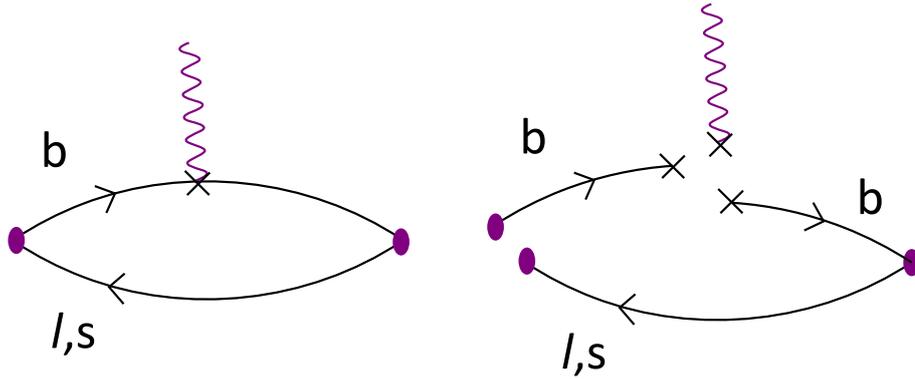
Figure 6.11: Feynman diagram used to calculate Z_V^{bb} (left). The current insertion is on the *b* quark, and the light or strange quark acts as a spectator. In the right part of the Figure, this diagram is broken down into sections to show the way the sequential source propagator is constructed.

Figure 6.12. We note that the statistical error on the points is extremely small, due to sources of error cancelling in the ratio. In general, larger error appearing in this ratio for *B* is a sign that the fixed sink time must be adjusted and the correlators re-calculated, taking care not to introduce additional excited state contamination.
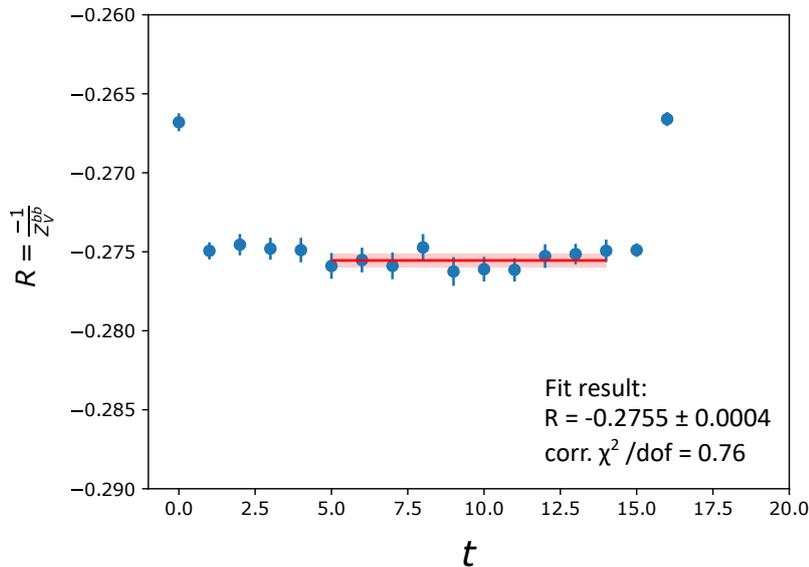
Figure 6.12: Ratio R of three point and two point correlators against lattice time coordinate t. Constant fit is shown in red, the fixed sink time is 16. This plot is for the central b in the tuning star for $(\kappa_l, \kappa_s) = (0.12104, 0.12062)$

When Z_V^{bb} has been fit for all seven b quarks, we can interpolate to the calculated best tuning from the ensemble and thus determine the Z_V^{bb} value corresponding to our 'best' B. In practice, we calculate $\sqrt{Z_V^{bb}}$ at the bootstrap level and quote the error in $\sqrt{Z_V^{bb}}$ as this is the quantity that is directly used in the calculation of the decay constant. The outcome of this process is shown in Figure 6.13.

For some ensembles, we observe some quadratic or nonlinear behaviour in $\sqrt{Z_V^{bb}}$
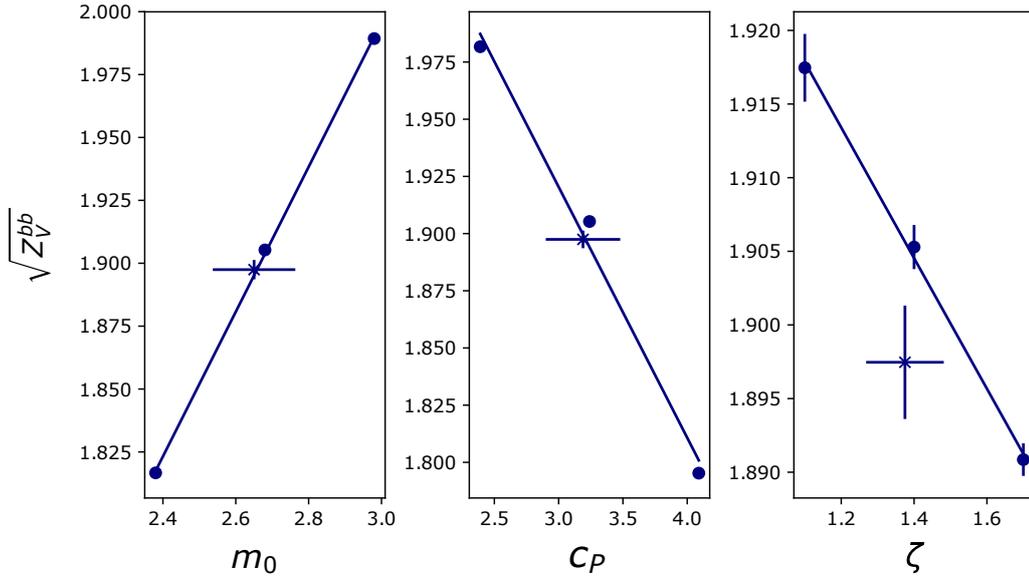
Figure 6.13: Calculated Z_V^{bb} values for the 7 b quarks in the tuning star for the $(\kappa_l, \kappa_s) = (0.12104, 0.12062)$ ensemble. The 'best' Z_V^{bb} calculated by interpolating in 3D to the optimal tuning parameters is marked by an x with errorbars shown.

relative to the c_P or ζ parameters, although this effect is usually small.

The calculated values of $\sqrt{Z_V^{bb}}$ for all ensembles are presented in Figure 6.14. Behaviour of the normalisation term is very linear with respect to a^2 as expected. Z_V^{bb} is also approximately constant for each lattice spacing even as M_π^2 varies. We note, however, that three point functions have not yet been produced for the $64^3 \times 96$ near-physical-pion lattice. In the rest of this thesis, the average value of Z_Φ calculated from all other $\beta = 5.65$ ensembles is used for the near-physical-pion lattice.

Z_V^{ll} and Z_V^{ss} were calculated on several ensembles, including at smaller lattice volumes to check for volume effects. These results are presented in Figure 6.15. The same colour scheme is used as in the rest of this thesis, with the exception of the new colour used to represent the single $24^3 \times 48$ $\beta = 5.4$ ensemble at the SU(3) symmetric point. We observe a significant difference between Z_V^{ss} calculated on the $24^3 \times 48$ and $32^3 \times 64$ lattices, indicating that the $24^3 \times 48$ lattices are too small and suffer from systematic volume errors. This also confirms that $32^3 \times 64$ lattices are sufficiently large, as no significant change is seen between results for $32^3 \times 64$ and $48^3 \times 64$ ensembles.

As we move further from the SU(3) symmetric point, the values of Z_V^{ll} and Z_V^{ss} separate. In the ensembles measured, this effect is 1-2 %. Unfortunately, this data is not currently available in all ensembles, especially those closer to the physical pion mass. Instead, an average of Z_V^{ll} and Z_V^{ss} is used for each lattice spacing and this known 1-2 % effect is included as a systematic in our final calculation of the decay constant.

The combined renormalisation constant, Z_Φ is shown for each ensemble in Figure
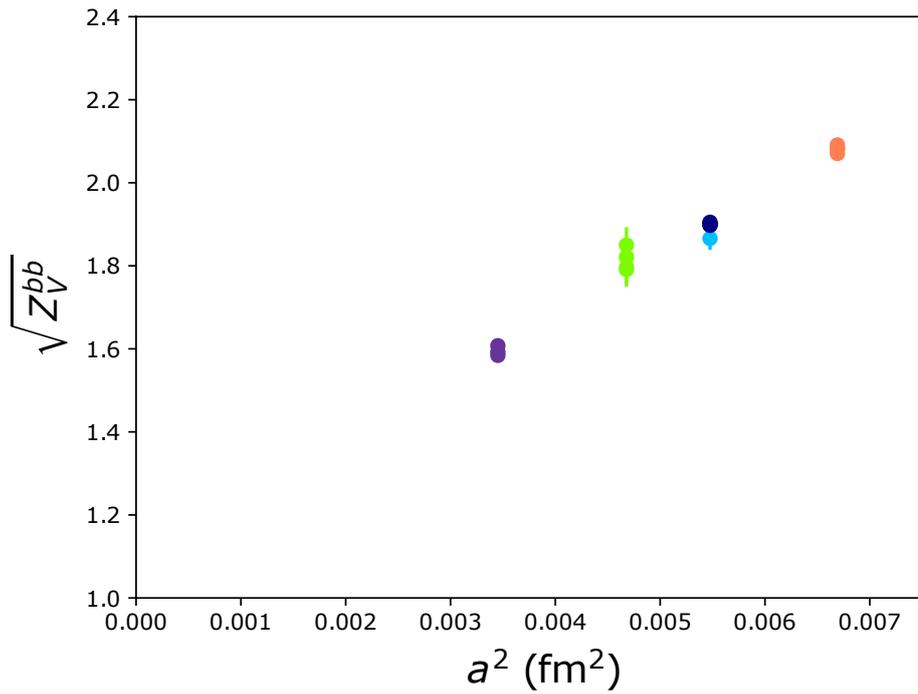
Figure 6.14: The heavy quark renormalisation for the decay constant against the lattice spacing a^2. The result is approximately linear across the region studied, and trends toward 1 in the continuum limit.
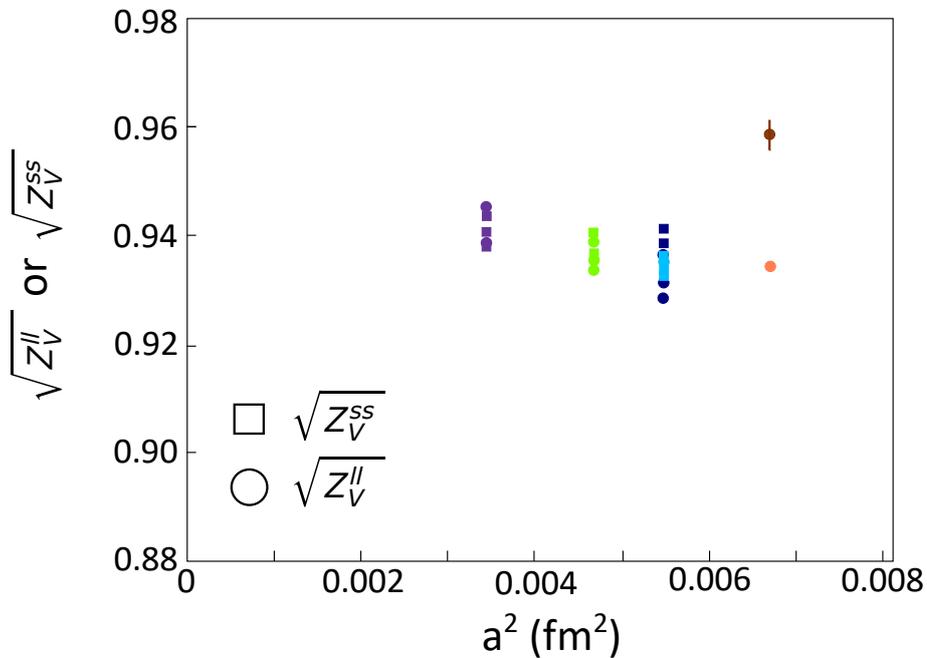
Figure 6.15: The light or strange quark renormalisation for the decay constant against the lattice spacing a^2. We note that the square symbol is also used for the SU(3) symmetric points.
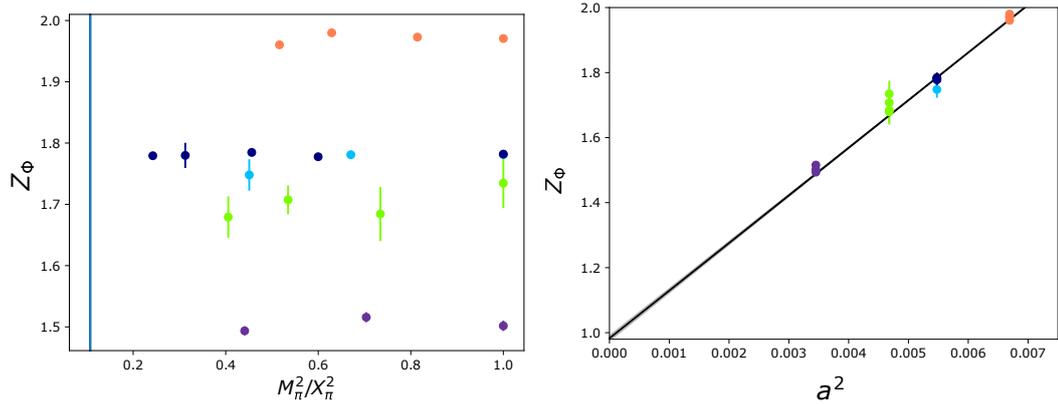
Figure 6.16: The heavy quark normalisation for the decay constant against the light quark splitting (left) or against the lattice spacing a^2 (right).

6.16. We see that the value of Z_Φ is approximately constant with respect to M_π^2/X_π^2. If we consider the behaviour with respect to lattice spacing a, we see Z_Φ trending toward 1 as we move from the coarsest lattices (orange) toward finer lattice spacings (purple). A simple linear fit to the Z_Φ against a^2 plot yields a y-intercept value of 0.982 ± 0.008.

6.4.2 Applying tuning to f_B

The decay constant itself is calculated using the currelator fit for the B or B_s meson mass ($\mathcal{C}_{PP}$), combined with a new correlator fit to $\mathcal{C}_{AP}$ which includes the axial operator (see Section 5.4.3). The shape of the effective mass of this correlator is shown in Figure 6.17.

Similarly to the tuning process, the decay constants for the B_l and B_s mesons are calculated for each b quark, and we select a fit window that has acceptable performance for all seven b in the tuning star. The 'best' decay constant values using the tuned b on each ensemble are calculated as follows:

1. Tuning is performed on 200 bootstraps, as detailed in Section 6.3. The additional fit to the smeared-point correlator is performed using the same set of 200 bootstraps as the smeared-smeared correlator used to calculate the mass.

2. The lattice decay constants (Φ_B^0) corresponding to the B_l and B_s mesons are calculated according to Equation 5.97 for each of the seven b quarks and for each bootstrap. The AP and PP correlators are fit individually.

3. Linear fits in terms of m_0, c_P, or ζ are calculated for Φ_B and Φ_{B_s} on each bootstrap. The matching best tuning for the individual bootstrap is used to interpolate a 'best' Φ_B and Φ_{B_s} on each bootstrap. We then take the mean to obtain our final decay constant values for the ensemble.
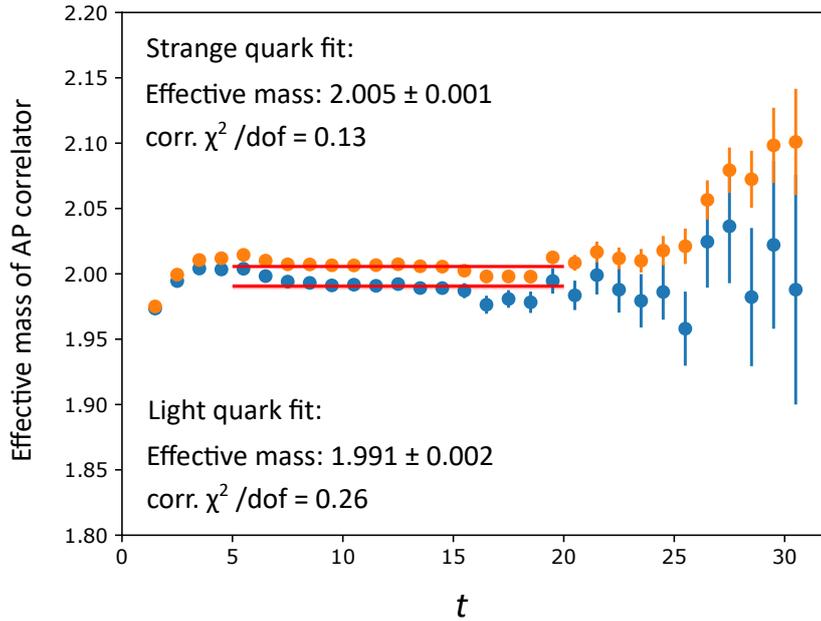
Figure 6.17: Effective mass for the AP correlator of the *bl* (blue) and *bs* mesons (orange) for the central *b* quark on the $(\kappa_l, \kappa_s) = (0.12104, 0.12062)$ ensemble. The correlator fits are shown in red.

4. The 'best' renormalisation constant Z_Φ is calculated separately and applied to the mean Φ_B and Φ_{B_s}, in order to produce f_B and f_{B_s}. The flavour-singlet value $f_{X_B} = \frac{1}{3}(2f_B + f_{B_S})$ is also calculated.

An example outcome of the tuning process for the decay constant is shown in Figure 6.18. In this plot, we present f_B in physical units but note that $Z_\Phi = 1$ is applied to show the behaviour of the decay constant fits only.

As shown, the behaviour of the decay constant with respect to the three *b* quark parameters is very linear. This remains linear as we choose ensembles with greater splitting between the light and strange quark masses. The value extracted for the decay constant, however, is sensitive to the choice of fit windows, and also on the fits used in tuning. Some of these effects can be mitigated by considering ratios of the decay constants, in particular f_B/f_{X_B}. These systematic effects will be described in detail in the next Chapter in the context of the SU(3) symmetry breaking across ensembles. The individual values of f_B and f_{B_s} will also be presented.
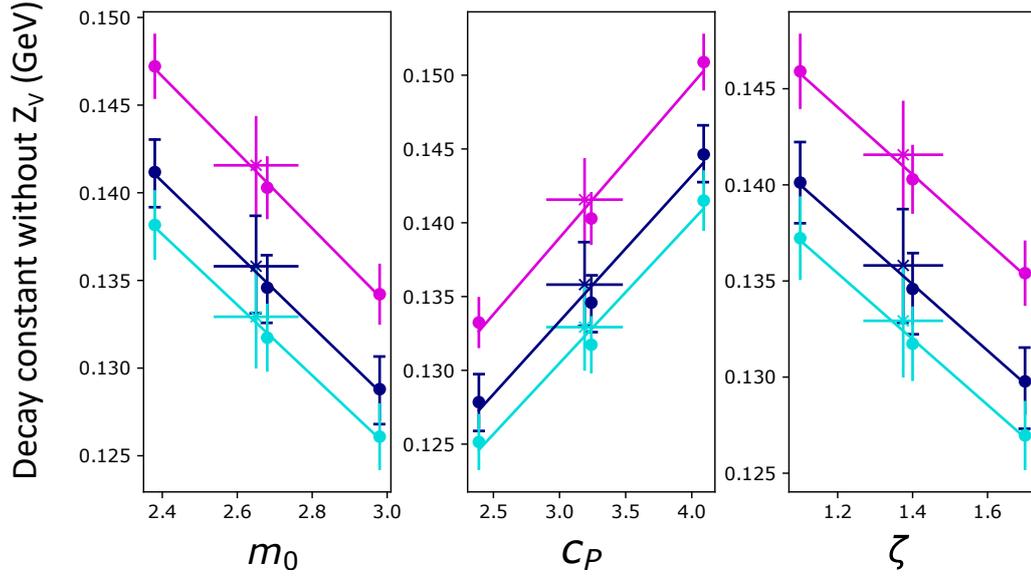
Figure 6.18: Calculated decay constant values from lattice results (with $Z_\Phi = 1$) for all 7 b quarks on the $(\kappa_l, \kappa_s)=(0.12104,0.12062)$ ensemble. f_B (cyan), f_{B_s} (magenta) and $f_{X_B} = \frac{1}{3}(2f_B + f_{B_s})$ (navy) are shown. The interpolated 'best' decay constant for each set is marked with an $\times$.

6.5 Summary

In this Chapter, we have described the QCDSF lattice configurations used in this study, including an overview of the SU(3) symmetry breaking approach. The tuning process for b quarks on these lattices has been explained in detail, and our default correlator fitting process is described. Other fitting methods are briefly discussed, and the comparison of different fits in the context of the decay constant will be explained further in Chapter 7.

The calculation of the decay constant from correlators was also described. The calculated renormalisation constants $Z_V^{bb/ll/ss}$ are presented, and volume and lattice spacing effects are discussed in detail. In Chapter 7, the decay constant will be compared across all ensembles.

Chapter 7

SU(3) breaking of the decay constant

Having completed the overview of tuning the b quarks and calculating the decay constant on each ensemble, in this Chapter we will focus on the results of our SU(3) breaking analysis across all ensembles. Investigation of different types of fits to the SU(3) breaking ratios are used to ensure our analysis is robust, and to investigate possible sources of systematic error.

As part of this discussion of error, we combine an overview of the SU(3) symmetry breaking and the effects of the SU(3) splitting across different ensembles with a quantitative study of extrapolations of the ratio f_{B_s}/f_B to the physical pion mass. Finally, our calculated value for f_{B_s}/f_B is presented with both statistical and systematic error estimates.

7.1 Ratios of the decay constant

7.1.1 SU(3) breaking expansion

As our lattice ensembles are chosen to keep the average quark mass ($\overline{m}$) constant, any increase δm_s in the strange quark mass must be matched by a decrease in the light quark mass such that $\delta m_s = -2\delta m_l$. When this leads to increases (or decreases) in measured quantities involving the strange quark mass, we expect this to be balanced by decreases (or increases) in light quark observables [72]. This simplifies extrapolations toward the physical point and allows us to make a Taylor expansion in the quark masses from the SU(3) symmetric point. Similarly, any flavour singlet quantity which contains a 2:1 ratio of light to strange quarks is expected to remain constant as we approach the physical point.

We first consider the ratio f_B/f_{X_B}, where $f_{X_B} = \frac{1}{3}(2f_B + f_{B_s})$ is the 'average' or flavour-singlet decay constant. The SU(3) breaking in these ratios of f_B (or f_{B_s})

can be modelled by adapting the partially-quenched NLO equations for the light
and strange pseudoscalar mesons in Bornyakov et al [81]:

$$
\begin{aligned}
\frac{f_B(q\bar{b})}{f_{X_B}} = {} & 1 + G(\delta\mu_q) + (H_1 + H_2)\delta\mu_q^2 \\
& - (\tfrac{2}{3}H_1 + H_2)(\delta m_u^2 + \delta m_d^2 + \delta m_s^2) \\
& + \dots
\end{aligned}
\tag{7.1}
$$

where $\delta\mu_q = \mu_q - m_0$ represents the distance between the valence quark mass μ_q and
the SU(3) symmetric mass m_0, and similarly δm are differences in the sea quark
masses. Compared to the original SU(3) breaking expansions of decay constants
where both quarks are lighter [81], in this expansion the heavy b quarks can be
ignored.

Ratios of the decay constants are calculated at the bootstrap level from the
individual fits to f_B and f_{B_s}. Dependence on the renormalisation constant Z_V^{bb} is
removed in the ratio, and because we approximate $Z_V^{ll} = Z_V^{ss}$, the ratio results are
independent of Z_Φ (see Section 6.4.1).

In Figure 7.1, we show the calculated decay constant ratios f_B/f_{X_B} and f_{B_s}/f_{X_B}
for all ensembles against $\delta\mu_q$, the mass of the lighter quark in the B or B_s meson.
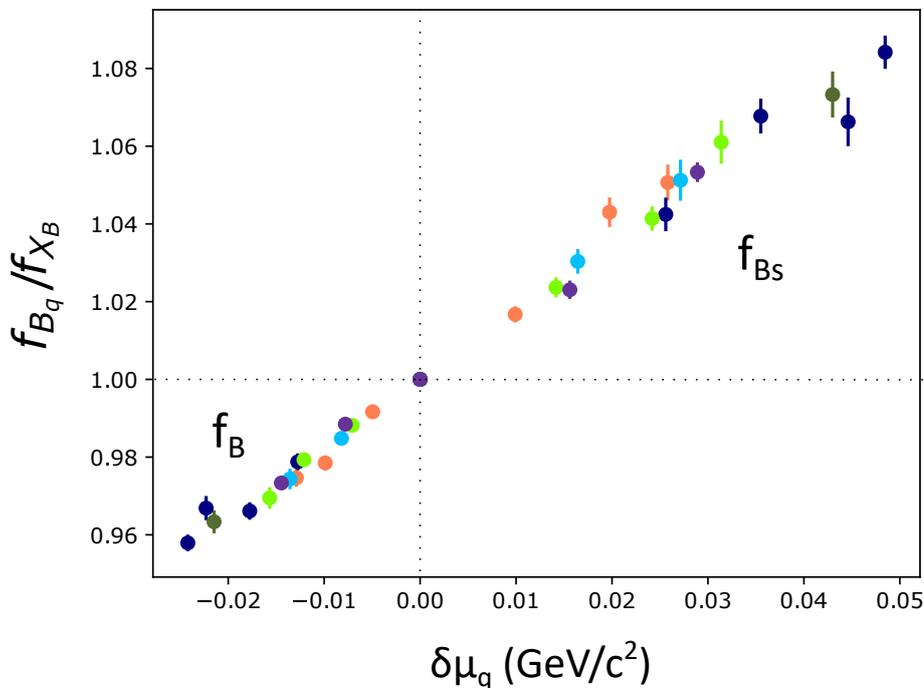

Figure 7.1: f_B/f_{X_B} and f_{B_s}/f_{X_B} for a variety of lattice ensembles, against the
change in quark mass relative to the matching SU(3) symmetric point ($\delta\mu_q$). The
symmetric point is marked by the dotted lines. Colours for ensembles continue to
match those in Figure 6.2.

We observe that for all ensembles, this expansion is mostly linear, although we can see some hints of the expected quadratic behaviour for larger values of $\delta\mu_q$. The coefficients G, H_1, and H_2 (see Equation 7.1) are expected to be different for each lattice spacing, and in practice the coefficients H_1 and H_2 are difficult to separate and constrain, as all of our current data uses light and strange quarks matched to the light and strange quarks in the sea of the ensemble. In this expansion, the target physical $\delta\mu_q$ is also different for each lattice spacing. To better compare across all ensembles, we instead must consider an expansion in another variable related to the SU(3) symmetry breaking, scaled in terms of the SU(3) and physical points. We choose the mass breaking ratio in the light quarks, defined

$$\frac{M_\pi^2}{X_\pi^2} = \frac{3M_\pi^2}{(M_\pi^2 + 2M_K^2)} \tag{7.2}$$

for the pion and kaon masses calculated on each lattice ensemble. This type of expansion has already been used in Section 6.1 to illustrate the SU(3) splitting in the masses of the pion and kaon. As QED effects are not included, we use the neutral pion and kaon whenever we calculate our 'physical' value of M_π^2/X_π^2, matching our $N_f = 2 + 1$ assumption that $m_u = m_d$.

Our calculated decay constant ratios f_B/f_{X_B} and f_{B_s}/f_{X_B} for all ensembles are shown against this SU(3) breaking ratio M_π^2/X_π^2 in Figure 7.2. It should be noted that the statistical error here is mostly a result of error propagation in interpolating the fit to the 'best' B meson on each ensemble: the ratio itself has very small statistical error for each individual b quark.

The FLAG2016 values used for comparison are calculated from the world average result for f_{B_s}/f_B on $N_f = 2 + 1$ samples [65]. There were no additional $N_f = 2 + 1$ results for f_B or f_{B_s} in the FLAG 2019 update that would change this average [32]. The observed splitting between the decay constants f_B and f_{B_s} as we approach the physical pion mass from the SU(3) symmetric point is shown to be approximately linear, although both linear and quadratic fits are provided. This 'fan plot' combines results from all ensembles, and similarly, the fits presented use all data points shown in the Figure. In the next section, we discuss different possible fits to this SU(3) breaking in further detail.

7.1.2 Fitting to the SU(3) breaking

While we wish to make a prediction for the SU(3) breaking at the physical point, we are also interested in the behaviour of fits to subsets of the data, especially fits to all ensembles with the same lattice spacing. We also investigate the effect of excluding the set of ensembles at $\beta = 5.5$ where $\overline{m}$ does not match the physical

Figure 7.2: f_B/f_{X_B} and f_{B_s}/f_{X_B} for a variety of lattice ensembles. The linear and quadratic fits shown are constrained to pass through the fixed point (1,1).

value (navy blue in each Figure), or exclude the near-physical lattice which has $m_\pi L < 4$. Additional properties of these ensembles can be found in Section 6.1.1.

As the SU(3) breaking ratio is known to be one at the SU(3) symmetric point where $m_l = m_s$, we choose fit functions that expand around this known point (1,1) in f_B/f_{X_B} vs M_π^2/X_π^2. The functional form of the fit functions used are shown in Table 7.1. In particular, we include a fit with some a^2 dependence in the linear coefficient of M_π^2/X_π^2 in order to be able to present a continuum extrapolation for the decay constant ratios.

Table 7.1: Table of the functional forms of fit functions used. Constant coefficients are labelled G (linear), H (quadratic) or C (constant y-intercept).

Type of fit	Functional form
Linear	$G_0 \left(M_\pi^2/X_\pi^2 - 1\right) + 1$
Quadratic	$H \left(M_\pi^2/X_\pi^2 - 1\right)^2 + G_0 \left(M_\pi^2/X_\pi^2 - 1\right) + 1$
Quadratic with a^2	$H \left(M_\pi^2/X_\pi^2 - 1\right)^2 + \left(G_0 + G_1 a^2\right) \left(M_\pi^2/X_\pi^2 - 1\right) + 1$

These fits are performed for relevant subsets of the data: on each set of ensembles that share the same value of a^2, plus some additional variations to exclude some ensembles that may have different systematic error from the others. Plots of linear and quadratic fit coefficients for each type of fit to multiple sets of ensembles are presented in a grid format in Figure 7.3.
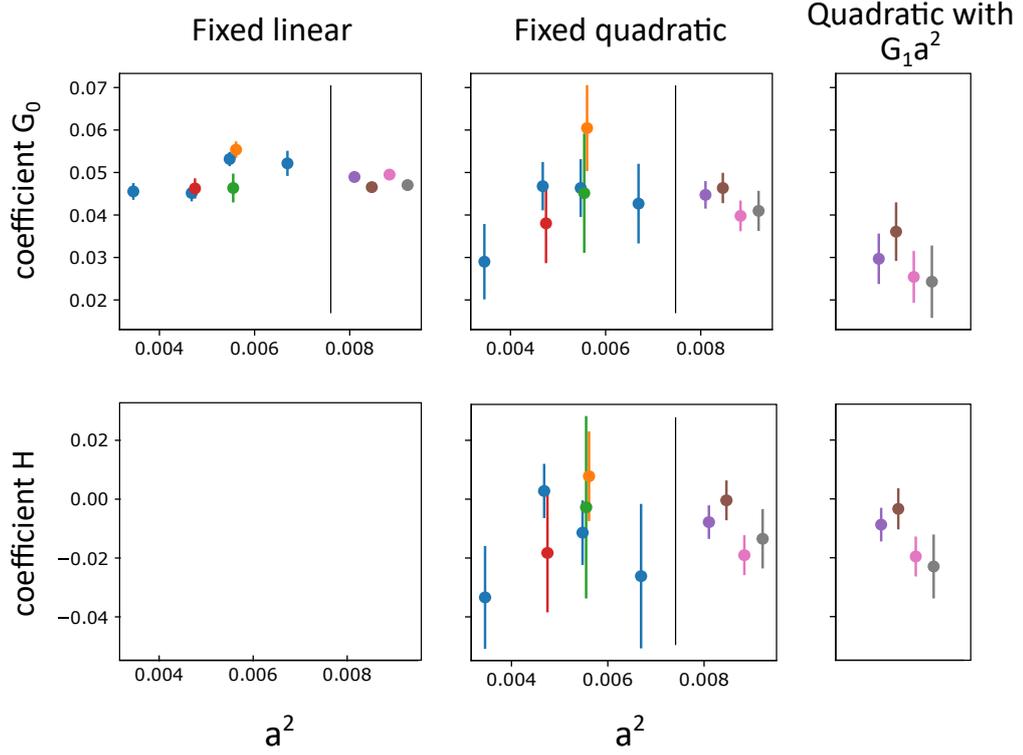
Figure 7.3: Fit coefficients for various fits to the SU(3) symmetry breaking f_B/f_{X_B} against M_π^2/X_π^2 in Figure 7.2. Linear fit coefficients (top row) and quadratic fit coefficients (bottom row) are shown against the a^2 of the ensembles used in the fit when applicable. The functional forms of the fits are presented in Table 7.1.

Combining information from the top row of the Figure, we see that there is no obvious dependence of the linear fit coefficient G_0 on the lattice spacing, with the value of G_0 appearing to be well-approximated by a constant across all ensembles. Fits incorporating a possible linear dependence on a^2 are nonetheless included in the last column of the grid for comparison. We notice that when the $G_1 a^2$ term is included, the fit value of G_0 is smaller. Using an average value of a^2, such as $a^2 = 0.0055$, and using the central value for G_1, we find that $G_1 a^2$ is approximately 0.016, such that $(G_0 + G_1 a^2)$ is 0.046. This is within the range of values for the G_0 coefficient in the quadratic fit without a^2, and close to some of the G_0 coefficient values for the linear-only fit.

We also find that some of the quadratic coefficients H are found to be consistent with zero. These results broadly conform to our expectations that the splitting in f_B/f_{X_B} as we approach the physical point should be mostly linear. In order to compare our results and error sizes to other groups, however, we need to consider a quantity that other groups do calculate directly.

7.1.3 The ratio f_{B_s}/f_B

In most studies, SU(3) symmetry breaking in the decay constants is reported in terms of the ratio f_{B_s}/f_B. This ratio is also used in the calculation of f_B itself from most groups, as f_B calculated directly has a larger error than the combination of f_{B_s} and the ratio. By extrapolating our calculated f_{B_s}/f_B result to the physical point, we will be able to compare our results to the FLAG averages. The ratio f_{B_s}/f_B for all ensembles is shown in Figure 7.4. As in Figure 7.2, the $N_f = 2+1$ FLAG average values are marked for visual comparison.

We observe that our ratio f_{B_s}/f_B is smaller than the world average. As discussed in Section 6.4.1, our calculation of f_{B_s}/f_B is missing the 1-2% difference between Z_V^{ss} and Z_V^{ll} with quark masses close to the physical SU(3) splitting, which would be expected to enhance f_{B_s}/f_B at these low values of M_π^2/X_π^2 in particular. This systematic will be included in our final prediction for f_{B_s}/f_B in Section 7.4.

As in the previous Section, we make multiple fits to the decay constant ratio in order to assess the impact of lattice ensemble effects. The fit coefficients are displayed in Figure 7.5. As we have a FLAG value of f_{B_s}/f_B for comparison, we also present extrapolations of each fit to the physical pion and kaon masses. The extrapolated results for different fit types are displayed in Figure 7.6, with key values also presented in Table 7.2. For fit functions containing a^2, the extrapolation to $M_\pi^2/X_\pi^2 = 1$ also includes the continuum extrapolation to $a^2 = 0$.

As for the f_B/f_{X_B} case, many of the quadratic coefficient predictions are consistent with zero. Similarly, we also see a change in the coefficient G_0 in the case
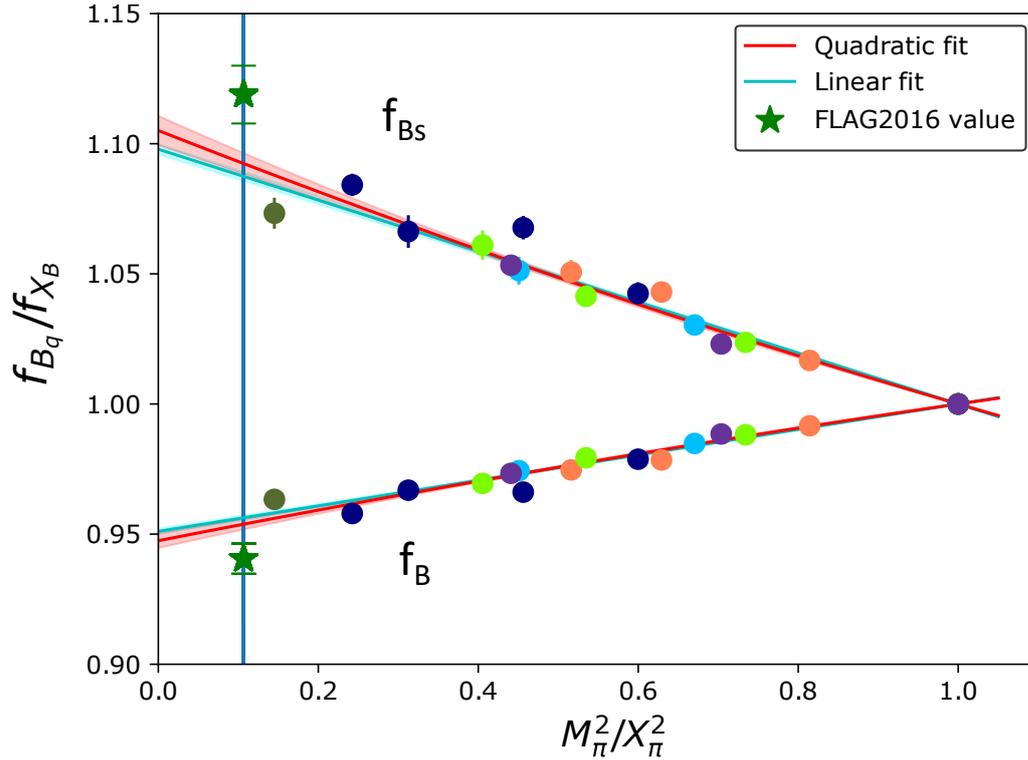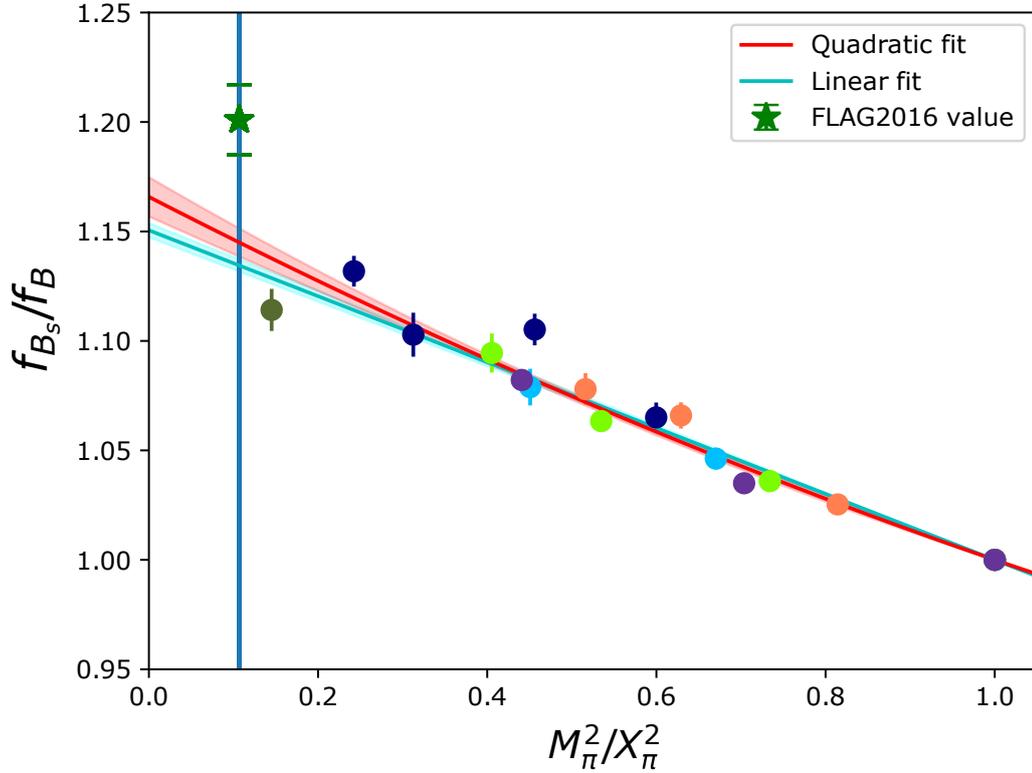
Figure 7.4: f_{B_s}/f_B for a variety of lattice ensembles. The linear and quadratic fits shown are constrained to pass through the fixed point (1,1).

Table 7.2: Extrapolated values of the f_{B_s}/f_B ratio for different fit types. The fits for $m_\pi L > 4$ include all ensembles except the near-physical pion mass $\beta = 5.65$ ensemble.

Data	Fit type	Value at physical	stat. error	χ^2/dof fit
All ensembles	Linear	1.134	0.003	1.8
	Quadratic	1.145	0.006	1.8
	Quadratic with a^2	1.105	0.015	1.3
$m_\pi L > 4$	Linear	1.136	0.003	1.8
	Quadratic	1.159	0.008	1.3
	Quadratic with a^2	1.120	0.015	0.9
FLAG value		1.201	0.016	
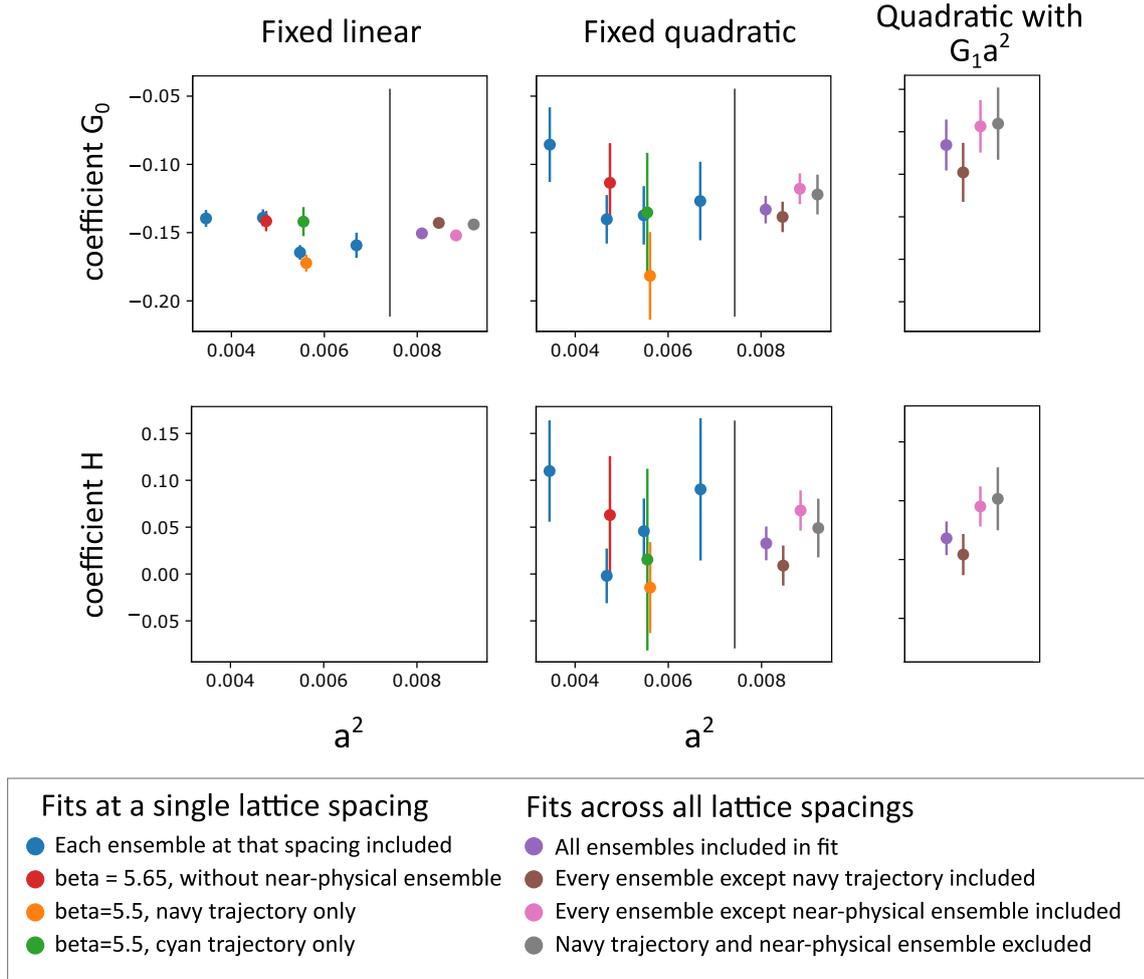
Figure 7.5: Fit coefficients for various fits to the ratio f_{B_s}/f_B against M_π^2/X_π^2 in Figure 7.4. Linear fit coefficients (top row) and quadratic fit coefficients (bottom row) are shown against the a^2 of the ensembles used in the fit when applicable. The functional forms of the fits are presented in Table 7.1.
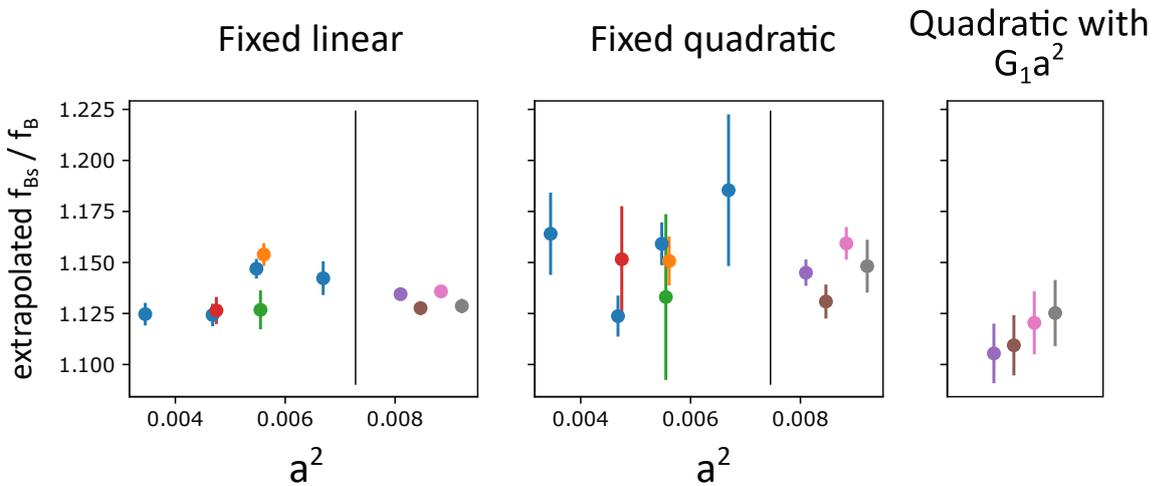
Figure 7.6: Extrapolated values corresponding to the physical point, for various fits to f_{B_s}/f_B. The extrapolated values correspond to the same fits as in Figure 7.3.

where $G_1 a^2$ is also included. When the extrapolation to the physical point is performed, we see that the fits including a^2 produce a lower expected value than the other fits. From the linear and quadratic fits at individual a^2, we see somewhat of a downward trend as a goes to zero that explains this behaviour in the fit including a^2, but we could also argue for a constant relationship between f_{B_s}/f_B and a^2 in these ensembles.

For further comparison with FLAG, we also need to consider systematic errors in the fit instead of just statistical error. In the next Section, we check our assumptions about the flavour breaking expansion and the tuning, and in Section 7.3.1, we begin to quantify systematic errors from the tuning and from correlator fits in the decay constant.

7.2 Expansions and flavour singlet quantities

The ratio f_B/f_{X_B} was shown to be approximately linear with the ensembles studied, as expected according to our flavour-breaking expansion where $\overline{m}$ is held constant. With this model, however, we also expect that other flavour-singlet quantities will be approximately constant as we break from the SU(3) symmetric point. We wish to test the limits of this assumption on these ensembles, and also check if we can see effects from the tuning or variation in lattice spacing.

As part of the tuning process, we choose the 'best' value of the b-quark tuning parameters m_0, c_P, and ζ such that our calculated value of the spin-averaged and flavour-averaged mass is held constant (see Section 6.3). More explicitly, we are tuning to

$$\overline{M}_{X_B} = \frac{1}{3}\left(2\overline{M}_{B_l} + \overline{M}_{B_s}\right) \tag{7.3}$$

$$= \frac{1}{12}\left(2M_{B_l} + 6M_{B_l^*} + M_{B_s} + 3M_{B_s^*}\right) \ . \tag{7.4}$$

This means that the B and B_s meson masses are not individually tuned or constrained. We need to check separately that the behaviour of M_B^2, $M_{B_s}^2$, and $X_B^2 = \frac{1}{3}(2M_B^2 + M_{B_s}^2)$ match our SU(3) breaking assumptions, namely that X_B^2 is approximately constant as we move toward the physical point and that the splitting in the B and B_s meson masses is approximately linear.

The splitting in the B meson masses, M_B^2/X_B^2, is shown in Figure 7.7 as an analogy to the decay constant splitting in Figure 7.2. The values of X_B^2 are also shown in Figure 7.8.

Comparing Figure 7.7 to Figure 7.2, we observe that the splitting in M_B^2/X_B^2 seems to have distinctly different slopes for each different lattice spacing and light
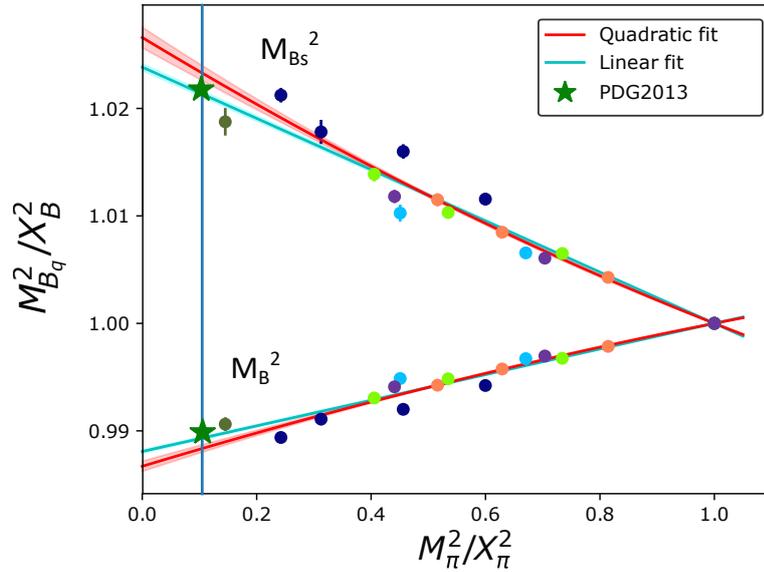
Figure 7.7: M_B^2/X_B^2 and $M_{B_s}^2/X_B^2$ against the light quark mass splitting M_π^2/X_π^2. The linear and quadratic fits shown are applied to all ensembles together.
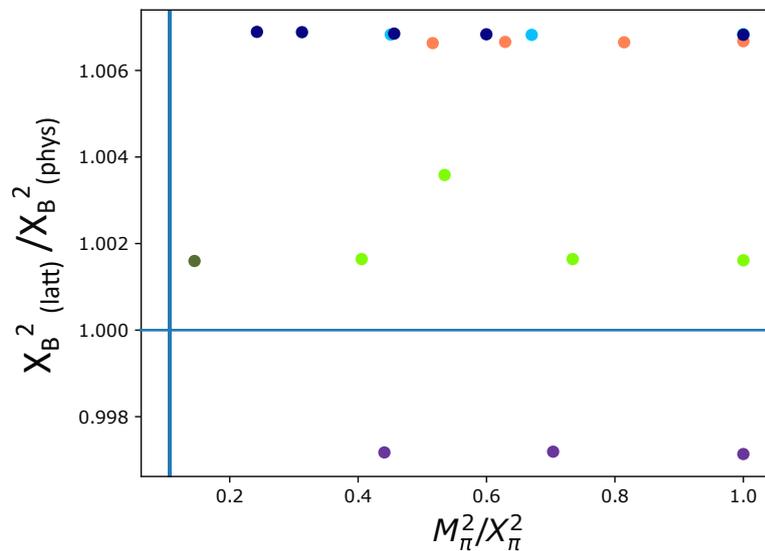
Figure 7.8: The flavour-singlet mass squared (X_B^2) on each ensemble, compared with the physical value of X_B^2 calculated from PDG results. We notice that the sets of ensembles with the same tuning star all have the same value of X_B^2.

quark trajectory. In particular, we see that the $\beta = 5.5$ ensemble with a different value of $\overline{m}$ (shown in navy) is particularly separated from the other trajectories with $\overline{m}$ closer to the physical.

We also observe in Figure 7.8 that the calculated value of X_B^2 is slightly different on each of the trajectories, but that the values are all within 0.5% of the physical value. The value of X_B^2 is approximately constant for each trajectory, matching our expectation from the flavour breaking expansion. The only exception to this rule is the κ_l=0.122130 ensemble for $\beta = 5.65$, which has a slightly different tuning star to the other $\beta = 5.65$ ensembles (see grey bands in Figure 6.10). This difference in X_B^2 shows us a clear indication of a small systematic error present in the tuning process that we have been somewhat able to mitigate by choosing the same tuning star for each ensemble.

The behaviour of f_B is expected to be less controlled by the tuning process than the B meson masses, but we nonetheless expect to see f_{X_B} approximately constant in the flavour breaking expansion from the SU(3) symmetric point toward the physical. The calculated flavour-singlet decay constant, f_{X_B}, is shown against the light quark splitting or against a^2 in the left and right sides of Figure 7.9 respectively.
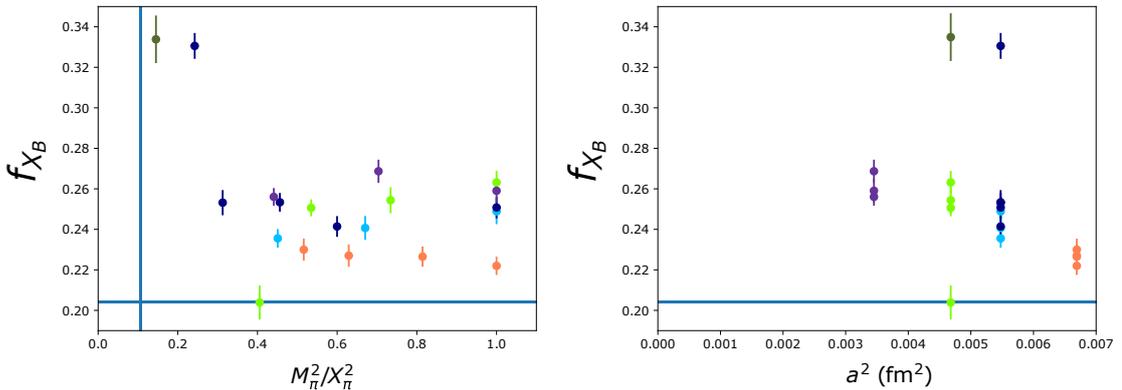
Figure 7.9: Calculated f_{X_B} for each ensemble, against the splitting M_π^2/X_π^2 (left) or against lattice spacing a^2 (right). The horizontal line shows the expected physical value of f_{X_B}.

Nearer to the SU(3) symmetric point, f_{X_B} appears to have an approximately constant value for each lattice spacing. Unfortunately, as we approach the physical point, this assumption is broken and we observe very different central values for f_{X_B}. We hypothesise that these differences might be related to challenges in calculating the decay constant as the light quark mass decreases. The problem cannot be resolved by selecting a new, combined fit window for all seven B and B_s meson pairs in the ensemble. Separate fit window limits for B_s and B_l, possibly also for each b in the ensemble, may be required to resolve this issue. The techniques presented in Section 6.2.1 and 6.2.2 for visualising the fits for multiple b quarks on each ensemble will be crucial for addressing this systematic in future studies.

Putting aside these outlier values of f_{X_B} for a moment, the rest of the ensembles show some linear behaviour of f_{X_B} with respect to the lattice spacing a^2. The decay constants calculated in this work, however, do not include the improvement term for the axial vector operator (see Equation 5.95) and these terms must be taken into account if direct calculations of f_B or f_{B_s} are to be made. In general, we expect the improvement term $c_1 \Phi_B^1$ to be larger at larger a^2, which may suggest our calculated f_{X_B} would remain larger than the world average.

Overall, we conclude that the general behaviour of the masses and decay constants matches with our expectations in terms of splitting relative to the SU(3) breaking. Further investigation is required into the behaviour of the decay constant, in particular the effect in f_{X_B} that causes the values to separate from their approximately-constant behaviour as we approach the physical point. In the next Section, we search for sources of systematic error that might produce this effect, including our fits to correlators and the tuning process itself.

7.3 Sources of systematic error

7.3.1 Propagating systematic error from fitting

Changing fits in the tuning

The best tuning of m_0, c_P, and ζ chosen by our tuning procedure is dependent on the correlator fits of each ensemble. As the fits are changed, small variations in the calculated masses, hyperfine splitting, or dispersion ratio will change the position of the interpolated best fit. In this Section, we consider changes to the 'best' set of tuning parameters, and observe the effect of this on our calculated SU(3) symmetry breaking of the decay constants.

In this short study, we choose to focus on changing the fits in the hyperfine splitting only, as this value is calculated from a single correlator fit rather than multiple. Additionally, we avoid changes to the fit of the B meson mass, as this value is also used directly in the calculation of the decay constant and thus cannot be used to quantify changes in the decay constant based on the tuning alone. Instead, we mimic the effect of more general changes in the best parameters by using the b quark at the centre value of the tuning star as another point of comparison. While this choice of tuning is somewhat unphysical, this is nonetheless suitable for a first overview of the systematic from changes in ζ in particular.

Returning to the discussion of hyperfine splitting, the fit window for the hyperfine splitting is chosen by default to be from $t = 5$ to $t = 14$. Based on the shape of this correlator (see Figure 6.4), for comparison we select an earlier fit window starting at $t = 1$, with the endpoint being chosen individually for each ensemble. This endpoint

must be chosen such that the fit is acceptable for all seven of the b quarks in the ensemble. While the χ^2 of fit can be somewhat useful for this task, we make the final decision using the grid of correlators (seen in Figure 6.5) so that we can ensure that the fit window matches the data well and doesn't extend into the noise region on any individual b.

An example of the effect of this changed fit on the calculated values for each b on a single ensemble, and thus on the overall tuning, is shown in Figure 7.10. The output tuning parameters themselves for two ensembles are shown in Figure 7.11.
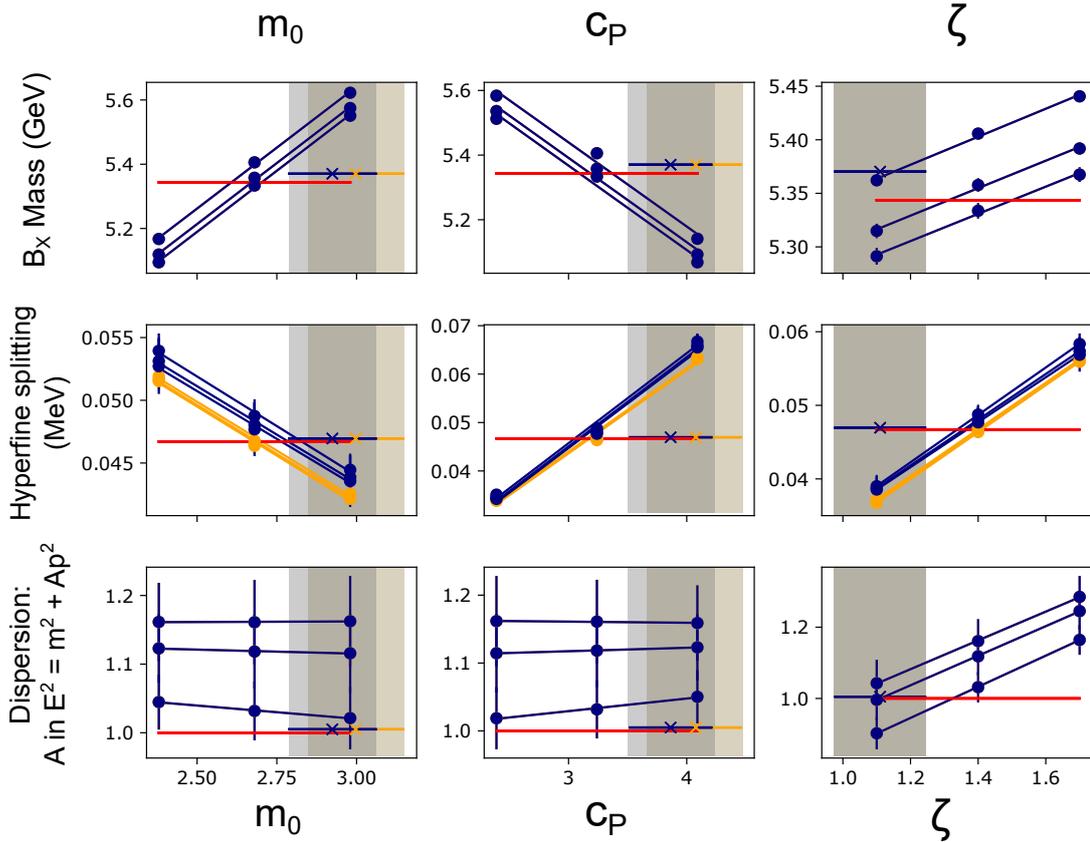
Figure 7.10: Plots of each physical quantity against each of the tuning parameters (m_0, c_P, ζ) for the $\beta = 5.5, (\kappa_l, \kappa_s) = (0.121145, 0.120413)$ ensemble, for the default tuning (blue) and the tuning with an earlier fit window on the hyperfine splitting (orange). The tuned best-fit parameters for each scenario are marked by an $\times$ and shaded band. We note that the fits for the mass and dispersion ratio remain unchanged between the blue and orange sets.

We observe that changing the fits for the hyperfine splitting does not affect the tuned value of the ζ parameter. On the $\kappa_l = 0.121145$ sample, we also notice that the change in the fit to the splitting has caused the best tuning to be just outside of the range of the tuning star, rather than just inside as with the default fit.

Now that we have changed the tuning, we wish to see how this propagates to the calculation of the decay constant. We recall that the decay constant for each ensemble is calculated on each bootstrap from a fit to the decay constant across
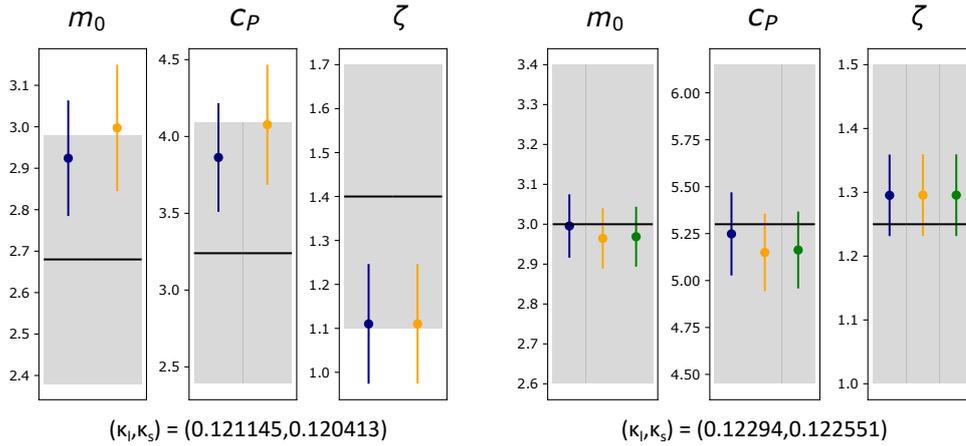
Figure 7.11: The outcome of the tuning process on the $\beta = 5.5, (\kappa_l, \kappa_s) = (0.121145, 0.120413)$ ensemble (left) and the $\beta = 5.8, (\kappa_l, \kappa_s) = (0.12294, 0.122551)$ ensemble (right) with changed fit windows for the hyperfine splitting. The default fit is shown in blue, with additional earlier fit windows shown in orange and green. The grey band shows the extent of the tuning star.

all 7 b quarks, using an interpolation (or extrapolation) to the value of the decay constant corresponding to the best tuning value on that bootstrap.

First, we consider the effect of the tuning on ratios of the decay constant. Figure 7.12 shows the effect on f_B/f_{X_B} of changing the tuning on each ensemble. The effect of using different fit windows for the hyperfine splitting and the effect of using the central value in the tuning star are shown side by side for comparison. We notice that the f_B/f_{X_B} are almost completely unaffected by the change in tuning. Even when the central value of the b tuning star is used, the values remain within the errorbars of the values calculated using the default fits. The behaviour of the ratio f_{B_s}/f_B, shown in Figure 7.13, is similarly unaffected by the change in tuning.

Unlike the decay constant ratios which can cancel systematics, we expect to see some systematic shifts in the value f_{X_B} as the tuned values of (m_0, c_P, ζ) are changed. The plot of f_{X_B} against M_π^2/X_π^2 is shown in Figure 7.14.

At the default tuning, there was already significant variation in the plot of f_{X_B} (see Figure 7.9), and we see that changing the tuning causes further variation. While the variations with respect to the earlier fits to the hyperfine splitting are small, f_{X_B} can be significantly affected by using the central value on the tuning star, particularly for the ensembles with lighter pion masses. In particular, the f_{X_B} value for the $\beta = 5.65$ $\kappa_l = 0.122167$ sample is of interest. At the default tuning, the f_{X_B} value is approximately 0.2, though using the central value on the tuning star yields a predicted value close to 0.5, removed from the plot to allow us to see other values.

Overall, the lack of change in either of the decay constant ratios leads us to believe that these ratios are independent of the b quark tuning in general, and thus
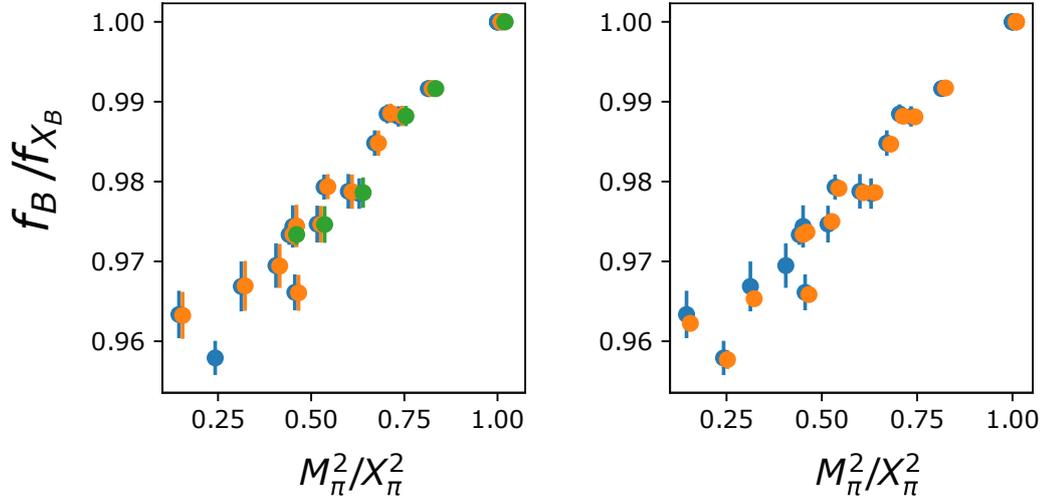
Figure 7.12: Comparison of values of f_B/f_{X_B} for different tunings based on selection of earlier fit windows for hyperfine splitting $B^* - B$ (left) or using the central value in the tuning star (right). The default tuning is shown in blue in both plots. In the left plot, two plateau lengths are tested on some ensembles.
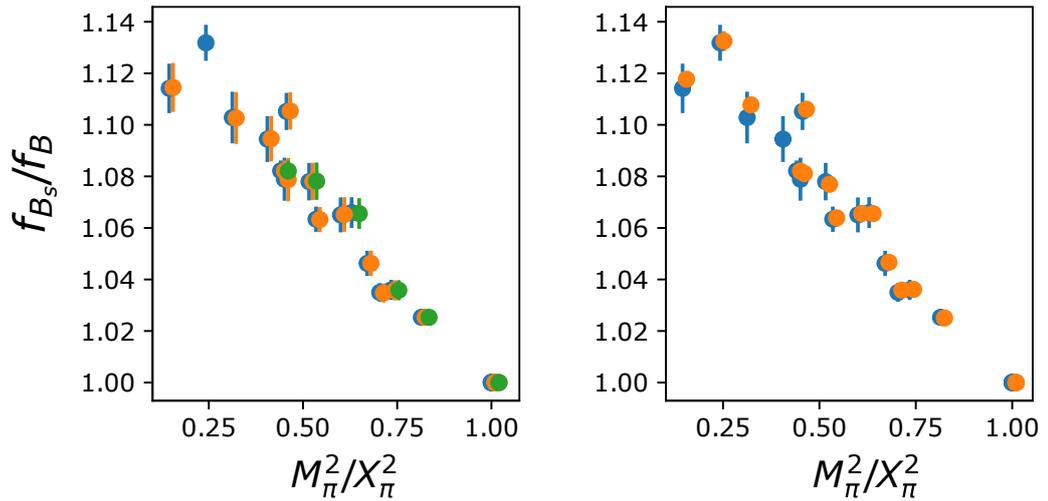
Figure 7.13: Comparison of values of f_{B_s}/f_B for different tunings based on selection of earlier fit windows for hyperfine splitting $B^* - B$ (left) or using the central value in the tuning star (right). The default tuning is shown in blue in both plots.
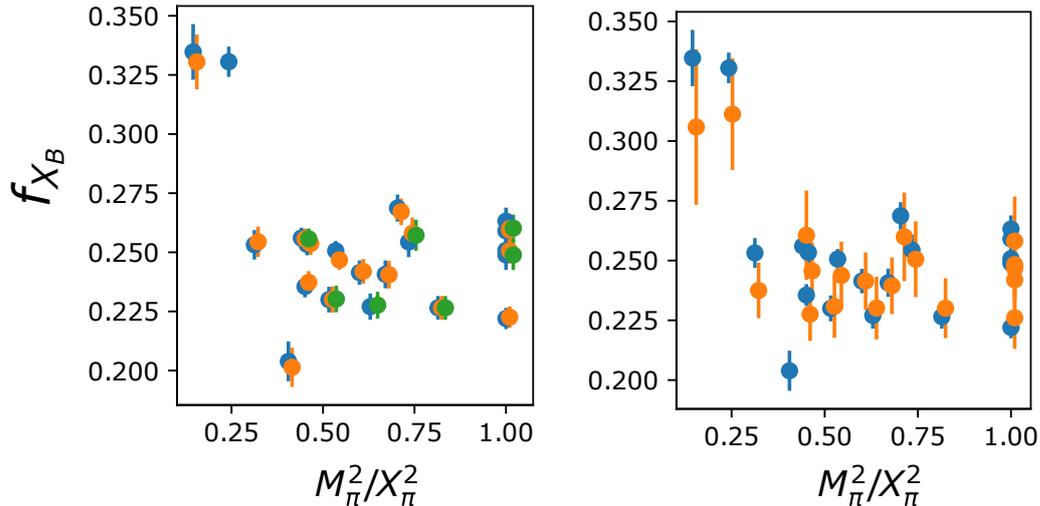
Figure 7.14: Comparison of values of f_{X_B} for different tunings based on selection of earlier fit windows for hyperfine splitting $B^* - B$ (left) or using the central value in the tuning star (right). The default tuning is shown in blue in both plots.

the impact of the tuning need not be considered. Similarly, the lack of change in f_{X_B} as a result of the fit window on the hyperfine splitting leads us to believe that the decay constant value is not particularly sensitive to small changes in m_0 or c_P directly. We do observe a lot of variation in f_{X_B} from using the central value of the tuning star, so large changes in m_0 and c_P and possibly any changes in ζ in the tuning are expected to change predictions of f_B or f_{B_s}. This will have to be examined further in future studies. Given the size of these effects, however, we also seek to measure the effect of changing the fits used directly in the calculation of the decay constant as a contribution to the systematic error calculated in this work.

Changing fits in the decay constant

While smeared-sink correlators are used to fit the masses of the mesons, the decay constant also requires fits to a point-sink correlator (see Section 5.4.3), and this merits further study. Unlike all other correlators used, the calculation of the decay constant is sensitive to the amplitude of the exponential fit rather than the exponent. In this study, we choose later fit windows staring at $t = 9$ or $t = 10$ to avoid contamination from excited states and try to ensure we are collecting a plateau, especially on the B_l meson. The results of using later fits to the decay constant is shown for f_B/f_{X_B} and f_{B_s}/f_B in Figure 7.15.

Unlike the changes for the tuning, the decay constant ratios are more strongly affected by the fit windows used to calculate the decay constant. This effect is larger for ensembles with lighter pion masses. Examining the fits for these samples, despite the plateau appearing to form later in general on all seven b, the extent of the plateau can vary when comparing different b on the same ensemble. The
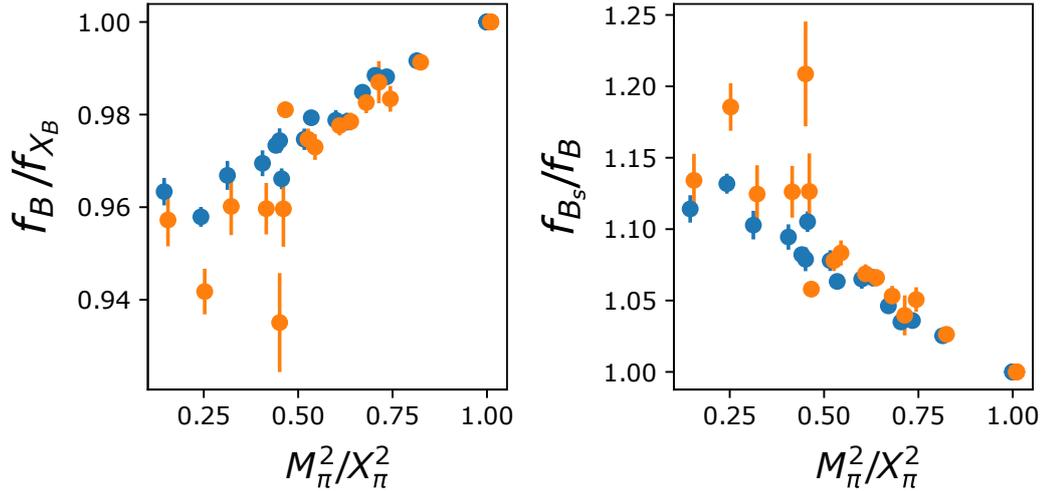
Figure 7.15: Comparison of values of f_B/f_{X_B} (left) or f_{B_s}/f_B (right) for different fits to the correlators in the decay constant. The default fit is shown in blue with the new, later fits in orange.

correlators for the $\beta = 5.5$, $\kappa_l = 0.121166$ ensemble are shown in Figure 7.16 as an example of this effect. We see from this Figure that at these lighter pion masses, the assumption that we can use the same fit window for the decay constant across all b quarks—and indeed, for both the B_l and B_s—is breaking down.

We also calculate f_{X_B} for each ensemble, shown in Figure 7.17.

Despite the changes in the central values of f_{X_B}, we observe that the sizes of the errorbars are similar between the two sets of fit windows for each ensemble. This indicates that the interpolation or extrapolation to the best tuning is about the same, and that the decay constants for each b quark also have approximately the same amount of error in most cases. Interestingly, while re-fitting the decay constant caused a lot of change in the ratios for pion masses close to the physical, the changes in f_{X_B} are not as extreme. Overall we see a general trend where the later fit causes the value of f_{X_B} to drop on almost all ensembles, indicating that this decay fit window is a strong source of systematic error if we wanted to make predictions for f_{B_s}/f_B, or f_B and f_{B_s} individually.

Extrapolating to the physical point

In order to try to quantify and compare the effects of changing the tuning or changing the correlator fits, we consider their effect on extrapolations of f_{B_s}/f_B to the physical point. The extrapolated values of f_{B_s}/f_B for each type of fit and for each of the four types of tunings or correlator fits examined are shown in Figure 7.18 and also in Table 7.3.

We observe that the χ^2/dof for the fits to the b at the centre of the tuning star are particularly large. This is because almost all of the statistical error in the ratio
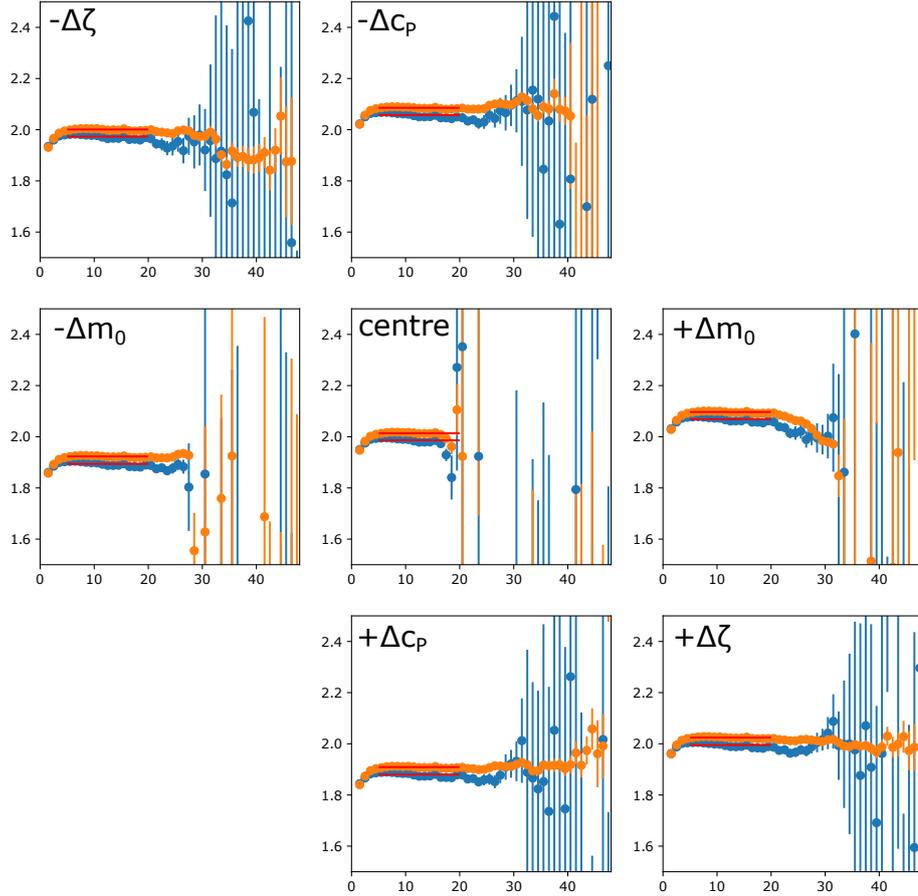
Figure 7.16: Correlator fits to the effective mass of the point sink $\mathcal{C}_A P$ correlator for the B_l (blue) and B_s (orange) mesons on the $(\kappa_l, \kappa_s)=(0.121166,0.120371)$ ensemble, shown for all 7 b quark parameters in the tuning 'star'. We notice that the plateau region for the central b is smaller than for the other parameter sets, and that the plateau for the B_s meson is longer than for the B_l meson for some of the b quarks.
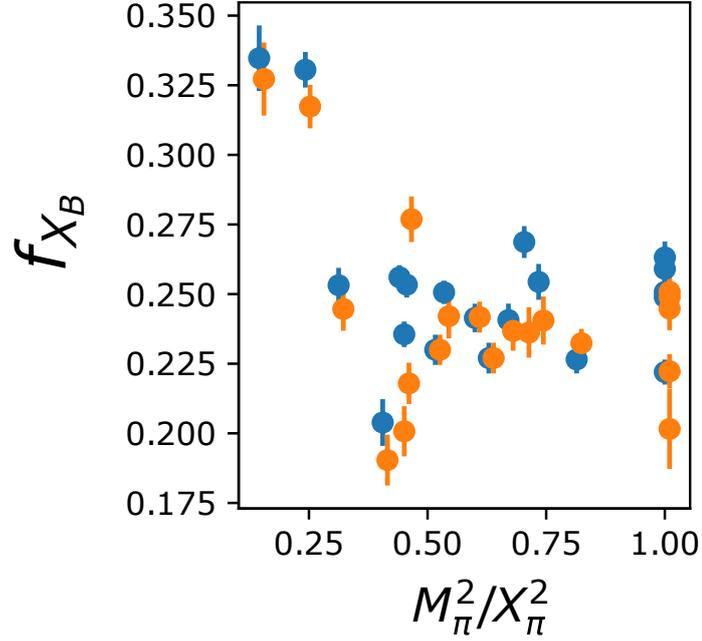
Figure 7.17: Comparison of values of f_{X_B} for different fits to the point sink corre-
lator. The default tuning is shown in blue.
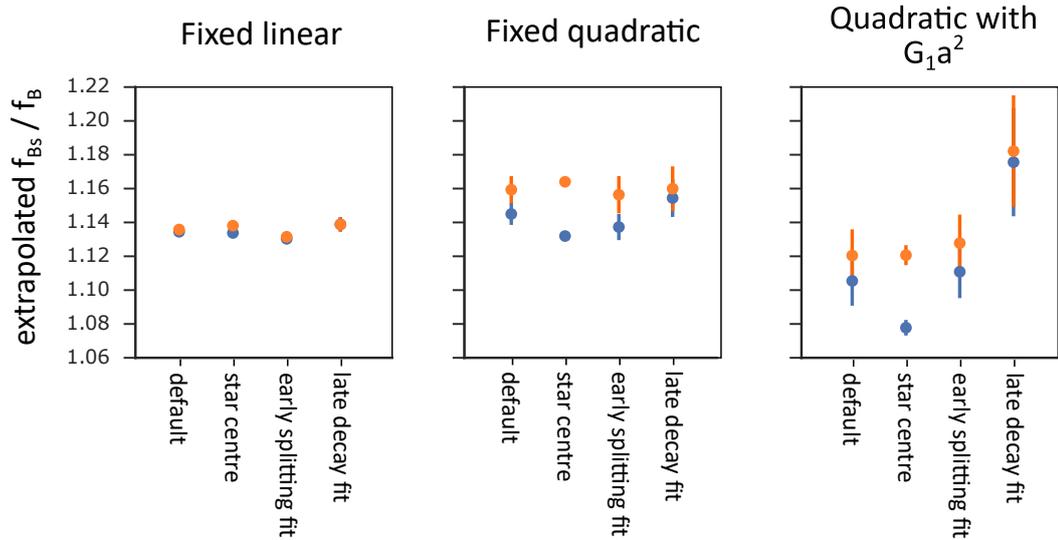
Figure 7.18: Extrapolated values corresponding to the physical point, for various
fits to f_{B_s}/f_B. All ensembles used in fit (blue), only ensembles with $m_\pi L > 4$ used
in fit (orange)

Table 7.3: Extrapolated values of the f_{B_s}/f_B ratio for different fit types. The fits for $m_\pi L > 4$ include all ensembles except the near-physical pion mass $\beta = 5.65$ ensemble. The 'interpolated star centre' results include statistical error from interpolating to the central b, while the 'star centre' results use only the central b correlators.

Data	Ratio fit type	Fit window sample	value at physical	χ^2/dof
All ensembles	Linear	default	1.134 ±0.003	1.8
		star centre	1.134 ±0.001	26.5
		intpl. star centre	1.135 ±0.003	1.7
		early splitting fit	1.130 ±0.003	1.5
		late decay fit	1.139 ±0.004	3.5
	Quadratic	default	1.145 ±0.006	1.8
		star centre	1.132 ±0.001	27.8
		intpl. star centre	1.145 ±0.006	1.6
		early splitting fit	1.137 ±0.008	1.6
		late decay fit	1.154 ±0.011	3.6
	Quadratic with a^2	default	1.105 ±0.015	1.3
		star centre	1.078 ±0.005	20.7
		intpl. star centre	1.112 ±0.013	1.2
		early splitting fit	1.111 ±0.016	1.4
		late decay fit	1.176 ±0.032	3.7
$m_\pi L > 4$	Linear	default	1.136 ±0.003	1.8
		star centre	1.138 ±0.001	22.4
		intpl. star centre	1.136 ±0.003	1.7
		early splitting fit	1.131 ±0.003	1.5
		late decay fit	1.139 ±0.004	3.7
	Quadratic	default	1.159 ±0.008	1.3
		star centre	1.164 ±0.002	16.5
		intpl. star centre	1.157 ±0.008	1.3
		early splitting fit	1.156 ±0.011	1.2
		late decay fit	1.160 ±0.013	3.7
	Quadratic with a^2	default	1.120 ±0.015	0.9
		star centre	1.121 ±0.006	13.4
		intpl. star centre	1.124 ±0.014	0.8
		early splitting fit	1.128 ±0.017	1.0
		late decay fit	1.182 ±0.033	3.9
FLAG 2016			1.201 ±0.016	

for each of the other cases comes from the interpolation to the best tuning. Without this interpolation, the error region for each point is much smaller and the best fit line will be many more standard deviations away from the central value for each ensemble. If we apply the statistical error from interpolation to the case for the central b, we now obtain a predicted physical fit and a χ^2 that is very similar to default fitting case.

We also notice that the near-physical point that is excluded from the $m_\pi L < 4$ fit has a strong effect on the final extrapolated value of f_{B_s}/f_B. This is somewhat unsurprising, as ensembles closest to the physical point are also furthest from the centre of the SU(3) expansion, and thus will have the greatest impact on the expected quadratic component of the fit. We can see that the extrapolated values with and without the near-physical point are much more similar for the case with the late decay fit, as using the late decay fit results in much more variation of the ensemble values in general (as seen in Figure 7.15).

Another interesting result is that relative to the simple quadratic fit, the physical prediction using the a^2 fit is larger in the case with the later decay fits, but smaller in all other cases. The coefficient of a^2, G_1, is negative and not consistent with 0 for all cases studied in this thesis except when the later decay fits are used. This change may be in part due to the larger values of f_{B_s}/f_B closer to the physical point, which are not equally distributed among all sets of ensembles.

7.4 Discussion of results

Our studies of the effect of tuning changes on the decay constant results conclude that small changes in the best value of m_0 or c_P do not significantly affect the results for the decay constant ratios calculated. For future work concerning individual values of f_B or f_{B_s}, however, systematics from tuning—particularly involving changes in ζ—must be studied quantitatively.

Fits to the correlators used in the decay constant directly are much more crucial as a systematic in calculations of the decay constant ratios. Particularly at lighter pion masses, plateaus occur later and the B_l and B_s mesons may require independent fit windows. Further investigation is required to make these results more robust, and taking a single fit to the ratio of correlators used in the decay constant should also be considered as an option in future to reduce sources of error.

In Table 7.4, we attempt to quantify sources of systematic error in our quadratic extrapolation of f_{B_s}/f_B discussed in this thesis.

Overall, we choose to take the quadratic fit with a^2 on all ensembles with the

Table 7.4: Summary of known sources of systematic error in calculation of f_{B_s}/f_B using the continuum extrapolation and quadratic extrapolation to the physical point. For a conservative estimate, errors are assumed to be uncorrelated with one another such that the total systematic is calculated in quadrature.

Source	-	+	Note
Z_V^{bb} value	0	0	Cancelled in ratio
Z_V^{ss}/Z_V^{ll}	0	0.023	2% systematic expected, see Section 6.4.1
Changes to b tuning	0.007	0.007	Difference between 'intpl. star centre' and 'default'
Fitting to ensembles with light pion masses	0.015	0.015	Difference between all ensembles and $m_\pi L > 4$ for 'default' fits
Correlator fits used in the decay constant	0.07	0.07	Difference between 'late decay fit' and 'default'
TOTAL SYSTEMATIC	**-0.071**	**+0.076**	

default fit window set-up as our final, central value result. We calculate

$$\frac{f_{B_s}}{f_B} = 1.105 \pm 0.015 \text{ (statistical)} \, ^{+0.076}_{-0.071} \text{ (systematic)}$$

as our final result in this thesis. The difference in size between the upper and lower systematic error in this calculation can be eliminated by explicitly calculating Z_V^{ss}/Z_V^{ll} on the lattice ensembles with lighter pion masses, and this is currently underway. Similarly, further studies of the correlator fits to the decay constant—especially on these lighter ensembles—will also further constrain the systematic error from fitting and improve our calculation of f_{B_s}/f_B overall.

7.5   Summary

In this Chapter, we have presented SU(3) breaking calculations for the B meson decay constant. The splitting between f_B/f_{X_B} and f_{B_s}/f_{X_B}, and the ratio f_{B_s}/f_B, are both presented across all ensembles. Multiple different fits are trialled in order to extrapolate to the physical pion mass.

Sources of systematic error were also studied. We qualitatively examined the SU(3) breaking expansion by considering the splitting between the mass of the B and B_s meson with this b quark action and tuning formalism. The impact of fits to correlators on the final extrapolation result was also assessed. We recorded a final, extrapolated result of $f_{B_s}/f_B = 1.105 \pm 0.015$ with purely statistical error from a quadratic fit expanded about (1,1) in f_{B_s}/f_B against M_π^2/X_π^2 with an additional extrapolation to the physical point. We also estimate an additional 7% error from

systematics related to tuning and fitting, which may include up to an additional 2% increase in the central value of f_{B_s}/f_B if Z_V^{ss}/Z_V^{ll} can be calculated for the lighter ensembles.

Chapter 8

Conclusion

8.1 Thesis summary

In this thesis we have explored the physics of flavour in both an experimental and theoretical context. Results are presented for B tagging studies using Belle II's Full Event Interpretation software designed for rare decays, and the SU(3) symmetry breaking in the B meson decay constant is investigated in detail using lattice QCD.

8.1.1 B meson tagging at Belle II

The Belle II detector and software environment is described. For rare decays, the Full Event Interpretation software is trained to reconstruct possible tag B mesons in a variety of channels. Testing and benchmarking of this software was performed for multiple Monte Carlo versions, and bugs were classified and identified. Skimming scripts to produce FEI tags for the collaboration were also developed. Preliminary studies of the semitauonic decay reconstruction relevant to the $R(D^{(*)})$ analysis were also performed to improve our understanding of the FEI when used by analysts.

8.1.2 b quarks for QCDSF

Tuning was performed to produce b quarks and B mesons on the lattice for a variety of $N_f = 2 + 1$ ensembles produced by the QCDSF collaboration [72]. A simple correlator fit method using the same fit window for each b quark was selected, though future possibilities of using a different fit selection regime were also discussed. The decay constants f_B and f_{B_s} were produced on each ensemble, and the renormalisation constants Z_V^{ll} and Z_V^{bb} were calculated.

Using the results of all ensembles, we examine SU(3) breaking effects and make extrapolated predictions of the SU(3) breaking at the physical point. We observe no strong a^2 dependence in the results for f_B/f_{X_B} or f_{B_s}/f_B. Neither of these ratios

have strong contributions to systematic error as a result of fits to the hyperfine splitting used in the tuning. Some effect of a change in the tuned value of ζ is observed. The strongest systematic is from the fits used directly in the calculation of the decay constant, and further studies of these fits will be required in future work.

Our final result for the decay constant ratio f_{B_s}/f_B is 1.105±0.015 in a continuum extrapolation with only statistical errors. We note that this central value may under-estimate f_{B_s}/f_B by up to 2 %, based on the systematic from Z_V^{ss}/Z_V^{ll} on these ensembles.

8.2 Future Work

8.2.1 Further b measurements on the lattice

The b quark tuning is one of the more intensive parts of any lattice study including the b quark. As this tuning has already been performed, a variety of future b studies on the QCDSF $N_f = 2 + 1$ ensembles are possible. Future studies of f_{B^*} and f_{B^*}/f_B could be made relatively quickly using the correlators already produced for this thesis. Further extensions to this work may include calculation of $B \to D^{(*)}\ell\nu$ form factors, or even calculation of the $B_c \to J/\psi\ell\nu$ form factors where even a first-look study is urgently needed to reduce theory uncertainty on recent LHCb measurements. Studies of B_c in particular would draw on existing tuning of the c quark also performed at the University of Adelaide for QCDSF. [82]

Extensions may also be made to studies of baryons. The Λ_b baryon is important to studies of $|V_{cb}|/|V_{ub}|$, and such a study would be a complement to existing light baryon studies at the University of Adelaide. Further lattice studies of this decay are in demand, as only one modern, unquenched simulation of the form factors was available at the time of the Belle II Physics Book [28].

8.2.2 Developing the FEI with Belle II data

Data collection using the full Belle II detector has begun this year, following the successful Phase II run in 2018. This Phase 3 data is already being used for calibration studies, including calibration of the FEI and FEI skims. Diagnostic scripts developed in this thesis will have continuing application to future studies, with new PhD students from the University of Adelaide continuing to work with the skimming group and FEI validation over the coming years. This next stage will be crucial for rare decay studies at Belle II, as the collaboration shifts to using skimmed data for all analyses.

Appendix A

Error in purity calculations

The purity is calculated from measured counts of events. As these two individual counts that make up the numerator and denominator are often small in the context of low-branching-ratio or poorly reconstructed channels, it becomes prudent to consider the Poisson counting error in both the numerator and denominator in order to calculate an error in the purity. This allows us to compare the purity across different Monte Carlo samples and determine whether the purity has changed even when low numbers of events are reconstructed.

A.1   Theoretical overview

We consider the calculation of the purity in the context of B meson reconstruction. Let us write the measured purity p as $p = n/N$, for n the number of correctly reconstructed B mesons, and N the total number of B mesons reconstructed. We wish to calculate the error in p, and we simplify this scenario to consider only a symmetric error region about a central value.

We take a Bayesian approach to this problem, and given that each reconstructed B is either correctly or incorrectly reconstructed, we can use binomial statistics to determine the correct probabilities. The probability that the purity is p given a sample measurement n/N is

$$\mathrm{P}\left(p|n \text{ of } N\right) = p^n \left(1 - p\right)^{N-n} \frac{(N + 1)!}{(N - n)!n!},\qquad\text{(A.1)}$$

so the mean sampled purity is

$$\bar{p} = \int_0^1 dp\; p\; \mathrm{P}\left(p|n \text{ of } N\right) \qquad\text{(A.2)}$$

$$= \frac{n + 1}{N + 2} \qquad\text{(A.3)}$$

APPENDIX A. ERROR IN PURITY CALCULATIONS

using binomial integration statistics and assuming a continuous distribution of values of true purity. This implies that our desired error σ_p can be written

$$\sigma_p = \sqrt{\int_0^1 dp \, (p - \bar{p})^2 \, \mathrm{P}\,(p|n \text{ of } N)} \tag{A.4}$$

$$= \sqrt{\int_0^1 dp \left(p - \frac{n+1}{N+2}\right)^2 p^n (1-p)^{N-n} \frac{(N+1)!}{(N-n)! \, n!}}. \tag{A.5}$$

Expanding the integral (labelled I), we can write,

$$I = \frac{(N+1)!}{(N-n)! \, n!} \left(\int_0^1 dp \, p^{n-2}(1-p)^{N-n} \right.$$
$$- \int_0^1 dp \, \frac{2(n+1)}{(N-2)} \times p^{n+1}(1-p)^{N-n} \tag{A.6}$$
$$\left. + \int_0^1 dp \, \left(\frac{n+1}{N+2}\right)^2 \times p^n (1-p)^{N-n} \right).$$

But we know that

$$\int_0^1 dp \, p^m (1-p)^n = \frac{n! \, m!}{(n+m+1)!}, \tag{A.7}$$

so I can be written

$$I = \left(\frac{(N+1)!}{(N-n)! \, n!} \right) \left(\frac{(n+2)!(N-n)!}{((N-n)+n+2+1)!} - \frac{2(n+1)}{(N+2)} \frac{(n+1)!(N-n)!}{((N-n)+n+1+1)!} \right.$$
$$\left. + \left(\frac{n+1}{N+2}\right)^2 \frac{n!(N-n)!}{(N+1)!} \right) \tag{A.8}$$

$$I = \left(\frac{(N+1)!}{n!} \right) \left(\frac{(n+2)!}{(N+3)!} - 2\frac{(n+1)(n+1)!}{(N+2)(N+2)!} + \left(\frac{n+1}{N+2}\right)^2 \frac{n!}{(N+1)!} \right) \tag{A.9}$$

$$\dots \tag{A.10}$$

$$= \frac{n+1}{N+2} \left(\frac{n+2}{N+3} - \frac{n+1}{N+2} \right), \tag{A.11}$$

such that the final result for the symmetric error in the purity p is written

$$\sigma_p = \sqrt{\frac{n+1}{N+2} \left(\frac{n+2}{N+3} - \frac{n+1}{N+2} \right)}. \tag{A.12}$$

Notably, if we consider $n \to \infty$ and $(N-n) \to \infty$ on our Bayesian result, we

obtain the Frequentist prediction for the error, which is

$$\sigma_p = \sqrt{\frac{p(1-p)}{N}} \tag{A.13}$$

$$= \sqrt{\frac{n(N-n)}{N^3}} \tag{A.14}$$

Appendix B

The Bootstrap Method

For a single sample of N elements x drawn from a true population, the sample mean $\overline{x}$ and sample standard deviation s_x can be used to approximate the mean μ and standard deviation σ of the full population. The error in the sample mean can be calculated

$$\Delta\bar{x} = \frac{s_x}{\sqrt{N}} \tag{B.1}$$

using a non-standard notation for the error in the sample mean ($\Delta\overline{x}$ instead of $s_{\overline{x}}$) to help distinguish it from the standard deviation of the sample.

In a more complicated situation, we may have a set of N collections (or elements) $x(t)$, and wish to make a fit with respect to discrete t. To make the best fit, we want to use the calculated mean and sample standard deviation of x at each position t as input to the fit, but this leaves us only a single fit value (say, f) for the entire set of N collections. Instead, we will need a way to produce multiple fit values, such that we can calculate $\overline{f}$ and s_f. Suppose each of these N collections $x(t)$ is labelled $x_i(t)$ for $i = 1...N$. If we assume that N is sufficiently large, and that these collections are independent and randomly drawn from the population, we are able to model collecting new samples from the population by creating new samples using the elements of our original sample of N elements. For the bootstrap method, these new samples are created by randomly selecting elements from our original sample with replacement to create new samples each consisting of N elements. For example, from our original sample $[x_1(t), x_2(t), ...x_N(t)]$ we might produce one new sample with elements $[x_4(t), x_1(t), x_N(t), x_1(t), ...]$.

From M bootstrap samples consisting of N elements, we then calculate f on each sample to produce a set of M values for f. We can then easily determine $\overline{f}$ and s_f, where s_f can also be used as the value for σ_f. A key property of the bootstrap method is that it can be used for quantities for which the error cannot otherwise be calculated, including fits (in this example), statistical measures such as the median, and any process where the final result is the outcome of multiple steps. For most

cases, $M > 100$ is sufficiently many bootstraps to ensure a good estimate, provided our original sample is sufficiently large and representative of the population [83].

We can also contrast the bootstrap with the jackknife approach. Rather than randomly resampling our N elements $x(t)$, instead we systematically remove one element at a time from the sample, to produce N new samples each consisting of $N-1$ elements. Because of the similarity of each jackknife sample to the original distribution, however, the jackknife may not perform as well as the bootstrap in determining the true error of quantities that are 'unsmooth', such as the median.

Appendix C

Publications

C.1   Journal publications

C.1.1   The Full Event Interpretation

Statement of Authorship

Title of Paper	The Full Event Interpretation: an exclusive tagging algorithm for the Belle II experiment
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Published in 'Computing and Software for Big Science', First online 25 February 2019 DOI 10.1007/s41781-019-0021-8

Principal Author

Name of Principal Author (Candidate)	Sophie Hollitt
Contribution to the Paper	Contribution to ongoing validation and calibration studies of the FEI. Debugging of FEI use in analysis software.
Overall percentage (%)	5
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.
Signature	Date 19/08/2019

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

    i.   the candidate's stated contribution to the publication is accurate (as detailed above);

    ii.   permission is granted for the candidate in include the publication in the thesis; and

    iii.   the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	A/Prof Paul Jackson (Principal Supervisor)
Contribution to the Paper	The candidate played an important role in the understanding of the software vital to this publication.
Signature	Date 19/08/2019

Name of Co-Author	
Contribution to the Paper	
Signature	Date

Please cut and paste additional co-author panels here as required.

162

ORIGINAL ARTICLE

The Full Event Interpretation

An Exclusive Tagging Algorithm for the Belle II Experiment

T. Keck[1] · F. Abudinén[2] · Florian U. Bernlochner[1] ⬚ · R. Cheaib[3] · S. Cunliffe[4] · M. Feindt[1] · T. Ferber[4] · M. Gelb[1] ·
J. Gemmler[1] · P. Goldenzweig[1] · M. Heck[1] · S. Hollitt[5] · J. Kahn[6] · J.-F. Krohn[7] · T. Kuhr[6] · I. Komarov[4] · L. Ligioi[2] ·
M. Lubej[8] · F. Metzner[1] · M. Prim[1] · C. Pulvermacher[1] · M. Ritter[6] · J. Schwab[1] · W. Sutcliffe[1] · U. Tamponi[9] · F. Tenchini[4] ·
N. E. Toutounji[10] · P. Urquijo[7] · D. Weyland[1] · A. Zupanc[8]

Abstract

The full event interpretation is presented: a new exclusive tagging algorithm used by the high-energy physics experiment
Belle II. The experimental setup of Belle II allows the precise measurement of otherwise inaccessible *B* meson decay modes.
The Full Event Interpretation algorithm enables many of these measurements. The algorithm relies on machine learning
to automatically identify plausible *B* meson decay chains based on the data recorded by the detector. Compared to similar
algorithms employed by previous experiments, the Full Event Interpretation provides a greater efficiency, yielding a larger
effective sample size usable in the measurement.

Introduction

The Belle II experiment is located at the SuperKEKB elec-
tron–positron collider in Tsukuba, Japan, and was commis-
sioned in 2018. The experiment is designed to perform a

✉ T. Keck
   thomas.keck2@kit.edu

✉ Florian U. Bernlochner
   florian.bernlochner@kit.edu

[1] Karlsruhe Institute of Technology, Campus Süd, Institut für
   Experimentelle Teilchenphysik, Wolfgang-Gaede-Str. 1,
   76131 Karlsruhe, Germany

[2] Max-Planck-Institut für Physik, Munich, Germany

[3] University of Mississippi, Mississippi, MS, USA

[4] Deutsches Elektronen-Synchrotron, Hamburg, Germany

[5] University of Adelaide, Adelaide, Australia

[6] Ludwig Maximilians Universität, Munich, Germany

[7] University of Melbourne, Melbourne, Australia

[8] Jožef Stefan Institute, Ljubljana, Slovenia

[9] INFN - Sezione di Torino, Turin, Italy

[10] University of Sydney, Sydney, Australia

wide range of high-precision measurements in all fields
of heavy flavor physics, in particular it will investigate the
decay of *B* mesons [1]. For this purpose, the experiment is
expected to record about 40 billion collision events each
containing an Υ(4S) resonance, which at least 96% of the
time decays into exactly **two** *B* mesons (a $B\bar{B}$ pair). Each
B meson decays via various intermediate states into a set
of final-state particles, which are considered stable in the
Belle II detector. In general, charged final-state particles
are reconstructed as tracks in the central drift chamber and
in the inner silicon-based vertex detectors, whereas neutral
final-state particles are reconstructed as energy depositions
(called clusters) in the electromagnetic calorimeter. The
entire experimental setup of the detector and the collider is
described in more detail in Doležal and Uno [1].

The measurement of the branching fraction of rare decays
like $B \to \tau \nu_\tau$, $B \to K \nu \nu$ or $B \to l \nu \gamma$, with undetectable neu-
trinos in their final states, is challenging. However, the sec-
ond *B* meson in each event can be used to constrain the
allowed decay chains. This general idea is known as **tagging**.
Conceptually, each Υ(4S) event is divided into two sides:
the signal-side containing the tracks and clusters compatible
with the assumed signal B_{sig} decay the physicist is interested
in, e.g., a rare decay like $B \to \tau \nu$; and the tag-side containing

⧂ Springer

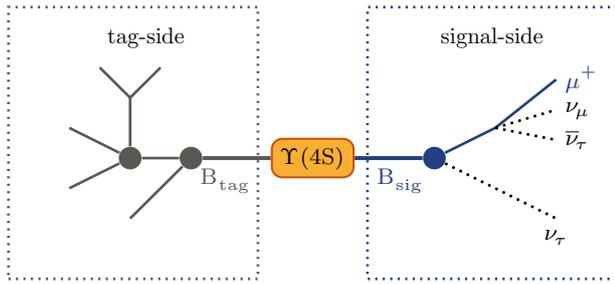**Fig. 1** Schematic overview of a $\Upsilon(4S)$ decay: (Left) a common tag-side decay $B_{tag}^- \to D^0(\to K_s^0(\to \pi^-\pi^+)\pi^-\pi^+)\pi^-$ and (right) a typical signal-side decay $B_{sig}^+ \to \tau^+(\to \mu^+\nu_\mu\bar{\nu}_\tau)\nu_\tau$. The two sides overlap spatially in the detector, therefore the assignment of a measured track to one of the sides is not known a priori

the remaining tracks and clusters compatible with an arbitrary B_{tag} meson decay. Figure 1 depicts this situation.

The initial four momentum of the produced $\Upsilon(4S)$ resonance is precisely known and no additional particles are produced in this primary interaction. Therefore, because of the relevant quantum numbers conservation, knowledge about the properties of the tag-side B_{tag} meson allows one to recover information about the signal-side B_{sig} meson which would otherwise be inaccessible. Most importantly, all reconstructed tracks and clusters which are not assigned to the B_{tag} mesons must be compatible with the signal decay of interest.

Ideally, a full reconstruction of the entire event has to take all reconstructed tracks and clusters into account to attain a correct interpretation of the measured data. The **Full Event Interpretation** (`FEI`) algorithm presented in this article is a new exclusive tagging algorithm developed for the Belle II experiment, embedded in the Belle II Analysis Software Framework (basf2) [2]. The `FEI` automatically constructs plausible B_{tag} meson decay chains compatible with the observed tracks and clusters, and calculates for each decay chain the probability of it correctly describing the true process using gradient-boosted decision trees. "Exclusive" refers to the reconstruction of a particle (here the B_{tag}) assuming an explicit decay channel.

Consequently, exclusive tagging reconstructs the B_{tag} independently of the B_{sig} using either **hadronic** or **semileptonic** B meson decay channels. The decay chain of the B_{tag} is explicitly reconstructed and therefore the assignment of tracks and clusters to the tag-side and signal-side is known.

In the case of a measurement of an exclusive branching fraction like $B_{sig} \to \tau\nu_\tau$, the entire decay chain of the $\Upsilon(4S)$ is known. As a consequence, all tracks and clusters measured by the detector should be already accounted for. In particular, the requirement of no additional tracks, besides the ones used for the reconstruction of the $\Upsilon(4S)$, is an extremely

powerful and efficient way to remove most reducible[1] backgrounds. This requirement is called the **completeness constraint** throughout this text.

In the case of a measurement of an inclusive branching fraction like $B_{sig} \to X_u l\nu$, all remaining tracks and clusters, besides the ones used for the lepton l and the B_{tag} meson, are identified with the X_u system. Hence, the branching fraction can be determined without explicitly assuming a decay chain for the X_u system.

The performance of an exclusive tagging algorithm depends on the **tagging efficiency** (i.e., the fraction of $\Upsilon(4S)$ events which can be tagged), the **tag-side efficiency** (i.e., the fraction of $\Upsilon(4S)$ events with a correct tag) and on the quality of the recovered information, which determines the **tag-side purity** (i.e., the fraction of the tagged $\Upsilon(4S)$ events with a correct tag) of the tagged events.

The exclusive tag typically provides a pure sample (i.e., purities up to 90% are possible). But, this approach suffers from a low tag-side efficiency, just a few percent, since only a tiny fraction of the B decays can be explicitly reconstructed due to the large amount of possible decay channels and their high multiplicity. The imperfect reconstruction efficiency of tracks and clusters further degrades the efficiency.

Both the quality of the recovered information and the systematic uncertainties depend on the decay channel of the B_{tag}, therefore we distinguish further between hadronic and semileptonic exclusive tagging.

Hadronic tagging considers only hadronic B decay chains for the tag-side [3, Section 7.4.1]. Hence, the four momentum of the B_{tag} is well-known and the tagged sample is very pure. A typical hadronic B decay has a branching fraction of $\mathcal{O}(10^{-3})$. As a consequence, hadronic tagging suffers from a low tag-side efficiency and can only be applied to a tiny fraction of the recorded events. Large combinatorics of high-multiplicity decay channels further complicate the reconstruction and require tight selection criteria.

Semileptonic tagging considers only semileptonic $B \to Dl\nu$ and $B \to D^*l\nu$ decay channels [3, Section 7.4.2]. Due to the presence of a high-momentum lepton, these decay channels can be easily identified and the semileptonic tagging usually yields a higher tag-side efficiency compared to hadronic tagging due to the large semileptonic branching fractions. On the other hand, the semileptonic tag will miss kinematic information due to the neutrino in the final state of the decay. Hence, the sample is not as pure as in the hadronic case.

To conclude, the `FEI` provides a hadronic and semileptonic tag for $B^\pm$ and B^0 mesons. This enables the measurement of exclusive decays with several neutrinos and

---

[1] Reducible background has distinct final-state products from the signal.
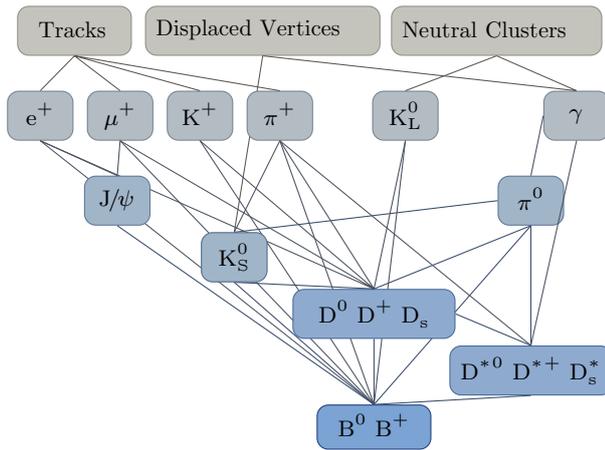
Fig. 2 Schematic overview of the FEI. The algorithm operates on objects identified by the reconstruction software of the Belle II detectors: charged tracks, neutral clusters and displaced vertices. In six distinct stages, these basics objects are interpreted as final-state particles (e^+, μ^+, K^+, π^+, K_L^0, γ) combined to form intermediate particles (J/ψ, π^0, K_s^0, D, D^*) and finally form the tag-side B mesons

inclusive decays. In both cases, the FEI provides an explicit tag-side decay chain with an associated probability.

Methods

The FEI algorithm follows a hierarchical approach with six stages, visualized in Fig. 2. Final-state particle candidates are constructed using the reconstructed tracks and clusters, and combined to intermediate particles until the final B candidates are formed. The probability of each candidate to be correct is estimated by a multivariate classifier. A multivariate classifier maps a set of input features (e.g., the four momentum or the vertex position) to a real-valued output, which can be interpreted as a probability estimate. The multivariate classifiers are constructed by optimizing a loss function (e.g., the misclassification rate) on Monte Carlo simulated $\Upsilon(4S)$ events and are described later in detail.

All steps in the algorithm are configurable. Therefore, the decay channels used, the cuts employed, the choice of the input features, and hyper-parameters of the multivariate classifiers depend on the configuration. A more detailed description of the algorithm and the default configuration can be found in Keck [4] and in the following we give a brief overview over the key aspects of the algorithm.

Combination of Candidates

Charged final-state particle candidates are created from tracks assuming different particle hypotheses. Neutral final-state particle candidates are created from clusters and displaced vertices constructed by oppositely charged tracks.

Each candidate can be correct (signal) or wrong (background). For instance, a track used to create a π^+ candidate can originate from a pion traversing the detector (signal), from a kaon traversing the detector (background) or originates from a random combination of hits from beam background (also background).

All candidates available at this stage are combined to intermediate particle candidates in the subsequent stages, until candidates for the desired B mesons are created. Each intermediate particle has multiple possible decay channels, which can be used to create valid candidates. For instance, a B^- candidate can be created by combining a D^0 and a π^- candidate, or by combining a D^0, a π^- and a π^0 candidate. The D^0 candidate could be created from a K^- and a π^+, or from a K_s^0 and a π^0.

The FEI reconstructs more than 100 explicit decay channels, leading to $\mathcal{O}(10000)$ distinct decay chains.

Multivariate Classification

The FEI employs multivariate classifiers to estimate the probability of each candidate to be correct, which can be used to discriminate correctly identified candidates from background. For each final-state particle and for each decay channel of an intermediate particle, a multivariate classifier is trained which estimates the signal probability that the candidate is correct. To use all available information at each stage, a network of multivariate classifiers is built, following the hierarchical structure in Fig. 2.

For instance, the classifier for the decay of $B^- \rightarrow D^0\pi^-$ would use the signal probability of the D^0 and π^- candidates, as input features to estimate the signal probability of the B^- candidate created by combining the aforementioned D^0 and π^- candidates.

Additional input features of the classifiers are the kinematic and vertex fit information of the candidate and its daughters. The multivariate classifiers used by the FEI are trained on Monte Carlo simulated events. The training is fully automatized and distributed using a map reduce approach [5]. Monte Carlo simulated data used to train the FEI is partitioned. At each reconstruction stage, the partitioned data is distributed to nodes where the reconstruction is performed and training datasets are produced (the mapping stage). The reduction stage consists of merging the training datasets and training multivariate classifiers with these training datasets.

The available information flows from the data provided by the detector through the intermediate candidates into the final B meson candidates, yielding a single number which can be used to distinguish correctly from incorrectly identified B_{tag} mesons. The process is visualized in Fig. 2. This allows one to tune the trade-off between tag-side efficiency and tag-side purity of the algorithm by requiring a minimal

signal probability. By contrast, most exclusive measurements by Belle, which used the previous FR algorithm, chose a working point near the maximum tag-side efficiency as described in "Previous work" section.

Combinatorics

It is not feasible to consider all possible B meson candidates created by all possible combinations. The amount of possible combinations scales with the factorial in the number of tracks and clusters. This problem is known as **combinatorics** in high-energy physics. Furthermore, it is not worthwhile to consider all possible B meson candidates, because all of them are wrong except for two in the best-case scenario.

The FEI uses two sets of the so-called **cuts**. A cut is a criterion that a candidate has to fulfill to be considered further. For instance, one could demand that the beam-constrained mass of the B meson candidate is near the nominal mass 5.28 GeV of a B meson particle, or that a μ^+ candidate has a high muon particle identification likelihood, which combines sub-detector information to identify muons.

Directly after the creation of the candidate (either from a track/cluster, or by combining other candidates), but before the application of the multivariate classifier, the FEI uses loose and fast **pre-cuts** to remove wrongly identified candidates (background), without losing signal. The main purpose of these cuts is to save computing time and to reduce the memory consumption. These **pre-cuts** are applied separately for each decay channel.

At first, a very loose fixed cut is applied on a quantity which is fast to calculate, e.g., the energy for photons, the invariant mass for D mesons, the energy released in the decay for D^* mesons, or the beam-constrained mass for hadronic B mesons. Second, the remaining candidates are ranked according to a quantity, which is fast to calculate (usually the same quantity as above is used here). Only the n (usually between 10 and 20) best candidates in each decay channel are further considered, the others are discarded. This best candidate selection ensures that each decay channel and each event receives roughly the same amount of computing time.

Next, the computationally expensive parts of the reconstruction are performed on each candidate: the matching of the reconstructed candidates to the generated particles (in case of simulated events), the vertex fitting, and the multivariate classification.

After the multivariate classifiers have estimated the signal probability of each candidate, the candidates of different decay channels can be compared. Here the FEI uses tighter **post-cuts** to aggressively remove incorrectly reconstructed candidates using all the available information. The main purpose of these cuts is to restrict the number of candidates per particle to a manageable number.

At first, there is a loose fixed cut on the signal probability, to remove unreasonable candidates. Second, the remaining candidates are ranked according to their signal probability. Only the m (usually between 10 and 20) best candidates of the particle (i.e., over all decay channels) are further considered, the others are discarded. This best candidate selection ensures that the amount of candidates produced in the next stage is reasonably low and can be handled by the computing system.

Performance

Applying the FEI to $\mathcal{O}(1$ billion$)$ events is a CPU-intensive task. An optimized runtime and a small memory footprint are key for a practical application and to save computing resources. The FEI spends most CPU time on vertex fitting (38%), particle combination (27%), and classifier inference (15%). All three tasks have been carefully optimized.

The FEI uses only a fast and simple unconstrained vertex fit during the reconstruction, and feeds the calculated information into its multivariate classifiers. The user can refit the whole decay chain of the final B candidates, including mass and/or interaction point profile constraints if desired. A dedicated fitter (called FastFit) based on a Kalman Filter [6] was implemented for the FEI, which requires drastically less computing time than the default implementation used by Belle II and yields very similar results. Due to this fitter, an overall speedup of the FEI of 2.74 was observed. The FastFit code is licensed under GPLv3 and available on GitHub [7].

As explained in "Combinatorics" section, the number of candidates which have to be processed scales as the factorial of the multiplicity of the channel. In previous approaches, the runtime and the maximum memory consumption was dominated by a few high-multiplicity events and tight cuts had to be applied to high-multiplicity channels. By contrast, the FEI addresses the combinatorics problem by performing best candidate selections during the reconstruction of the decay chain instead of fixed cuts. As a consequence, for each event and each decay channel, the FEI processes the same number of candidates in vertex fitting and classifier inference, i.e., consumes similar amounts of CPU time. Moreover, the maximum memory consumption is limited due to the fixed number of best candidates per event, which is a key requirement for using the computing infrastructure.

Finally, the FEI uses FastBDT [8], a gradient-boosted decision tree (BDT) implementation, as its default multivariate classification algorithm. The algorithm was originally designed for the FEI to speed up the training and application phase. Compared to other popular BDT implementations such as those provided by TMVA [9], SKLearn [10] and XGBoost [11], it originally improved the execution time by more than one order of magnitude, both in training

and application. In addition, an improved classification quality was observed. Most of the time when using `FastBDT` is spent during the extraction of the necessary features, therefore no further significant speedups can be achieved by employing a different method.

Automatic Reporting

The `FEI` includes an automatic reporting system called `Full Event Interpretation Report (FEIR)`.

The `FEIR` contains efficiencies and purities for all particles and decay channels at different points during the reconstruction. Individual reports containing control plots for each multivariate classifier and input variables are also automatically created. Control plots include receiver operating characteristics (ROC) curves, which show the tag-side efficiency against purity. Additionally, for each classifier, the purity is plotted as a function of classifier output, to check for a linear relationship as this confirms the classifier output can be treated as a probability. This built-in monitoring capability upgrades the `FEI` from a black-box to a white-box algorithm, which the user can understand and inspect on all levels of reconstruction.

Previous Work

Previous experiments have already developed and successfully employed tagging algorithms. To compare the algorithms, the maximal achievable tag-side efficiency is of particular interest, because it is directly related to the signal selection efficiency of the measurement. On the other hand, the achievable tag-side purity is only of limited use, because the achievable final purity of the final selection used for the measurement is dominated by the completeness constraint. Hence, most of the incorrect tags can be easily discarded and the final purity depends strongly on the considered signal decay channel. Moreover, signal-side independent ROC curves are not available for most of the previously employed algorithms. The area under the ROC curve allows one to compare the performance of the tagging algorithms.

The BaBar experiment [12] used the `Semi-Exclusive B reconstruction (SER)` algorithm for hadronic tagging [3, Section 7.4.1.1]. The algorithm used exclusive D and D^* mesons candidates as a seed, and combined those with up to five charmless hadrons to form a B_{tag} without assuming an exclusive B decay mode. The tag-side efficiency and tag-side purity of each B decay chain was extracted by fitting the beam-constrained mass [3, Section 7.1.1.2] spectrum of the constructed B_{tag} meson candidates. The beam-constrained mass is defined as $M_{\text{bc}} = \sqrt{E_{\text{beam}}^2/c^4 - p_B^2/c^2}$ where p_B denotes the three momentum of the reconstructed

B meson candidate and E_{beam} denotes half of the centre-of-mass energy of the colliding electron–positron pair. The maximum hadronic tag-side efficiency achieved by this algorithm was 0.2% for $B^0\bar{B}^0$ and 0.4% for B^+B^-, with a tag-side purity around 30%. The tag-side purity could be further increased by rejecting B meson candidates from low-purity decay chains. The semileptonic tag was usually constructed by combining an exclusive D or D^* meson with a lepton. The maximum semileptonic tag-side efficiency was typically 0.3% for $B^0\bar{B}^0$ and 0.6% for B^+B^- with an unknown tag-side purity.

The Belle experiment [13] used the so-called `Full Reconstruction (FR)` algorithm [14] for hadronic tagging [3, Section 7.4.1.2]. The `FR` introduced an hierarchical approach, which is still used by its successor and is presented in this article (see "Methods" section). The tag-side efficiency and tag-side purity was extracted by fitting the beam-constrained mass spectrum of the constructed B_{tag} meson candidates. The maximum hadronic tag-side efficiency achieved by this algorithm was 0.18% for $B^0\bar{B}^0$ and 0.28% for B^+B^-, with a tag-side purity around 10%. Multivariate classifiers [15] were used to estimate the signal probability of each candidate. The tag-side purity could be further increased by requiring a minimal signal probability. Variants of the FR were used for semileptonic tagging (see [16, 17]). The maximum semileptonic tag-side efficiency was 0.31% for $B^0\bar{B}^0$ and 0.34% for B^+B^-, with a typical tag-side purity of 5%.

Compared to the previously employed algorithms, the `FEI` provides a greater tagging and tag-side efficiency, with a equal or better tag-side purity. The improvements with respect to the FR can be attributed equally to the additional

Table 1 Summary of the maximum tag-side efficiency of the Full Event Interpretation and for the previously used exclusive tagging algorithms

	$B^{\pm}$ (%)	B^0 (%)
Hadronic		
`FEI` with FR channels	0.53	0.33
`FEI`	0.76	0.46
FR	0.28	0.18
SER	0.4	0.2
Semileptonic		
`FEI`	1.80	2.04
FR	0.31	0.34
SER	0.3	0.6

For the `FEI` simulated data from the last official Monte Carlo campaign of the Belle experiment were used. The maximum tag-side efficiency on recorded data is lower (see "Hadronic tag" section). The numbers for the older algorithms (see "Previous work" section), are not directly comparable due to different selection criteria, like best candidate selections and selections to suppress non-$\Upsilon(4S)$ events

decay channels and the new candidate selection criteria. The reported maximum tag-side efficiencies for the previously used exclusive tagging algorithms are summarized in "Results" section, Table 1. The stated efficiencies are not directly comparable due to different selection criteria, e.g., a threshold on the beam-constrained mass or the deviation of the nominal energy from the reconstructed energy $\Delta E = E_{\text{beam}} - E_B$ with E_B denoting the energy of the B candidate, best candidate selections, or cuts on the event shape used to suppress background from non-$\Upsilon(4S)$ events.

Results

The `FEI` algorithm was developed for the Belle II experiment. To quantify the improvements with respect to the previously used FR algorithm, the `FEI` is applied to data recorded by the Belle experiment. Simulated events and recorded data from the Belle experiment are converted into the new Belle II data format [4, Chapter 2]. This conversion tool was used to validate the entire Belle II analysis software and will be described in a separate publication [18]. The remainder of this article focuses on the results obtained for the hadronic tag on data recorded by the Belle experiment. The results for the semileptonic tag and for Belle II are based on simulated events and are only summarized briefly. A detailed validation of the entire algorithm can be found in Keck [4, Chapter 4].

Hadronic Tag

The performance of the hadronic tag provided by the `FEI` using simulated and recorded Belle events is studied and compared to the previously used FR algorithm.

At first, the considered decay channels of the `FEI` are restricted to the set of hadronic decay channels used by the FR. The performance of the `FEI` to the FR are compared using the same hardware and the same simulated charged (neutral) $B\bar{B}$ Belle events. The `FEI` required 33% less computing time and achieved a maximum tag-side efficiency of 0.53% (0.33%) on simulated events, which is significantly higher than the previously reported tag-side efficiencies (see "Previous work" section). The increase in the maximum tag-side efficiency is due to the improved candidate selection criteria, in particular the best candidate selections.

Second, all decay channels of the `FEI` are used, including the 38 additional hadronic decay channels. The performance of the `FEI` to the FR using the same hardware and the same simulated charged (neutral) Belle events are then compared. The `FEI` required 48% more computing time and achieved a maximum tag-side efficiency of 0.76% (0.46%) on simulated events. The further increase in the maximum tag-side efficiency is due to the additional decay channels.

As mentioned before, the maximum tag-side efficiency is an important performance indicator for exclusive measurements, which can employ the completeness constraint to achieve a high final purity. The achieved maximum tag-side efficiencies are summarized in Table 1.

To validate the results for the hadronic tag obtained from the simulation study, we conducted exclusive measurements of ten different semileptonic B decay channels using the full $\Upsilon(4S)$ dataset recorded by Belle. The branching fractions of the considered semileptonic decay channels are well-known from independent untagged measurements. The branching fraction of those well-known decay channels is measured using the hadronic tag, taking into account all known disagreements between simulation and data, e.g., in the particle identification performance and the track reconstruction efficiency. We assume that the remaining disagreement between simulation and data is caused by the tag-side. Therefore, the ratio ϵ of the measured and the expected branching fraction is proportional to the ratio of the tag-side efficiency on recorded data and simulated events. Our assumption is supported by the compatibility of the extracted ratios within their uncertainties. Figure 3 summarizes the results for the ten decay channels. The ratios averaged over all control channels for the charged and neutral B_{tag} mesons are

$$\epsilon_{\text{charged}} = 0.74^{+0.014}_{-0.013} \pm 0.050$$
$$\epsilon_{\text{neutral}} = 0.86^{+0.045}_{-0.050} \pm 0.054,$$

where the first uncertainty is statistical and the second systematic. The systematic uncertainties arises from the signal-side, e.g., through uncertainties on the particle identification performance or the track reconstruction efficiency.

A detailed description of the control measurements, including results for each tag and control channel, can be found in Schwab [19]. A similar study was conducted in the past for the FR by Sibidanov et al. [20], yielding a similar overall ratio of $\epsilon_{\text{comb.}} = 0.75 \pm 0.03$. The rather large discrepancy between simulated events and recorded data is caused by the uncertainty on the branching fractions and decay models of the simulated B decay channels used for the tag-side and the large number of multivariate classifiers involved in the process.

The uncertainty on the tag-side efficiency of the `FEI` is one of the most important systematic uncertainties in the measurement of branching fractions of rare decays. The tag-side efficiency can be corrected using the extracted ratios. It is possible to apply this correction as a function of the tag-side decay channel and signal probability. A measurement which uses the ratios to correct the tag-side efficiency is performed relative to the considered calibration decay channels. The systematic uncertainty of the correction is given by the uncertainty of the ratios.
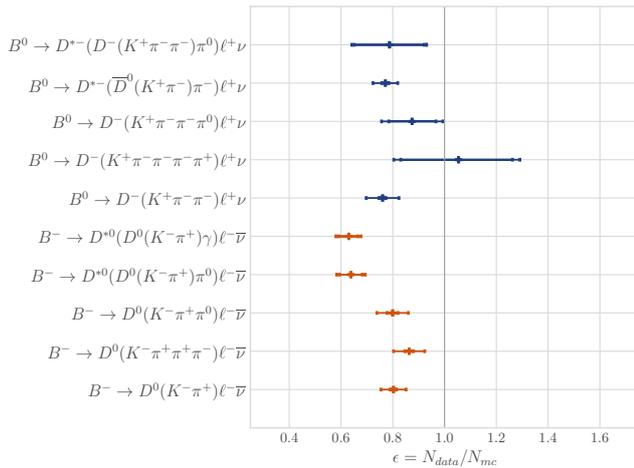
$B^0 \to D^{*-}(D^-(K^+\pi^-\pi^-)\pi^0)\ell^+\nu$

$B^0 \to D^{*-}(\overline{D}^0(K^+\pi^-)\pi^-)\ell^+\nu$

$B^0 \to D^-(K^+\pi^-\pi^-\pi^0)\ell^+\nu$

$B^0 \to D^-(K^+\pi^-\pi^-\pi^+)\ell^+\nu$

$B^0 \to D^-(K^+\pi^-\pi^-)\ell^+\nu$

$B^- \to D^{*0}(D^0(K^-\pi^+)\gamma)\ell^-\overline{\nu}$

$B^- \to D^{*0}(D^0(K^-\pi^+)\pi^0)\ell^-\overline{\nu}$

$B^- \to D^0(K^-\pi^+\pi^0)\ell^-\overline{\nu}$

$B^- \to D^0(K^-\pi^+\pi^+\pi^-)\ell^-\overline{\nu}$

$B^- \to D^0(K^-\pi^+)\ell^-\overline{\nu}$

$\epsilon = N_{data}/N_{mc}$

Fig. 3 The ratios calculated by measuring ten semileptonic decay channels on converted Belle data using the hadronic tag. The procedure is described in Schwab [19]

To compare the hadronic tag provided by the FEI and the FRin a well-defined manner, which is independent of the signal-side, both algorithms are applied to the same set of ten million events. These events are randomly sampled from the full $\Upsilon(4S)$ dataset of 772 million events recorded by the Belle experiment. After the tag-side reconstruction, only B meson candidates are kept, which fulfill cuts on the beam-constrained mass of $M_{\mathrm{bc}} > 5.24\,\mathrm{GeV}$ and on the deviation of the reconstructed energy from the nominal energy of $-0.15\,\mathrm{GeV} < \Delta E < 0.1\,\mathrm{GeV}$ calculated on the candidate. In addition, a best candidate selection is performed, taking the B meson candidate with the highest signal probability in each event.

The same cuts on the beam-constrained mass $M_{\mathrm{bc}} > 5.24\,\mathrm{GeV}$ and the deviation of the reconstructed energy from the nominal energy $-0.15\,\mathrm{GeV} < \Delta E < 0.1\,\mathrm{GeV}$ were applied and only the best (i.e., the highest signal probability) B meson candidate in each event was used.

From this dataset, we determined the tag-side efficiency and tag-side purity for different cuts on the signal probability. We followed the procedure established in previous publications [3, Chapter 7.1]. For different cuts on the signal probability, extended unbinned maximum likelihood fits of the beam-constrained mass spectrum are performed. The signal peak consisting of correct B_{tag} mesons is modeled with a Crystal Ball function [21], whereas the background is described using an ARGUS function [22]. The Gaussian mean of the Crystal Ball function was fixed to the B meson mass and its power law exponent was fixed to $m = 4$ based on the expected shape obtained from Monte Carlo simulations. The location and the width of the ARGUS were fixed using the known kinematic end point of the spectrum. All other parameters: the normalization of both functions, the width of the Crystal Ball, and the remaining

shape parameters of both functions were adjusted by the fit. The tag-side efficiency and tag-side purity are determined in a window of $5.27\,\mathrm{GeV} < M_{\mathrm{bc}} < 5.29\,\mathrm{GeV}$ using the fitted yields of the signal and background component.

In addition, we checked for a potential peaking combinatorial background component, which would bias the results. This test was done using ten million events recorded 60 MeV below the $\Upsilon(4S)$ resonance. This dataset does not contain B mesons, hence no signal is expected. The fitted signal yields were compatible with zero.

The resulting ROC curves are shown in Figs. 4 and 5 for charged and neutral B_{tag} mesons, respectively. The FEI exhibits a larger overall tag-side efficiency compared to the FR. We observe a slightly better performance for the FR than reported in Feindt et al. [14]. Both algorithms perform equally well when requiring a high tag-side purity. We suspect this is because there are only a finite number of cleanly identifiable B_{tag} meson candidates and both algorithms identify them with similar performance. The results for tag-side purities above 70% cannot be extracted reliably and depend strongly on the chosen signal or background fit model. For practical applications, the low tag-side purity regions is of particular interest for exclusive measurements. The beam-constrained mass distributions corresponding to the low-purity region with about 15% tag-side purity and the high-purity region with approximatively 80% tag-side purity are shown in Figs. 6 and 7, respectively, for the charged B_{tag}.

The maximum tag-side efficiency on recorded data is not determinable by this method, as the fits are restricted to the best B_{tag} candidates. However, a significant contribution to the improvement of the FEI compared to the FRis the increased number of provided candidates per event. A physics measurement will benefit from these additional tag-side candidates by first combining them with potential signal-side candidates, applying the completeness constraint (i.e.,
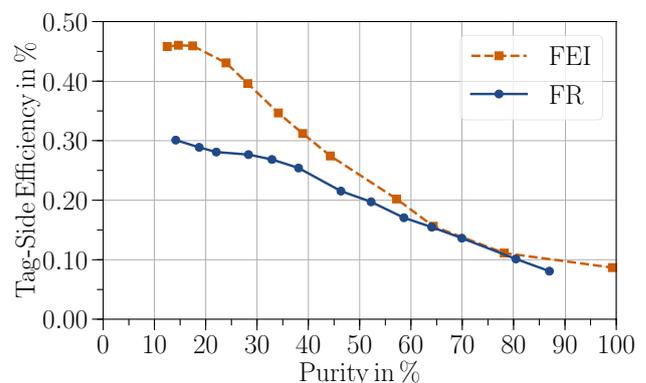

Fig. 4 Receiver operating characteristic of charged B_{tag} mesons extracted from a fit of the beam-constrained mass on converted Belle data. The FEI outperforms the FR algorithms performance at low and high purity
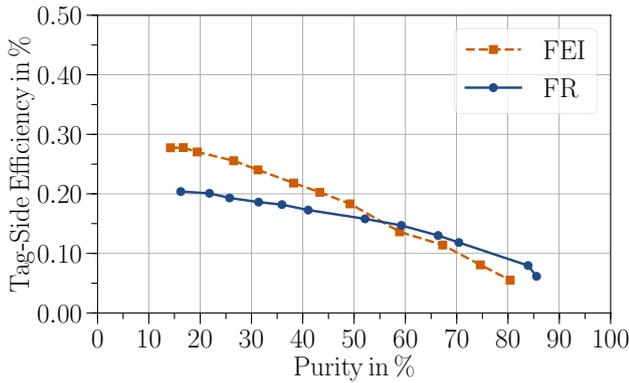
Fig. 5 Receiver operating characteristic of neutral B_{tag} mesons extracted from a fit of the beam-constrained mass on converted Belle data. The FEI outperforms the FR algorithms performance at low and intermediate purity. At high purity, the tag-side efficiency cannot be extracted reliably
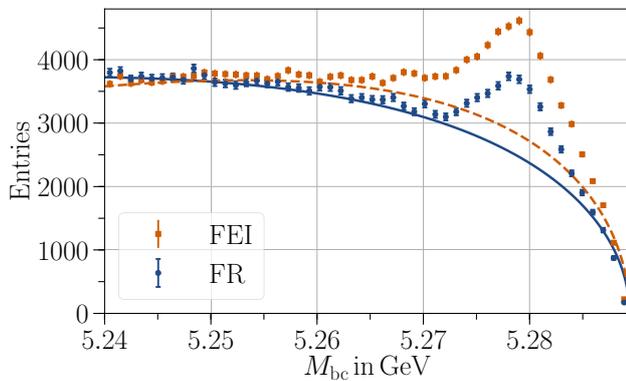

Fig. 6 Beam-constrained mass distribution of charged B_{tag} mesons in the low tag-side purity region on converted Belle data
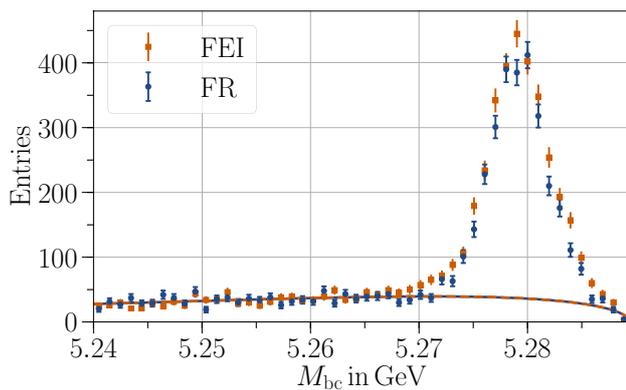

Fig. 7 Beam-constrained mass distribution of charged B_{tag} mesons in the high tag-side purity region on converted Belle data

requiring no additional tracks in the event), and performing the best B_{tag} candidate selection as the final step of the selection procedure. This procedure was successfully used by several measurements to validate the expected improvements on recorded data: [4, 19, 23].

Semileptonic Tag

The performance of the semileptonic tag provided by the FEI is studied using simulated Belle events. The maximum tag-side efficiencies are summarized in Table 1. Receiver operating characteristics extracted from simulated events can be found in Keck [4]. The results obtained from simulated events, and the fact that the hadronic and semileptonic tag only share five out of six reconstruction stages, indicate a significant increase in the maximum tag-side efficiency. The semileptonic tag was successfully used by Keck [4] to determine the branching fraction of $B \to \tau \nu_\tau$ on the full $\Upsilon(4S)$ dataset recorded by the Belle experiment, with a smaller relative statistical uncertainty than obtained previously. However, no studies with well-known calibration channels as described in Kronenbitter [24] and no signal-side independent determination of the ROCs as described in Kirchgessner [16], are available yet.

Outlook for Belle II

As the Belle II reconstruction software is still being optimized and no large recorded experimental data set was available at the time of writing, hence the final tag-side efficiency cannot be determined reliably for Belle II at this point. Preliminary results can be found in [4] which indicate a worse overall performance. This is likely due to the increased beam background caused by the higher luminosity of the collider, which does lead to additional tracks and neutral energy depositions. This additional detector activity is not yet fully rejected by the Belle II reconstruction algorithms [4] and future improvements are likely possible.

Discussion

The multivariate classifiers used by the FEI are trained on Monte Carlo simulated events. Depending on the training procedure and the type of events provided to the training, the multivariate classifiers of the FEI are optimized for different objectives.

In this article, we presented a so-called **generic** adaption of the FEI. The generic refers to that the FEI was trained independently of any specific signal-side using 180 million simulated $\Upsilon(4S)$ events. This setup optimizes the tag-side efficiency of a "generic" $\Upsilon(4S)$.

Other versions of the `FEI` exist which optimize the tag-side efficiency of specific signal events like $B \rightarrow \tau\nu_\tau$. The so-called **specific** `FEI` is trained on the remaining tracks and clusters after a potential signal B meson was already identified. The training uses simulated $\Upsilon(4S)$ events and simulated signal events. As a consequence, the classifiers can be specifically trained to identify correctly reconstructed B_{tag} mesons for signal events and can focus on reducing non-trivial background which is not discarded by the completeness constraint. The specific `FEI` was first introduced as a proof of concept by Keck [25] and used in Metzner [23].

Roughly half of the improvements with respect to the previous algorithm can be attributed to the additionally considered decay channels. Future extensions are currently investigated which use semileptonic D meson decays, baryonic decays and decays including K_L^0 particles.

It should also be noted that the `FEI` algorithm can be applied, with little modification, to the $\Upsilon(4S)$ resonance. This resonance decays into a pair of $B^{(*)}B^{(*)}$ and $B_s^{0(*)}B_s^{0(*)}$ mesons. The powerful completeness constraint can still be applied in this situation.

Conclusion

The `Full Event Interpretation` is a new exclusive tagging algorithm developed for the Belle II experiment that will be used to measure a wide range of decays with a minimum of detectable information. The algorithm exploits the unique setup of B factories and significantly improves the tag-side efficiency compared to its predecessor algorithms.

The tag-side efficiency for hadronically tagged B mesons was validated and calibrated using Belle data. Furthermore, the hadronic and the semileptonic tag provided by `FEI` have already been used in several validation measurements [4, 19, 26] using the full $\Upsilon(4S)$ dataset recorded by the Belle experiment. Similar studies and measurements for Belle II are anticipated as soon as the experiment records a sufficient amount of collision events.

There are several ways that the `FEI` algorithm could be further refined and applied to so far unexplored applications. These will provide an exciting and fruitful area of future research.

References

1. Abe T et al (2010) Belle II technical design report. Technical report, KEK-REPORT-2010-1. arXiv:1011.0352
2. Moll Andreas (2011) The software framework of the Belle II experiment. J Phys Conf Ser. https://doi.org/10.1088/1742-6596/331/3/032024
3. Bevan AJ (2014) The physics of the B factories. Eur Phys J C 74:3026. https://doi.org/10.1140/epjc/s10052-014-3026-9
4. Keck T (2017) Machine learning algorithms for the Belle II experiment and their validation on Belle data. PhD thesis, KIT. https://doi.org/10.5445/IR/1000078149
5. Dean J, Ghemawat S (2004) MapReduce: simplified data processing on large clusters. In: OSDI'04: sixth symposium on operating system design and implementation, San Francisco, CA, pp 137–150
6. Fruhwirth R (1987) Application of Kalman filtering to track and vertex fitting. Instrum Methods Nucl. https://doi.org/10.1016/0168-9002(87)90887-4
7. https://github.com/thomaskeck/FastFit. Accessed 02 Oct 2017
8. Keck T (2017) FastBDT: a speed-optimized multivariate classification algorithm for the Belle II experiment. Comput Softw Big Sci 1(1):9. https://doi.org/10.1007/s41781-017-0002-8
9. Hocker A et al (2007) TMVA—toolkit for multivariate data analysis. PoS ACAT:040
10. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay E (2011) Scikit-learn: machine learning in Python. J Mach Learn Res 12:2825–2830
11. Chen T, Guestrin C (2016) Xgboost: a scalable tree boosting system. CoRR. arXiv:abs/1603.02754
12. Aubert B et al (2013) The BABAR detector: upgrades, operation and performance. Nucl Instrum Methods A. https://doi.org/10.1016/j.nima.2013.05.107
13. Abashian A et al (2002) The Belle detector. Nucl Instrum Methods A479:117–232. https://doi.org/10.1016/S0168-9002(01)02013-7
14. Feindt M, Keller F, Kreps M, Kuhr T, Neubauer S, Zander D, Zupanc A (2011) A hierarchical NeuroBayes-based algorithm for full reconstruction of B mesons at B factories. Nucl Instrum Methods A654:432–440. https://doi.org/10.1016/j.nima.2011.06.008
15. Feindt M, Kerzel U (2006) The NeuroBayes neural network package. Nucl Instrum Methods 559(1):190–194. https://doi.org/10.1016/j.nima.2005.11.166
16. Kirchgessner K (2012) Semileptonic tag side reconstruction. Master's thesis, KIT. http://ekp-invenio.physik.uni-karlsruhe.de/record/48181
17. Kronenbitter B (2015) Measurement of the branching fraction of $B^+ \rightarrow \tau^+\nu_\tau$ decays with the semileptonic tagging method. Phys Rev D 92(5):051102. https://doi.org/10.1103/PhysRevD.92.051102

18. Gelb M et al (2018) B2BII: data conversion from Belle to Belle II. Comput Softw Big Sci 2(1):9. https://doi.org/10.1007/s41781-018-0016-x

19. Schwab J (2017) Calibration of the full event interpretation for the Belle and the Belle II experiment. Master's thesis, KIT. https://ekp-invenio.physik.uni-karlsruhe.de/record/48931

20. Sibidanov A (2013) Study of exclusive $B \to X_u \ell \nu$ decays and extraction of $|V_{ub}|$ using full reconstruction tagging at the Belle experiment. Phys Rev D 88(3):032005. https://doi.org/10.1103/PhysRevD.88.032005

21. Skwarnicki T (1986) PhD thesis, Institute for Nuclear Physics, Krakow

22. Albrecht H (1990) Search for Hadronic $b \to u$ decays. Phys Lett B 241:278–282. https://doi.org/10.1016/0370-2693(90)91293-K

23. Metzner F (2016) Analysis of $B^+ \to \ell^+ \nu_\ell \gamma$ decays with the Belle II analysis software framework. Master's thesis, KIT. https://ekp-invenio.physik.uni-karlsruhe.de/record/48845

24. Kronenbitter B (2014) Measurement of the branching fraction of $B^+ \to \tau^+ \nu_\tau$ decays at the Belle experiment. PhD thesis, KIT. https://ekp-invenio.physik.uni-karlsruhe.de/record/48604

25. Keck T (2014) The full event interpretation for Belle II. Master's thesis, KIT. https://ekp-invenio.physik.uni-karlsruhe.de/record/48602

26. Gelb M et al (2018) Search for the rare decay $B^+ \to \ell^+ \nu_\ell \gamma$ with improved hadronic tagging. Phys Rev D98(11):112016. https://doi.org/10.1103/PhysRevD.98.112016

C.2 Conference Proceedings

C.2.1 The 39th International Conference on High Energy Physics (ICHEP 2018)

Statement of Authorship

Title of Paper	Measurements of R(D(*)) and other missing energy decays at Belle II
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Conference proceedings for ICHEP 2018, published in Proceedings of Science

Principal Author

Name of Principal Author (Candidate)	Sophie Hollitt
Contribution to the Paper	Presented Belle II Collaboration prospects at the time of the ICHEP conference, for R(D*), B→τυ, and B→Kυυ studies. Authored full text of paper.
Overall percentage (%)	100
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.
Signature	Date 19/08/2019

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

    i.   the candidate's stated contribution to the publication is accurate (as detailed above);

    ii.   permission is granted for the candidate in include the publication in the thesis; and

    iii.   the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	
Contribution to the Paper	
Signature	Date

Name of Co-Author	
Contribution to the Paper	
Signature	Date

Please cut and paste additional co-author panels here as required.

Measurements of R(D(*)) and other missing energy decays at Belle II

Sophie Hollitt[*]

On behalf of the Belle II Collaboration
E-mail: sophie.hollitt@adelaide.edu.au

The Belle II experiment and SuperKEKB energy-asymmetric e^+e^- collider have already successfully completed Phase 1 and 2 of commissioning with first collisions seen in April 2018. The design luminosity of SuperKEKB is 8×10^{35} cm^{-2}s^{-1} and the Belle II experiment aims to record 50 ab^{-1} of data, a factor of 50 more than the Belle experiment. With this much data, decays sensitive to physics beyond the Standard Model can be studied with unprecedented precision. We present prospects for studying lepton flavor non-universality in $B \to D^{(*)}\tau\nu$ modes. Prospects for other missing energy modes sensitive to physics beyond the Standard Model such as $B^+ \to \tau^+\nu$ and $B \to K^{(*)}\nu\bar{\nu}$ are also discussed.

The 39th International Conference on High Energy Physics (ICHEP2018)
4-11 July, 2018
Seoul, Korea

---

[*]Speaker.

1. Introduction

In this era of high-luminosity experiments, we can continue to study smaller and more subtle discrepancies between experiment and Standard Model (SM) expectations. With a design instantaneous luminosity of 8×10^{35} cm^{-2}s^{-1} – 40 times the instantaneous luminosity of KEKB at Belle–SuperKEKB and the Belle II experiment have a key role to play in studying beyond SM physics and existing flavour anomalies. In these proceedings we focus on prospects for the $R(D^{(*)})$ anomaly and the rare decay modes $B \to \tau \nu$ and $B \to K^{(*)} \nu \overline{\nu}$. These modes are good candidates for studying possible new physics, and the first statistically significant measurements in the latter two modes are eagerly anticipated.

Since first collisions in April 2018, Phase 2 studies of the Belle II detector have been ongoing. About 500 pb^{-1} of data has been collected with a near-complete Belle II arrangement: all outer detectors and a partial central vertex detector[1]. Additional parts of the vertex detector are currently being installed ready for Belle II Phase 3 data collection, starting in early 2019. Data collection will continue until 2025, and at least 50 ab^{-1} of data is expected at the $\Upsilon(4S)$ resonance.

2. Full Event Reconstruction for Missing Energy

As Belle II is an almost hermitic detector surrounding an e^+e^- interaction, $B\overline{B}$ signal events can be well constrained by reconstructing all visible final state particles. For early rare decay analyses, we consider reconstructing both the signal B and a tag B meson decaying in a purely hadronic mode, which - when combined with the known centre-of-mass energy of the e^+e^- collision - allows us to also indirectly measure the missing energy and momentum from neutrino(s) in the signal decay. At Belle II, this is achieved using the Full Event Interpretation[2], an update to Belle's Full Reconstruction algorithm.

In brief, the Full Event Interpretation (FEI) is a collection of multivariate classifiers. From an initial classifier on tracks and vertex information, further classifiers are trained to reconstruct intermediate particles and the final B meson. The FEI allows reconstruction of 55 hadronic B decay channels - 24 more channels than Belle's equivalent Full Reconstruction - and also increases the number of D and D^* channels from 30 to 43 to further increase the number of B decay chains reconstructed. Overall, the FEI improves the tag reconstruction efficiency to 0.5% compared to Belle's 0.2% in Monte Carlo studies, despite increased background expected at Belle II.

Additional studies of FEI performance on Belle II's Phase 2 data are in progress.

3. Belle II Prospects

3.1 $B \to D^{(*)} \tau \nu$ and $R(D^{(*)})$

Current discrepancies between experimentally-measured $R(D^{(*)})$ and the expected SM values give tantalising hints of possible lepton flavour universality violation, especially when combined with additional 2-3σ anomalies in other channels such as the similar $R(J/\Psi)$ at LHCb. The current HFLAV world average experimental combination of $R(D)$ and $R(D^*)$ is at 3.78σ from the SM[3], including measurements from Belle, BaBar, and LHCb. Belle II will be able to make a simultaneous measurement of $R(D)$ and $R(D^*)$ at higher precision than before, shedding new light on this

1

flavour puzzle. A projection of the expected uncertainty of Belle II measurements of $R(D)$ and $R(D^*)$ at 50 ab^{-1} is shown in Figure 1, with additional predictions in Table 1. Belle II at 5 ab^{-1} is expected to have smaller uncertainty than the current world average.
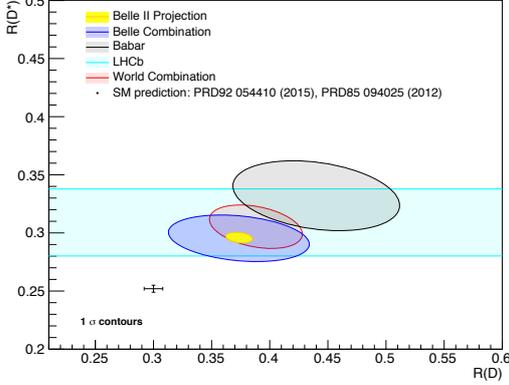

Figure 1: Projection of $R(D^*)$ and $R(D)$ uncertainties for the full Belle II dataset (yellow), compared to current experimental status and world averages. Figure also seen in [4].

	Observable	Uncertainty	
		stat.	syst.
1 ab^{-1}	$\mathscr{B}(B \to \tau\nu)$	29%	13%
5 ab^{-1}	$\mathscr{B}(B \to \tau\nu)$	13%	7%
	$R(D)$	6.0%	3.9%
	$R(D^*)$	3.0%	2.5%
	$P_\tau(D^*)$	0.18	0.08
50 ab^{-1}	$\mathscr{B}(B \to \tau\nu)$	4%	5%
	$R(D)$	2.0%	2.5%
	$R(D^*)$	1.0%	2.0%
	$P_\tau(D^*)$	0.06	0.04

Table 1: Prospects for uncertainties on $B \to D^{(*)}\tau\nu$ and $B \to \tau\nu$ observables at Belle II from Monte Carlo studies with hadronic FEI B tags. Data summarised from [4].

Theory uncertainties in $|V_{cb}|$ and the $B \to D^{(*)}l\nu$ form factors largely cancel in $R(D^{(*)})$. The remaining predicted systematic uncertainty in these measurements is dominated by uncertainty in decays involving excited D resonances (often labelled D^{**}), which contribute to $B \to D^{(*)}\tau\nu$ backgrounds via missing soft pions in reconstruction. Studies of both $B \to D^{**}\tau\nu$ and $B \to D^{**}l\nu$ decays are planned to help further reduce these uncertainties in the long term. Some additional excited D background may also be removed by taking advantage of improved vertexing at Belle II.

With the full Belle II dataset, measurement of additional observables in $B \to D^{(*)}\tau\nu$ will also be possible. Both the tau polarisation $P_\tau(D^*)$ and q^2 distribution will be available, and if the $R(D^{(*)})$ anomalies continue, this polarisation may be able to give insight into the type of new physics effects causing the differences from the Standard Model. In particular, at 50 ab^{-1} $P_\tau(D^*)$ may be able to discriminate between scalar, vector, or tensor New Physics[4].

3.2 $B^+ \to \tau^+\nu$

Although the $B^+ \to \tau^+\nu$ channel has been observed as a world average, no single experiment has reached 5σ discovery individually. The most recent Belle result has 24% uncertainty[5], and higher luminosity at Belle II will allow us to improve on this result.

The Belle II measurement of $B \to \tau\nu$ – one of many "golden mode" studies[4] – will be made using FEI hadronic B tag and a one-prong tau decay in the signal B. Correctly reconstructed events will have no additional charged tracks and will not have missed any neutral particles arising from B decays. At Belle and BaBar, the number of beam background photons is small, and so the energy in the calorimeter unassigned to any particles in the decay gives an indication of the number of neutral particles that may have been missed in reconstruction. This calorimeter energy distribution will thus peak at 0 for correctly reconstructed signal events. At 40 times the luminosity, additional

2

constraints are needed at Belle II to reject beam background photons from consideration in this extra neutral energy distribution. Such constraints include the use of calorimeter cluster timing information to identify and cut out photons that are not consistent with e^+e^- bunch crossing times.

Despite the complication of additional beam background, we still expect to see some improvement in reconstruction efficiency of $B \to \tau\nu$ events at Belle II compared to Belle, due to improved detector precision and the FEI tag reconstruction. Prospects for measurements of $B \to \tau\nu$ with different amounts of Belle II data are shown in Table 1. A SM $B \to \tau\nu$ 5σ discovery is predicted around 2.6 ab^{-1}.

3.3 $B \to K^{(*)}\nu\overline{\nu}$

In the longer term, $B \to K^{(*)}\nu\overline{\nu}$ is a rare decay channel with a lot of potential for SM and new physics studies. Measurements of these modes will be possible at Belle II, and the exact factorisation of hadrons and leptons in this decay allows precise experimental measurement of the $B \to K^{(*)}$ form factors. New Physics could also enhance the branching ratio of this mode, and the neutrino final state provides a window into right-handed operators. Interpreted as a measurement of $B \to K^{(*)} + invisible$, this decay channel also constrains weakly interacting dark matter scenarios[6].

Monte Carlo studies for this decay mode use the FEI Hadronic B tag, event shape constraints, and quality constraints on the K or K^* to determine an appropriate signal region. The current Standard Model estimate of the branching ratio for this decay has 10% uncertainty, which could be matched in experiment with the full 50 ab^{-1} dataset at Belle II. Evidence of a SM branching ratio is expected at 4 ab^{-1} and 5σ discovery is predicted with approximately 18 ab^{-1} of Belle II data.

4. Conclusion

The Belle II physics program has the opportunity to produce precision studies on a large number of new rare modes and known flavour anomalies. Results for $B \to \tau\nu$ and $R(D^{(*)})$ of interest to the global flavour community will be available at or by 5 ab^{-1}, currently anticipated in 2021.

References

[1] P. Branchini, *The Belle II Experiment: Status and Prospects*, *Universe* **4** (2018) 101.

[2] T. Keck, F. Abudinen, F. U. Bernlochner, R. Cheaib, S. Cunliffe, M. Feindt et al., *The Full Event Interpretation – An exclusive tagging algorithm for the Belle II experiment*, `1807.08680`.

[3] HEAVY FLAVOR AVERAGING GROUP collaboration, Y. Amhis et al., *Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2016*, *Eur. Phys. J.* **C77** (2017) 895 [`1612.07233`].

[4] The BELLE-II Collaboration, *The Belle II Physics book*, `1808.10567`.

[5] BELLE collaboration, B. Kronenbitter, M. Heck, P. Goldenzweig, T. Kuhr, A. Abdesselam, I. Adachi et al., *Measurement of the branching fraction of $B^+ \to \tau^+\nu_\tau$ decays with the semileptonic tagging method*, *Phys. Rev. D* **92** (2015) 051102.

[6] C. Bird, P. Jackson, R. Kowalewski and M. Pospelov, *Dark matter particle production in b→s transitions with missing energy*, *Physical Review Letters* **93** (2004) 1 [`hep-ph/0401195`].

3

C.2.2 The 36th Annual International Symposium on Lattice Field Theory (LATTICE 2018)

Statement of Authorship

Title of Paper	Control of SU(3) symmetry breaking effects in calculations of B meson decay constant
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Conference proceedings for LATTICE 2018, published in Proceedings of Science, also available on arxiv.org 1811.06677

Principal Author

Name of Principal Author (Candidate)	Sophie Hollitt
Contribution to the Paper	Tuning of b quarks for study. Calculation of decay constants on each ensemble. Prepared text of this paper. Work uses gauge field configurations generated by group
Overall percentage (%)	90
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.
Signature	Date 19/08/2019

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	A/Prof Paul Jackson
Contribution to the Paper	Supervisory role. The candidate performed the entirety of the work published. I grant permission for this publications inclusion in the thesis
Signature	Date 19/08/2019

Name of Co-Author	
Contribution to the Paper	
Signature	Date

Please cut and paste additional co-author panels here as required.

180

Control of SU(3) symmetry breaking effects in calculations of B meson decay constant

Sophie Hollitt[*],[a] **R. Horsley,**[b] **P. D. Jackson,**[a] **Y. Nakamura,**[c] **H. Perlt,**[d] **P. E. L. Rakow,**[e] **G. Schierholz,**[f] **A. Schiller,**[d] **H. Stüben,**[g] **R. D. Young**[a] **and J. M. Zanotti**[a]

[a]*CSSM/CoEPP, Department of Physics, University of Adelaide, Adelaide SA 5005, Australia*

[b]*School of Physics and Astronomy, University of Edinburgh, Edinburgh EH9 3FD, United Kingdom*

[c]*RIKEN Advanced Institute from Computation Science, Kobe, Hyogo 650-0047, Japan*

[d]*Institut für Theoretische Physik, Universität Leipzig, 04103 Leipzig, Germany*

[e]*Theoretical Physics Division, Department of Mathematical Physics, University of Liverpool, Liverpool L69 3BX, United Kingdom*

[f]*Deutsches Elektronen-Synchrotron DESY, 23603 Hamburg, Germany*

[g]*RRZ, Univeristy of Hamburg, 20146 Hamburg, Germany*

E-mail: sophie.hollitt@adelaide.edu.au

Early *B*-physics experiments have left us with a number of puzzles in heavy flavour physics. New lattice calculations (with a greater understanding of QCD effects in the Standard Model) will be needed to support the increase in experimental precision to be achieved by upcoming experiments such as Belle II. We extend the CSSM/UKQCD/QCDSF studies of SU(3) flavour breaking effects by presenting new results for the decay constants f_B and f_{B_s}.

The 36th Annual International Symposium on Lattice Field Theory - LATTICE2018
22-28 July, 2018
Michigan State University, East Lansing, Michigan, USA.

---

[*]Speaker.

1. Introduction

Recent *B* physics experiments have left us with a number of heavy flavour physics puzzles [1], and as Belle II's physics run approaches, a more precise understanding of the QCD contribution to the Standard Model determination of heavy flavour observables is needed to help isolate possible new physics by reducing theory errors. *B* meson decay constants in particular are used in the determination of multiple CKM matrix elements: $|V_{td}|$ and $|V_{ts}|$ from $B^0\overline{B^0}$ and $B_s^0\overline{B_s^0}$ oscillations, and $|V_{ub}|$ and $|V_{cb}|$ from leptonic decays of *B* and B_c mesons respectively. Combined with precise lattice calculations of f_B, measurements of the $B \to \tau\nu$ branching ratio at Belle II can be used as an independent measurement of $|V_{ub}|$ and give further insight into the existing discrepancy between $|V_{ub}|$ measured from inclusive and exclusive decays.

As f_B is often calculated on the lattice via the ratio f_{B_s}/f_B [2], it is important to understand and control SU(3) breaking effects in the light and strange quarks, and study how these affect extrapolations of f_{B_s}/f_B. In this work, we compute f_B and f_{B_s} using a set of gauge field configurations that break SU(3) flavour in a controlled way, keeping the average of the lighter quark masses held fixed at the physical value.

2. Simulation Details

2.1 SU(3) breaking and quark actions

We use multiple ensembles of gauge field configurations with 2+1 flavours of non-perturbatively $\mathcal{O}(a)$ improved Wilson fermions. When extrapolating to the physical point using multiple lattice ensembles with different quark masses, it is common practice to choose the strange quark mass m_s to be held (approximately) fixed at its physical value. We instead follow the QCDSF process for choosing the masses of light and strange quarks in a 2+1 flavour formalism [3], where the value of $\overline{m} = \frac{1}{3}(2m_l + m_s)$ is kept constant to control symmetry breaking. In this approach, all flavour-singlet quantities are only affected by SU(3)-flavour breaking effects at $\mathcal{O}((\delta m)^2)$, and have been shown to stay approximately constant from the SU(3) symmetric point to the physical point [3]. A diagram showing this behaviour compared to a standard approach is given in the left plot of Figure 1. The *x*-axis is scaled with the flavour-singlet combination of light pseudoscalar meson masses, $X_\pi^2 = \frac{1}{3}(2m_K^2 + m_\pi^2)$.

In the specific case of *B*-mesons, we also expect flavour-singlet combinations of *B* meson properties to be approximately constant along this quark mass trajectory. We can thus use properties of the physical *B* flavour singlet as an appropriate target in tuning our *B*-mesons on the lattice. We label this *B* flavour singlet meson as $X_B = 1/3\,(2B_l + B_s)$ and then consider its mass (M_{X_B}) or decay constant (f_{X_B}) with an appropriate substitution.

We generate bottom quarks using a variant of the 'Fermilab action' or 'RHQ action' [4, 5]. This anisotropic clover-improved action has the form [6]

$$S_{lat} = a^4 \sum_{x,x'} \overline{\psi}(x') \left(m_0 + \gamma_0 D_0 + \zeta \vec{\gamma} \cdot \vec{D} - \frac{a}{2}(D^0)^2 - \frac{a}{2}\zeta(\vec{D})^2 + \sum_{\mu,\nu} \frac{ia}{4} c_P \sigma_{\mu\nu} F_{\mu\nu} \right)_{x,x'} \psi(x)$$
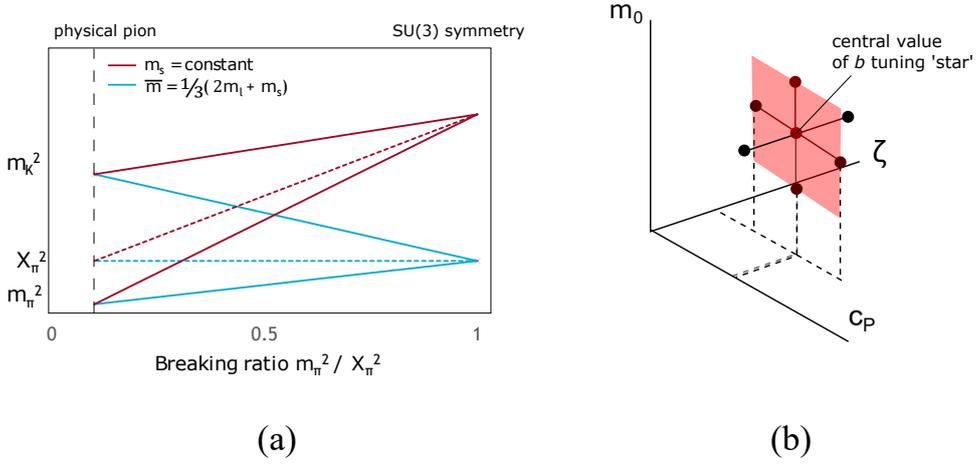
1

Figure 1: (a) Diagram of the evolution of m_π^2 and m_K^2 for constant strange quark mass m_s (red) or constant average quark mass (blue). The dashed lines show the evolution of the flavour singlet $X_\pi^2 = \frac{1}{3}(2m_K^2 + m_\pi^2)$ (b) Diagram of the tuning 'star' shape, which has one central value of (m_0, c_P, ζ) and then additional values at $\pm\Delta m$, $\pm\Delta c_P$, $\pm\Delta\zeta$

.

where m_0, c_P, and ζ are tuned as three free parameters. The 'best' B meson is selected by tuning the free parameters until the masses and hyperfine splitting of our calculated X_B and X_{B^*} mesons match the properties of the physical X_B and X_{B*}.

In practice, uncertainties on measured masses and splittings also result in uncertainty in the values of m_0, c_P, and ζ corresponding to the 'best' tuned B meson. We choose to generate multiple b-quarks per lattice ensemble in a 'tuning star' shape (see right-hand plot in Figure 1) and interpolate to the 'best' B, rather than generating only one 'best' b-quark per ensemble. Since we tune using SU(3) flavour singlets, we can employ the same set of seven b-quarks for each ensemble with the same lattice spacing and volume along the line of constant $\overline{m}$.

2.2 Plateau fitting for multiple b quarks

B-meson properties are calculated using fits to correlators and ratios of correlators. To maintain consistency, we choose to use the same fit window for each correlator across all 7 b-quark candidates on a fixed lattice ensemble. Selecting the best fit is assisted by using the correlated χ^2/d.o.f for the fit on each B meson correlator.

2.3 Lattice spacings and volumes

A variety of lattice spacings and lattice volumes are used in this work. Some details of the QCDSF gauge field ensembles are presented in Table 1. The interpolated 'best' values for m_0, c_P, and ζ on each ensemble are not shown here due to space constraints, but will be presented in future studies on the systematic uncertainties. For all results, the source locations for the calculated mesons are randomised to reduce correlations between neighbouring configurations in the ensemble.

2

β	a (fm)	Lattice volume	κ_{light}	κ_{strange}	m_π (MeV)	m_K (MeV)	
5.4	0.082	$24^3 \times 48$	0.11993	0.11993	413	413	
			0.120048	0.119695	325	448	
		$32^3 \times 64$	0.11993	0.11993	408	408	
			0.119989	0.119812	366	424	†
			0.120084	0.119623	290	450	†
5.5	0.074	$32^3 \times 64$	0.1209	0.1209	468	468	*
			0.12104	0.12062	357	505	*
			0.121095	0.120512	315	526	*
			0.121145	0.120413	258	537	*
		$32^3 \times 64$	0.12095	0.12095	403	403	
			0.12104	0.12077	331	435	
			0.121099	0.120653	270	454	
		$48^3 \times 96$	0.121166	0.120371	226	539	*
5.65	0.068	$32^3 \times 64$	0.122005	0.122005	421	421	
			0.122078	0.121859	361	448	
			0.12213	0.121756	310	463	
		$48^3 \times 96$	0.122005	0.122005	412	412	
			0.122078	0.121859	355	441	
			0.12213	0.121756	302	457	
			0.122167	0.121682	265	474	
		$64^3 \times 96$	0.122227	0.121563	155	480	†
5.8	0.059	$48^3 \times 96$	0.12281	0.12281	427	427	
			0.12288	0.12267	357	456	
			0.12294	0.122551	280	477	

Table 1: Table of lattice ensembles used in this work. * indicates ensembles with a different value of $\overline{m}$, further from the physical $\overline{m}$. † indicates ensembles where analysis is still in progress.

3. Calculating f_B on the lattice

The decay constant f_B is calculated from its lattice counterpart Φ_B via the equation

$$f_B = \frac{1}{a} Z_\Phi \left[\Phi_B^0 + c_A \Phi_B^1 \right]$$

where Φ_B is calculated from two-point correlators for axial and pseudoscalar operators:

$$\Phi_B = -\frac{\sqrt{2M_B}\mathscr{C}_{AP}}{\mathscr{C}_{PP}}, \qquad \mathscr{C}_{AP} = \frac{\langle \Omega | A_4 | B \rangle \, \langle B | P | \Omega \rangle}{2M_B}, \qquad \mathscr{C}_{PP} = \frac{\langle \Omega | P | B \rangle \, \langle B | P | \Omega \rangle}{2M_B}$$

and Z_Φ is calculated:

$$Z_\Phi = \rho_A^{bl} \sqrt{Z_V^{bb} Z_V^{ll}}.$$

The perturbative constant ρ_A^{bl} is set to 1 in this work, and similarly the higher-order correction coefficient c_A in f_B is set to 0. For determining $Z_V^{bb/ll}$, we compute meson three point functions of

3

the vector current and enforce charge conservation. This formulation of Φ_B is equivalent to that used in [7].

In these proceedings, we focus on the SU(3) flavour breaking effects in the ratio f_B/f_{X_B} so that most sources of systematic error cancel. The SU(3) breaking in these ratios of f_B or f_{B_s} can be modelled by adapting the partially-quenched NLO equations in Bornyakov et al [8]:

$$
\begin{aligned}
\frac{f_B(q\bar{b})}{f_{X_B}} = {} & 1 + G(\delta\mu_q) + (H_1 + H_2)\delta\mu_q^2 \\
& - (\tfrac{2}{3}H_1 + H_2)(\delta m_u^2 + \delta m_d^2 + \delta m_s^2) \\
& + \dots
\end{aligned}
\tag{3.1}
$$

where $\delta\mu_q$ represent the distance between the valence quark mass and the SU(3) symmetric mass, and similarly δm are differences in the sea quark masses. By fitting the coefficients in this equation to f_B and f_{B_s} on each set of ensembles, we can use the $\delta\mu$ and δm values corresponding to the physical point to extrapolate to a physical prediction of f_B/f_{X_B}.

4. Results

The decay constant ratios f_B/f_{X_B} and f_{B_s}/f_{X_B} are shown against the SU(3) breaking ratio m_π^2/X_π^2 in Figure 2. It should be noted that the statistical error is mostly a result of error propagation resulting from interpolating the fit to the 'best' B meson on each ensemble. Both linear and quadratic fits in m_π^2 are shown: it can be seen that while the data is mostly linear in the breaking ratio m_π^2/X_π^2, an additional quadratic term is required to describe the data well and approach the FLAG values [2] denoted by $\star$.

The same ratios are presented again in Figure 3, in terms of the light or strange quark mass $\delta\mu_q$ in Equation 3.1. We also include some additional B mesons with partially quenched light or strange quarks from the $32^3 \times 64$ β=5.4 $\kappa_l = \kappa_s = 0.11993$ ensemble. Each set of ensembles with a different lattice spacing will require a separate fit as part of our future extrapolation toward the continuum limit. We note that the trend is mostly linear in $\delta\mu_q$, indicating that these terms in Equation 3.1 will dominate in agreement with observations made in the light quark sector [8], although some hints of curvature (and thus higher order terms) are still visible for larger $\delta\mu_q$.

5. Conclusion

We have presented the current status of our investigation of SU(3) symmetry breaking effects in B meson decay constants. A large number of lattice ensembles have already been processed, though a few additional ensembles are required to complete our analysis of lattice discretisation effects and make extrapolations toward the physical and continuum limits in the near future. Studies of the B^* decay constant are also in progress.

4
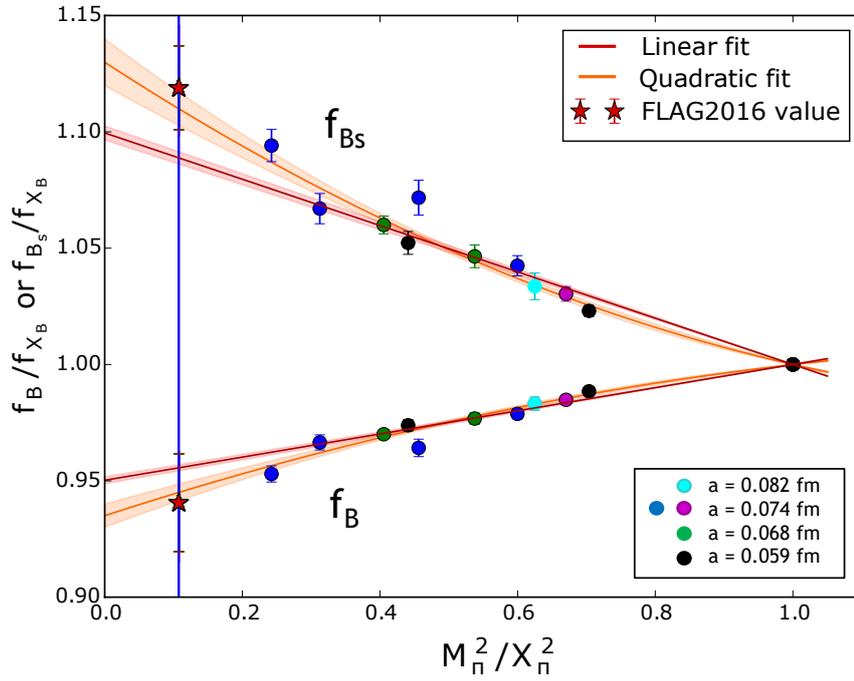
Figure 2: f_B/f_{X_B} and f_{B_s}/f_{X_B} for a variety of lattice ensembles.
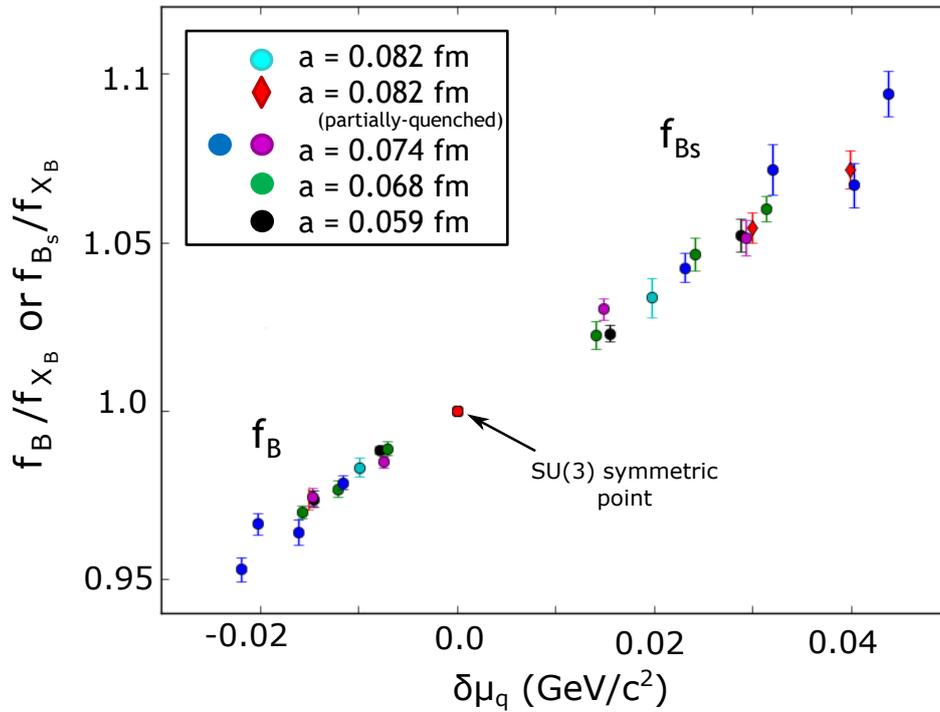
Figure 3: f_{B_q}/f_{X_B} against $\delta\mu_q$ in physical units.

5

Acknowledgments

The numerical configuration generation (using the BQCD lattice QCD program [9]) and data analysis (using the Chroma software library [10]) was carried out on the IBM BlueGene/Q and HP Tesseract using DIRAC 2 resources (EPCC, Edinburgh, UK), the IBM BlueGene/Q at NIC (Jülich, Germany), the Cray XC40 at HLRN (The North-German Supercomputer Alliance), the NCI National Facility in Canberra, Australia, and the iVEC facilities at the Pawsey Supercomputing Centre. These Australian resources are provided through the National Computational Merit Allocation Scheme and the University of Adelaide Partner Share supported by the Australian Government. This work was supported in part through supercomputing resources provided by the Phoenix HPC service at the University of Adelaide. The BlueGene codes were optimised using Bagel [11]. The Chroma software library [10], was used in the data analysis. This investigation has been supported by the Australian Research Council under grants FT120100821, FT100100005, FT130100018 and DP140103067 (RDY and JMZ).

References

[1] J. Albrecht, F. U. Bernlochner, M. Kenzie, S. Reichert, D. M. Straub and A. Tully, *Future prospects for exploring present day anomalies in flavour physics measurements with Belle II and LHCb*, arXiv:1709.10308v5.

[2] S. Aoki, Y.Aoki, D.Bečirević, C.Bernard, T.Blum, G. Colangelo et al., *Review of lattice results concerning low-energy particle physics*, *European Physical Journal C* **77** (2017) 1 [1607.00299].

[3] W. Bietenholz, V. Bornyakov, M. Göckeler, R. Horsley, W. G. Lockhart, Y. Nakamura et al., *Flavor blindness and patterns of flavor symmetry breaking in lattice simulations of up, down, and strange quarks*, *Physical Review D - Particles, Fields, Gravitation and Cosmology* **84** (2011) 1 [1102.5300].

[4] A. X. El-Khadra, A. S. Kronfeld and P. B. Mackenzie, *Massive fermions in lattice gauge theory*, *Physical Review D - Particles, Fields, Gravitation and Cosmology* **55** (1997) 3933 [9604004].

[5] H. W. Lin and N. Christ, *Nonperturbatively determined relativistic heavy quark action*, *Physical Review D - Particles, Fields, Gravitation and Cosmology* **76** (2007) 1 [0608005].

[6] Y. Aoki, N. Christ, J. Flynn, T. Izubuchi, C. Lehner, M. Li et al., *Nonperturbative tuning of an improved relativistic heavy-quark action with application to bottom spectroscopy*, *Physical Review D* **86** (2012) 116003 [1206.2554].

[7] O. Witzel, *B -meson decay constants with domain-wall light quarks and nonperturbatively tuned relativistic b-quarks*, in *LATTICE 2013*, pp. 0–6, 2013, arXiv:1311.0276v2.

[8] V. G. Bornyakov, R. Horsley, Y. Nakamura, H. Perlt, D. Pleiter, P. E. L. Rakow et al., *Flavour breaking effects in the pseudoscalar meson decay constants*, *Physics Letters B* **767** (2017) 366 [1612.04798].

[9] T. R. Haar, Y. Nakamura and H. Stüben, *An update on the BQCD Hybrid Monte Carlo program*, in *EPJ Web of Conferences*, vol. 175, p. 14011, EDP Sciences, 2018, 1711.03836.

[10] R. G. Edwards and B. Joó, *The Chroma software system for lattice QCD*, *Nuclear Physics B - Proceedings Supplements* **140** (2005) 832 [0409003].

[11] P. A. Boyle, *The BAGEL assembler generation library*, *Comput. Phys. Commun.* **180** (2009) .

6

C.2.3 The 26th International Nuclear Physics Conference (INPC 2016)

Statement of Authorship

Title of Paper	The role of SU(3)-flavour symmetry breaking in B-meson decay constants from Lattice QCD
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Conference proceedings for INPC 2016, published in Proceedings of Science.

Principal Author

Name of Principal Author (Candidate)	Sophie Hollitt
Contribution to the Paper	Tuning of B quarks for study. Calculation of decay constants for each ensemble. Prepared text of this paper. Work over gauge field configurations generated by group.
Overall percentage (%)	90
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.
Signature	Date 19/08/2019

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

    i.   the candidate's stated contribution to the publication is accurate (as detailed above);

    ii.   permission is granted for the candidate in include the publication in the thesis; and

    iii.   the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	A/Prof Paul Jackson
Contribution to the Paper	Supervisory role. The candidate performed the entirety of the published work and permission is granted for the inclusion of this publication in the thesis.
Signature	Date 19/08/2019

Name of Co-Author	
Contribution to the Paper	
Signature	Date

Please cut and paste additional co-author panels here as required.

The role of SU(3)-flavour symmetry breaking in B-meson decay constants from Lattice QCD

Sophie Hollitt[∗]**, Paul Jackson, Ross Young, and James Zanotti**
CSSM and CoEPP, Department of Physics,
The University of Adelaide, Adelaide SA 5005, Australia
E-mail: sophie.hollitt@adelaide.edu.au

B-physics is currently facing a number of puzzles, and additional precision data - and precision calculations - will be required to pick these apart. In particular, as the Belle II detector approaches completion, there is increasing need for lattice calculations related to CKM matrix elements and to important *B* branching ratios such as $\mathscr{B}(B \to \tau \nu)$.

We present early results for f_B and f_{B_s} from the CSSM/QCDSF/UKQCD collaboration on a set of lattices with a fixed volume $32^3 \times 64$ and lattice spacing $a = 0.074$ fm. By varying the *u,d,s* quark masses while holding their average value constant, we are able to reliably control the SU(3)-flavour breaking effects.

[∗]Speaker.

PoS(INPC2016)272

1. Introduction

Results from Belle and BABAR were crucial for our understanding of *B*-physics and the unitarity in the CKM matrix, but these results have also left us with a number of puzzles where further understanding of QCD (and of our detectors) is required. As the Belle II experiment [1] approaches its first science run and LHCb continues to increase statistics, the pressure is on to improve errors on theoretical and lattice calculations ahead of future improvements to experimental precision. We choose to focus on the *B*-meson decay constants f_B and f_{B_s} which are crucial to Standard Model calculations of the branching ratio $\mathscr{B}(B \to \tau \nu)$, though the f_B and f_{B_s} also appear in calculations of CKM matrix elements $|V_{tb}|$ and $|V_{ts}|$ from measurements of CP violation in B^0 and $\bar{B}^0$ mesons.

The Flavour Lattice Averaging Group (FLAG) reviews [2, 3] indicate that there are several groups working on f_B and related *B* meson observables with an eye toward improved precision [4–9]. As a companion to this body of existing work, we investigate f_B and f_{B_s} with a focus on SU(3) symmetry breaking effects by choosing light and strange quarks with a constant average mass $\overline{m}$. This methodology comes from the UKQCD/QCDSF group and we follow a similar method in our f_B/f_{B_s} study as in their related study of f_π and f_K. [10]

2. Simulation Details

2.1 Treatment of light and strange quarks

When extrapolating to the physical point using multiple lattice ensembles with different quark masses, many groups will choose the strange quark mass m_s to be constant. We instead follow the UKQCD/QCDSF process for choosing the masses of light and strange quarks in a 2+1 formalism [10]. The value of $\overline{m} = \frac{1}{3}(2m_l + m_s)$ is kept constant to control symmetry breaking and remove effects of $\mathscr{O}(\delta m)$. In fact, all flavour singlet quantities are only affected by SU(3)-flavour breaking effects at $\mathscr{O}((\delta m)^2)$, and have been shown to stay approximately constant from the SU(3) symmetric point to the physical point. [10]

For the ensembles of lattice configurations used in this work, there is a mixture of ensembles where $\overline{m}$ is equal to the physical value of $\overline{m}$ and ensembles where $\overline{m}$ has a slightly different value. The relationship between m_l and m_s (or equivalently, the relationship between the pion and kaon masses) for different ensembles at $a = 0.0074$ fm [11] is shown graphically in Figure 1 (inset) and also displayed in Table 1.

2.2 Bottom quarks

We generate bottom quarks using the anisotropic clover-improved action [7]

$$S_{lat} = a^4 \sum_{x,x'} \overline{\psi}(x') \left(m_0 + \gamma_0 D_0 + \zeta \vec{\gamma} \cdot \vec{D} - \frac{a}{2}(D^0)^2 - \frac{a}{2}\zeta(\vec{D})^2 + \sum_{\mu,\nu} \frac{ia}{4} c_P \sigma_{\mu\nu} F_{\mu\nu} \right)_{x,x'} \psi(x)$$

and tune m_0, c_P and ζ to specify the mass, hyperfine splitting, and dispersion relation of the generated $B(*)$ or $B(*)_s$ mesons. This is a variant of the 'Fermilab action' or 'RHQ action' [12, 13]. We choose the 'best' tuning by considering a flavour singlet *B*-meson $X_B = \frac{1}{3}(2B_l + B_s)$ and selecting

1

the tuning parameters such that our calculated B and B_s mesons combine to create an X_B matching X_B for the physical B and B_s.

In practice, uncertainties on the measured mass, splitting, and dispersion relation also result in uncertainties in the values of m_0, c_P and ζ corresponding to the 'best' tuned B meson. We choose to always generate multiple b-quarks per lattice ensemble and interpolate to the 'best' B, rather than generating only one 'best' b-quark after completing the tuning process. This allows us to use the same set of seven b-quarks for each ensemble with the same lattice spacing and volume.

2.3 Additional Information

In this work, we use multiple ensembles of ~ 800 gauge field configurations with 2+1 flavours of non-perturbatively $\mathscr{O}(a)$ improved Wilson fermions. Further details of the configurations used in each ensemble are shown in Table 1 and Table 2.

(κ_l, κ_s)	m_π (MeV)	m_K (MeV)	# configs used	m_0	c_P	ζ
(0.12090,0.12090)	465	465	778	2.80 ± 0.13	3.60 ± 0.34	1.30 ± 0.11
(0.12104,0.12062)	360	505	758	2.65 ± 0.11	3.19 ± 0.29	1.37 ± 0.11
(0.121095,0.120512)	310	520	380	2.98 ± 0.22	4.03 ± 0.57	1.21 ± 0.16
(0.12095,0.12095)	400	400	400	2.69 ± 0.15	3.29 ± 0.39	1.50 ± 0.14
(0.12104,0.12077)	330	435	786	2.82 ± 0.13	3.59 ± 0.34	1.38 ± 0.10

Table 1: Lattice configurations and tuning results for ensembles with $V = 32^3 \times 64$, $a = 0.074$ fm, and $\beta = 5.5$ used in SU(3) breaking calculations

V	a (fm)	β	$\kappa_l = \kappa_s$	# configs used	m_0	c_P	ζ
$24^3 \times 48$	0.0818	5.4	0.11993	808	3.90 ± 0.32	4.64 ± 0.91	1.33 ± 0.24
$32^3 \times 64$	0.074	5.5	0.120900	778	2.80 ± 0.13	3.60 ± 0.34	1.30 ± 0.11
$32^3 \times 64$	0.0684	5.65	0.122005	410	2.74 ± 0.18	4.34 ± 0.48	1.13 ± 0.12

Table 2: Lattice configurations and tuning results for additional ensembles at the SU(3) symmetric point

The source locations for quarks are randomised to reduce correlations between neighbouring configurations in the ensemble. Future work will include either more configurations for each lattice ensemble where additional configurations are available, or additional source locations on the same configurations to increase statistics.

3. Calculating f$_B$ on the lattice

The decay constant f_B is calculated from its lattice counterpart Φ_B via the equation

$$f_B = \frac{\hbar c}{a} Z_\Phi \Phi_B$$

2

where Φ_B is calculated from two-point correlators for Axial and Pseudoscalar operators:

$$\Phi_B = -\frac{\sqrt{2M_B}\mathscr{C}_{AP}}{\mathscr{C}_{PP}}, \qquad \mathscr{C}_{AP} = \frac{\langle\Omega|A_4|B\rangle\langle B|P|\Omega\rangle}{2M_B}, \qquad \mathscr{C}_{PP} = \frac{\langle\Omega|P|B\rangle\langle B|P|\Omega\rangle}{2M_B}$$

and Z_Φ is calculated:

$$Z_\Phi = \rho_A^{bl}\sqrt{Z_V^{bb}Z_V^{ll}}.$$

This formulation of Φ_B is equivalent to that used in [6], which does not include the factor $2m$ in the correlators.

In practice, Z_V^{bb} has been calculated using a light spectator quark in the three-point correlator, and we use $\rho_{\text{lat}}^{bl} = 1$ for these early calculations of f_B. f_B can also be improved by letting Φ_B go to $\Phi_B^0 + c_A\Phi_B^1$ using an improvement coefficient c_A, but this has not yet been explored at the time of this work.

4. Results

4.1 SU(3) symmetry breaking in f_B and f_{B_s}

Following the procedure used in the light quark sector [14], we plot f_B and f_{B_s} relative to the average decay constant $f_{B_X} = \frac{1}{3}(2*f_B + f_{B_s})$ in order to cancel out some errors and look for SU(3)-flavour breaking behaviour. If SU(3) breaking of the lighter quarks is the main effect, we expect a linear fit for f_B/f_{B_X} against m_π^2/m_X^2 as was observed for f_π and f_K in [14]. These early results with a linear fit show good agreement with the FLAG2013 average of $N = 2 + 1$ flavour calculations. [2]
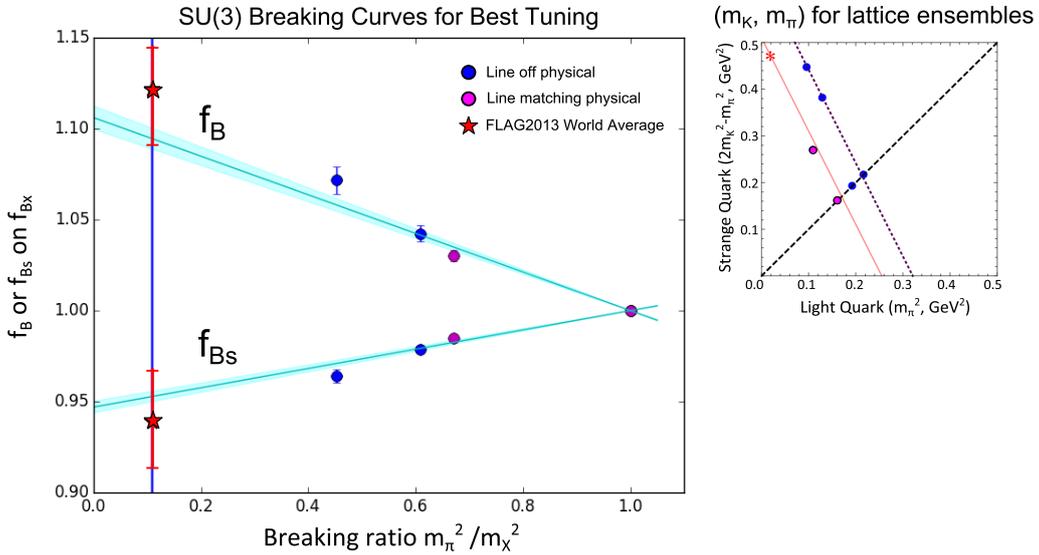

Figure 1: (Main Graph): Calculated values for f_B and f_{B_s} are plotted relative to the average decay constant $f_{B_X} = \frac{1}{3}(2*f_B + f_{B_s})$. (Inset): Legend of location of lattice ensembles on a plane representing the strange and light quark masses relative to the physical point, plot taken from [15]

3

4.2 Toward predictions for f_B and f_{B_s}

It is crucial to control as many sources of uncertainty as possible in order to have competitive calculated values for f_B and f_{B_s}. As part of an investigation of lattice artefacts, we examine the way that the normalisation factor Z_Φ changes as the lattice volume and lattice spacing are varied. The effect of this normalisation on f_B is shown in Figure 2, while the changes in the components $\sqrt{Z_V^{bb}}$ and $\sqrt{Z_V^{ll}}$ of Z_Φ are plotted against lattice spacing a^2 in Figure 3.
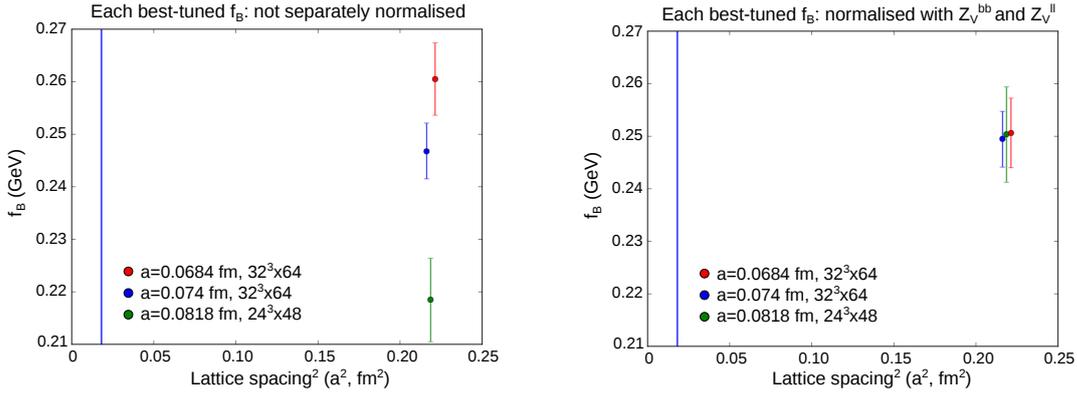
Figure 2: (Left) f_B, where the same Z_Φ value is used regardless of lattice spacing/lattice volume (Right) f_B, where Z_Φ is calculated for each lattice ensemble individually.
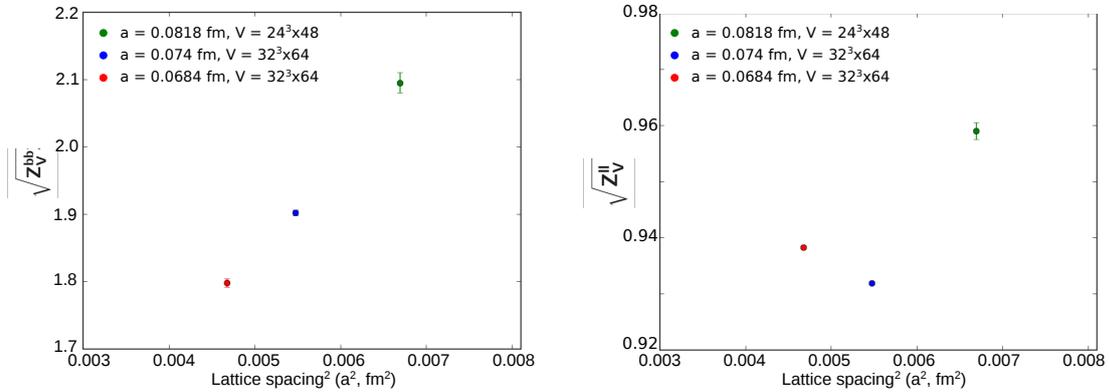
Figure 3: (Left) Normalisation factor $\sqrt{Z_V^{bb}}$ against lattice spacing squared. The relationship between the two is approximately linear. (Right) Normalisation factor $\sqrt{Z_V^{ll}}$ against lattice spacing squared. Z_V^{ll} appears to be affected most strongly by the lattice volume.

While $\sqrt{Z_V^{bb}}$ changes linearly with a^2 and appears to be unaffected by the lattice volume, the value of Z_V^{ll} is significantly different for the $24^3 \times 48$ and $32^3 \times 64$ lattices. We suspect that this difference is due to discretisation errors on the small $24^3 \times 48$ lattices, but this remains to be

4

checked in future work. Future work will also involve $48^3 \times 96$ lattices for additional comparison at $\beta = 5.65$ as well as at finer lattice spacings ($\beta = 5.8, 5.95$).

5. Conclusion

We have presented preliminary results for f_B and f_{B_s} with controlled SU(3) symmetry breaking by controlling the way light and strange quark masses are chosen. Investigation of systematic lattice discretisation effects is underway, and we look forward to further progress toward calculations of f_B and f_{B_s} with competitive errors.

Acknowledgments

References

[1] S. Uno, T. Abe, I. Adachi, K. Adamczyk and S. Ahn, *Belle II technical design report*, arXiv:1011.0352.

[2] S. Aoki, Y. Aoki, C. Bernard, T.Blum, G. Colangelo, M.DellaMorte et al., *Review of lattice results concerning low-energy particle physics*, *European Physical Journal C* **74** (2014) 1–179, [arXiv:1011.4408].

[3] S. Aoki, Y.Aoki, D.Bečirević, C.Bernard, T.Blum, G. Colangelo et al., *Review of lattice results concerning low-energy particle physics*, *European Physical Journal C* **77** (2017) 1–228, [arXiv:1011.4408].

[4] R. J. Dowdall, C. T. H. Davies, R. R. Horgan, C. J. Monahan and J. Shigemitsu, *B-meson decay constants from improved lattice nonrelativistic QCD with physical u, d, s, and c quarks*, *Physical Review Letters* **110** (2013) 1–5, [arXiv:1302.2644].

[5] N. H. Christ, J. M. Flynn, T. Izubuchi, T. Kawanai, C. Lehner, A. Soni et al., *B -meson decay constants from 2+1 -flavor lattice QCD with domain-wall light quarks and relativistic heavy quarks*, *Physical Review D - Particles, Fields, Gravitation and Cosmology* **91** (2015) 1–24.

[6] O. Witzel, *B -meson decay constants with domain-wall light quarks and nonperturbatively tuned relativistic b-quarks*, in *LATTICE 2013*, pp. 0–6, 2013. 1311.0276v2.

5

[7]   Y. Aoki, N. Christ, J. Flynn, T. Izubuchi, C. Lehner, M. Li et al., *Nonperturbative tuning of an improved relativistic heavy-quark action with application to bottom spectroscopy*, *Physical Review D* **86** (2012) 116003, [`arXiv:1206.2554`].

[8]   Y. Aoki, T. Ishikawa, T. Izubuchi, C. Lehner and A. Soni, *Neutral b meson mixings and b meson decay constants with static heavy and domain-wall light quarks*, *Physical Review D* **91** (2015) 114505.

[9]   H. Na, C. Monahan, C. Davies, R. Horgan, G. Lepage and J. Shigemitsu, *B and B_{s} meson decay constants from lattice QCD*, *Physical Review D* **86** (2012) 1–11, [`arXiv:1202.4914`].

[10] W. Bietenholz, V. Bornyakov, M. Göckeler, R. Horsley, W. G. Lockhart, Y. Nakamura et al., *Flavor blindness and patterns of flavor symmetry breaking in lattice simulations of up, down, and strange quarks*, *Physical Review D - Particles, Fields, Gravitation and Cosmology* **84** (2011) 1–43, [`arXiv:1102.5300`].

[11] V. G. Bornyakov, R. Horsley, R. Hudspith, Y. Nakamura, H. Perlt, D. Pleiter et al., *Wilson flow and scale setting from lattice QCD*, `arXiv:1508.05916`.

[12] Z. Sroczynski, *First results from the asymmetric O(a) improved Fermilab action*, *Nuclear Physics B (Proceedings Supplement)* **84** (1999) 971–973, [`9910004`].

[13] H. W. Lin and N. Christ, *Nonperturbatively determined relativistic heavy quark action*, *Physical Review D - Particles, Fields, Gravitation and Cosmology* **76** (2007) 1–20, [`0608005`].

[14] V. G. Bornyakov, R. Horsley, Y. Nakamura, H. Perlt, D. Pleiter, P. E. L. Rakow et al., *Flavour breaking effects in the pseudoscalar meson decay constants*, *Physics Letters B* **767** (2017) 366–373, [`arXiv:1612.04798`].

[15] P. E. Shanahan, R. Horsley, Y. Nakamura, D. Pleiter, P. E. L. Rakow, G. Schierholz et al., *Electric form factors of the octet baryons from lattice QCD and chiral extrapolation*, *Physical Review D - Particles, Fields, Gravitation and Cosmology* **90** (2014) 1–17, [`arXiv:1401.5862`].

[16] Y. Nakamura and H. Stüben, *BQCD - Berlin quantum chromodynamics program*, *Proceedings of Science Lattice 2010* (2010) 7, [`arXiv:1011.0199`].

[17] P. A. Boyle, *The BAGEL assembler generation library*, *Comput. Phys. Commun.* **180** (2009) .

[18] R. G. Edwards and B. Joó, *The Chroma software system for lattice QCD*, *Nuclear Physics B - Proceedings Supplements* **140** (2005) 832–834, [`0409003`].

PoS(INPC2016)272

6

Bibliography

[1] M. Kobayashi and T. Maskawa, *CP-Violation in the Renormalizable Theory of Weak Interaction*, Progress of Theoretical Physics **49** (1973) 652.

[2] S. W. Herb, D. C. Hom, L. M. Lederman, J. C. Sens, H. D. Snyder, J. K. Yoh et al., *Observation of a Dimuon Resonance at 9.5 GeV in 400-GeV Proton-Nucleus Collisions*, Physical Review Letters **39** (1977) 252.

[3] Particle Data Group collaboration, *Review of particle physics*, Phys. Rev. D **98** (2018) 030001.

[4] S. L. Glashow, J. Iliopoulos and L. Maiani, *Weak interactions with lepton-hadron symmetry*, Phys. Rev. D **2** (1970) 1285.

[5] J. Charles et al., *Current status of the Standard Model CKM fit and constraints on $\Delta F = 2$ New Physics*, Phys. Rev. **D91** (2015) 073007 [`arXiv:1501.05013`].

[6] BaBar, Belle collaboration, *The Physics of the B Factories*, Eur. Phys. J. **C74** (2014) 3026 [`arXiv:1406.6311`].

[7] J. J. Aubert, U. Becker, P. J. Biggs, J. Burger, M. Chen, G. Everhart et al., *Experimental Observation of a Heavy Particle J*, Physical Review Letters **33** (1974) 1404.

[8] J. E. Augustin, A. M. Boyarski, M. Breidenbach, F. Bulos, J. T. Dakin, G. J. Feldman et al., *Discovery of a narrow resonance in e^+e^- annihilation*, Physical Review Letters **33** (1974) 1406.

[9] D. Andrews, K. Berkelman, R. Cabenda, D. Cassel, J. DeWire, R. Ehrlich et al., *Observation of a fourth Upsilon state in e^+e^- annihilations*, Physical Review Letters **45** (1980) 219.

[10] H. Albrecht, U. Binder, P. Bckmann, G. Drews, R. Glser, G. Harder et al., *Argus: A universal detector at doris ii*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **275** (1989) 1 .

[11] H. Albrecht, A. Andam, U. Binder, P. Böckmann, R. Gläser, G. Harder et al., *Observation of $B^0 \bar{B}^0$ mixing, Physics Letters B* **192** (1987) 245.

[12] H. Albrecht, A. Andam, U. Binder, P. Bckmann, G. Drews, H. Ehrlichmann et al., *Physics with ARGUS, Physics Reports* **276** (1996) 224 .

[13] J. Bartelt, S. Csorna, Z. Egyed, V. Jain, D. Akerib, B. Barish et al., *Measurement of charmless semileptonic decays of b mesons, Physical Review Letters* **71** (1993) 4111.

[14] CLEO COLLABORATION collaboration, *First measurement of the rate for the inclusive radiative penguin decay $b \to s\gamma$, Phys. Rev. Lett.* **74** (1995) 2885.

[15] B. Aubert, A. Bazan, A. Boucham, D. Boutigny, I. D. Bonis, J. Favier et al., *The babar detector, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **479** (2002) 1 .

[16] B. Collaboration et al., *Technical design report, KEK report* **95** (1995) .

[17] BELLE collaboration, *A Measurement of CP violation in B0 meson decays with BELLE*, , `hep-ex/0010008`.

[18] BABAR collaboration, *First CP violation results from BaBar*, , `hep-ex/0011024`.

[19] BABAR COLLABORATION collaboration, *Observation of CP Violation in the B^0 Meson System, Phys. Rev. Lett.* **87** (2001) 091801.

[20] BELLE COLLABORATION collaboration, *Observation of large CP violation in the neutral B meson system, Phys. Rev. Lett.* **87** (2001) 091802.

[21] BABAR COLLABORATION collaboration, *Observation of the Bottomonium Ground State in the Decay $\Upsilon(3S) \to \gamma\eta_b$, Phys. Rev. Lett.* **101** (2008) 071801.

[22] S. Uno, T. Abe, I. Adachi, K. Adamczyk and S. Ahn, *Belle II technical design report*, `arXiv:1011.0352`.

[23] LHCb COLLABORATION collaboration, R. Antunes-Nobrega, A. Frana-Barbosa, I. Bediaga, G. Cernicchiaro, E. Correade-Oliveira, J. Magnin et al., *LHCb reoptimized detector design and performance: Technical Design Report*, Technical Design Report LHCb. CERN, Geneva, 2003.

[24] LHCb collaboration, R. Aaij, B. Adeva, M. Adinolfi, Z. Ajaltouni, S. Akar et al., *Test of lepton universality with $B^0 \to K^{*0}\ell^+\ell^-$ decays*, `arXiv:1705.05802`.

[25] Heavy Flavor Averaging Group collaboration, *Averages of b-hadron, c-hadron, and τ-lepton properties as of Summer 2016*, Eur. Phys. J. **C77** (2017) 895 [`arXiv:1612.07233`].

[26] The Belle Collaboration, A. Abdesselam, I. Adachi, K. Adamczyk, H. Aihara, S. A. Said et al., *Measurement of $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ with a semileptonic tagging method*, `arXiv:1904.08794`.

[27] S. Hirose, T. Iijima, I. Adachi, K. Adamczyk, H. Aihara, S. Al Said et al., *Measurement of the τ Lepton Polarization and R(D*) in the Decay $B \to D^*\tau^-\bar\nu_\tau$* , *Physical Review Letters* **118** (2017) 1.

[28] The BELLE-II Collaboration, *The Belle II Physics book*, `arXiv:1808.10567`.

[29] E. Waheed, P. Urquijo, I. Adachi, K. Adamczyk, H. Aihara, S. A. Said et al., *Measurement of the CKM Matrix Element $|V_{cb}|$ from $B^0 \to D^{*-}\ell^+\nu_\ell$ at Belle*, *arXiv e-prints* (2018) [`arXiv:1809.03290`].

[30] The Belle Collaboration, A. Abdesselam, I. Adachi, K. Adamczyk, H. Aihara, S. A. Said et al., *Precise determination of the CKM matrix element $|V_{cb}|$ with $\bar B^0 \to D^{*+}\ell^-\bar\nu_\ell$ decays with hadronic tagging at Belle*, `arXiv:1702.01521`.

[31] F. U. Bernlochner, Z. Ligeti, M. Papucci and D. J. Robinson, *Tensions and correlations in $|V_{cb}|$ determinations*, *Physical Review D* **96** (2017) 1 [`1708.07134`].

[32] S. Aoki, Y. Aoki, D. Becirevic, T. Blum, G. Colangelo, S. Collins et al., *FLAG Review 2019*, `arXiv:1902.08191`.

[33] F. U. Bernlochner, Z. Ligeti, M. Papucci and D. J. Robinson, *Tensions and correlations in $|V_{cb}|$ determinations*, *Physical Review D* **96** (2017) 1 [`1708.07134`].

[34] J. Flynn, T. Izubuchi, A. Jüttner, T. Kawanai, C. Lehner, E. Lizarazo et al., *Form factors for semi-leptonic B decays*, *Proceedings of Science* **Part F128557** (2016) 0 [`1612.05112`].

[35] A. Bazavov, C. Bernard, N. Brown, C. Detar, A. X. El-Khadra, E. Gámiz et al., *B- And D-meson leptonic decay constants from four-flavor lattice QCD*, *Physical Review D* **98** (2018) [`1712.09262`].

[36] A. V. Avilés-Casco, C. DeTar, D. Du, A. El-Khadra, A. Kronfeld, J. Laiho et al., *B → D*l v at non-zero recoil* , *EPJ Web of Conferences* **175** (2018) 13003.

[37] C. Hughes, C. T. Davies and C. J. Monahan, *New methods for B meson decay constants and form factors from lattice NRQCD*, *Physical Review D* **97** (2018) 1 [`1711.09981`].

[38] J. A. Bailey, T. Bhattacharya, R. Gupta, Y.-C. Jang, W. Lee, J. Leem et al., *Calculation of $\bar{B} → D^*\ell\bar{\nu}$ form factor at zero recoil using the Oktay-Kronfeld action*, `arXiv:1711.01786`.

[39] T. Kaneko, Y. Aoki, B. Colquhoun, H. Fukaya and S. Hashimoto, *$B → D^{(*)}\ell\nu$ form factors from $N_f = 2 + 1$ QCD with Möbius domain-wall quarks*, `arXiv:1811.00794`.

[40] A. Bussone, N. Carrasco, P. Dimopoulos, R. Frezzotti, P. Lami, V. Lubicz et al., *Mass of the b quark and B -meson decay constants from Nf=2+1+1 twisted-mass lattice QCD*, *Physical Review D* **93** (2016) 0 [`1603.04306`].

[41] D. Giusti, V. Lubicz, C. Tarantino, G. Martinelli, C. T. Sachrajda, F. Sanfilippo et al., *First Lattice Calculation of the QED Corrections to Leptonic Decay Rates*, *Physical Review Letters* **120** (2018) [`1711.06537`].

[42] C. Lehner, S. Meinel, T. Blum, N. H. Christ, A. X. El-Khadra, M. T. Hansen et al., *Opportunities for lattice QCD in quark and lepton flavor physics*, `arXiv:1904.09479`.

[43] K. Akai, K. Furukawa and H. Koiso, *SuperKEKB collider*, *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **907** (2018) 188.

[44] Hikino Hajime at KEK PR Office, "Electrons and positrons collide for the first time in the SuperKEKB Accelerator." `https://www.kek.jp/en/newsroom/2018/04/26/0700/`, 2018.

[45] D. Andrews, K. Berkelman, M. Billing, R. Cabenda, D. G. Cassel, J. W. DeWire et al., *Observation of three Upsilon states*, *Phys. Rev. Lett.* **44** (1980) 1108.

[46] T. Böhringer, F. Costantini, J. Dobbins, P. Franzini, K. Han, S. W. Herb et al., *Observation at the Cornell Electron Storage Ring*, *Phys. Rev. Lett.* **44** (1980) 1111.

[47] G. Finocchiaro, G. Giannini, J. Lee-Franzini, R. D. Schamberger, M. Sivertz, L. J. Spencer et al., *Observation of the Υ''' at the Cornell Electron Storage Ring*, Phys. Rev. Lett. **45** (1980) 222.

[48] D. Besson and T. Skwarnicki, *UPSILON SPECTROSCOPY : Transitions in the Bottomonium System*, Annual Review of Nuclear and Particle Science (1993) .

[49] Belle-II collaboration, *Belle II Technical Design Report*, `arXiv:1011.0352`.

[50] L. Andricek, "Operational experience with the DEPFET based PXD pixel detector at Belle II and possible improvements utilizing micro-channel cooling." IAS Mini-Workshop January 2019, `http://ias.ust.hk/program/shared_doc/2019/201901hep/program/` `20190117_2042_am_Ladislav%20Andricek.pdf`, 2019.

[51] T. Kuhr, C. Pulvermacher, M. Ritter, T. Hauth and N. Braun, *The Belle II Core Software*, `arXiv:1809.04299`.

[52] D. J. Lange, *The EvtGen particle decay simulation package*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **462** (2001) 152.

[53] T. Sjstrand, S. Mrenna and P. Skands, *PYTHIA 6.4 physics and manual*, Journal of High Energy Physics **2006** (2006) 026.

[54] S. Agostinelli, J. Allison, K. Amako, J. Apostolakis and H. Araujo, *Geant 4 — a simulation toolkit*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **506** (2003) 250.

[55] W. Waltenberger, W. Mitaroff and F. Moser, *Ravea detector-independent vertex reconstruction toolkit*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **581** (2007) 549 .

[56] J. Krohn, P. Urquijo and F. Abudin, *Global Decay Chain Vertex Fitting at B-Factories*, `1901.11198`.

[57] T. Keck, F. Abudinén, F. U. Bernlochner, R. Cheaib, S. Cunliffe, M. Feindt et al., *An Exclusive Tagging Algorithm for the Belle II Experiment*, Computing and Software for Big Science **8** (2019) 1.

[58] T. Keck, *FastBDT: A Speed-Optimized Multivariate Classification Algorithm for the Belle II Experiment*, Computing and Software for Big Science **1** (2017) 1.

[59] T. Keck, "FastFit library." `https://github.com/thomaskeck/FastFit`, 2019.

[60] N. Toutounji, *Reconstruction Methods for Semi-leptonic Decays of B -mesons with the Belle II Experiment*, Master's thesis, University of Sydney, 2018.

[61] T. Keck, *The Full Event Interpretation for Belle II*, Master's thesis, Karlsruher Instituts für Technologie (KIT), 2014.

[62] BELLE collaboration, *Measurement of the branching fraction of $B^+ \to \tau^+ \nu_\tau$ decays with the semileptonic tagging method*, Phys. Rev. D **92** (2015) 051102.

[63] J. A. Shifflett, "Standard model lagrangian (including neutrino mass terms)." `http://einstein-schrodinger.com/Standard_Model.pdf`, 2015.

[64] K. Olive, *Review of particle physics*, Chinese Physics C **38** (2014) 090001.

[65] S. Aoki, Y.Aoki, D.Bečirević, C.Bernard, T.Blum, G. Colangelo et al., *Review of lattice results concerning low-energy particle physics*, European Physical Journal C **77** (2017) 1 [`arXiv:1607.00299`].

[66] M. F. L. Golterman, *Staggered Mesons*, Nuclear Physics B **273** (1985) 663.

[67] G. P. Lepage, *Flavor-symmetry restoration and Symanzik improvement for staggered quarks*, Phys. Rev. D **59** (1999) 1.

[68] A. S. Kronfeld, *Lattice gauge theory with staggered fermions: how , where , and why (not)*, `0711.0699v`.

[69] A. Bazavov, D. Toussaint, C. Bernard, J. Laiho, C. DeTar, L. Levkova et al., *Nonperturbative qcd simulations with $2 + 1$ flavors of improved staggered quarks*, Rev. Mod. Phys. **82** (2010) 1349.

[70] E. Follana, Q. Mason, C. Davies, K. Hornbostel, G. P. Lepage, J. Shigemitsu et al., *Highly improved staggered quarks on the lattice with applications to charm physics*, Physical Review D - Particles, Fields, Gravitation and Cosmology **75** (2007) 1 [`0610092v1`].

[71] B. Sheikholeslami and R. Wohlert, *Improved continuum limit lattice action for qcd with wilson fermions*, Nuclear Physics B **259** (1985) 572.

[72] W. Bietenholz, V. Bornyakov, M. Göckeler, R. Horsley, W. G. Lockhart, Y. Nakamura et al., *Flavor blindness and patterns of flavor symmetry breaking in lattice simulations of up, down, and strange quarks*, *Physical Review D - Particles, Fields, Gravitation and Cosmology* **84** (2011) 1 [arXiv:1102.5300].

[73] W. Bietenholz, V. Bornyakov, M. Göckeler, R. Horsley, W. G. Lockhart, Y. Nakamura et al., *Flavor blindness and patterns of flavor symmetry breaking in lattice simulations of up, down, and strange quarks*, *Physical Review D - Particles, Fields, Gravitation and Cosmology* **84** (2011) 1 [arXiv:1102.5300].

[74] Z. Sroczynski, *First results from the asymmetric O(a) improved Fermilab action*, *Nuclear Physics B (Proceedings Supplement)* **84** (1999) 971 [9910004].

[75] H. W. Lin and N. Christ, *Nonperturbatively determined relativistic heavy quark action*, *Physical Review D - Particles, Fields, Gravitation and Cosmology* **76** (2007) 1 [0608005].

[76] N. H. Christ, M. Li and H.-W. Lin, *Relativistic heavy quark effective action*, *Phys. Rev. D* **074505** (2007) 1.

[77] C. Bernard, C. Detar, M. D. Pierro, A. X. El-Khadra, R. T. Evans, E. D. Freeland et al., *Tuning Fermilab heavy quarks in 2 + 1 flavor lattice QCD with application to hyperfine splittings*, *Phys. Rev. D* **034503** (2011) 1.

[78] Y. Aoki, N. Christ, J. Flynn, T. Izubuchi, C. Lehner, M. Li et al., *Nonperturbative tuning of an improved relativistic heavy-quark action with application to bottom spectroscopy*, *Physical Review D* **86** (2012) 116003 [arXiv:1206.2554].

[79] O. Witzel, *B -meson decay constants with domain-wall light quarks and nonperturbatively tuned relativistic b-quarks*, in *LATTICE 2013*, pp. 0–6, 2013, 1311.0276v.

[80] R. G. Edwards and B. Joó, *The Chroma software system for lattice QCD*, *Nuclear Physics B - Proceedings Supplements* **140** (2005) 832 [0409003].

[81] V. G. Bornyakov, R. Horsley, Y. Nakamura, H. Perlt, D. Pleiter, P. E. L. Rakow et al., *Flavour breaking effects in the pseudoscalar meson decay constants*, *Physics Letters B* **767** (2017) 366 [arXiv:1612.04798].

[82] Z. Koumi, *Isospin Violating Hadronic Mass Splittings using Lattice QCD+QED*, Ph.D. thesis, The University of Adelaide, 2019.

[83] B. Efron, *Nonparametric estimates of standard error: The jackknife, the bootstrap and other methods*, *Biometrika* **68** (1981) 589.