

Semileptonic $b \rightarrow c$ Form Factors from Lattice Quantum Chromodynamics

Euan McLean

Supervised by Prof. Christine Davies

*Submitted in fulfillment of the requirements for the degree of
Doctor of Philosophy*

April 2019



*The University of Glasgow
College of Science and Engineering*

Declaration of originality

This thesis is my own work, except where explicit attribution to others is made. Chapter 6 is based on work presented in [1], written in collaboration with Christine Davies, Andrew Lytle and Jonna Koponen. The work in Chapter 7 is presented in a paper by the same authors [2].

All results and figures presented are my own, except for Fig. 2.8 which is courtesy of the Particle Data Group [3], Fig. 2.3 from the HFLAV collaboration [4], and Fig. 3.7 from McNeile et. al [5]. I also quote results from an in-progress study [6] in Eq. (5.32).

'This hasn't gone anything like how I hoped' - Dan Hatton

Acknowledgments

I would like to thank my supervisor, Christine Davies, for her guidance and patience as I stumbled through this PhD making every possible mistake along the way.

I also owe a debt of gratitude to Brian Colquhoun and Andrew Lytle, who's advice and support in the opening stages of my time in Glasgow were indispensable. Without you I surely would have perished. You are top humans.

I would also like to thank Jonna Koponen, Chris Bouchard, Dan Hatton and Javad Komijani for an array of very useful conversations. I hope the Lattice [continues to] treat you well.

It also seems traditional to acknowledge the other students in the group, for like friendship or something. They were ok I guess.

Contents

1	Introduction	1
2	Motivation & Tools from the Continuum	3
2.1	Testing the Standard Model	3
2.2	Flavor-Changing Charged Currents	5
2.2.1	CKM Matrix Unitarity	8
2.2.2	Weak Decays	9
2.2.3	$b \rightarrow c$ Transitions and $ V_{cb} $	14
2.2.4	Flavor Anomalies & Lepton Flavor Violation	16
2.3	Strong Interaction Physics	19
2.3.1	Quantum Chromodynamics	19
2.3.2	Chiral Symmetry	21
2.4	Heavy Quark Physics	23
2.4.1	HQET	23
2.4.2	NRQCD	27
3	Lattice Quantum Chromodynamics	31
3.1	Lattice Gauge Fields	32
3.1.1	The Gauge Action	34
3.1.2	Symanzik Improvements of the Gauge Action	35
3.2	Lattice Fermions	38
3.2.1	The Naive Fermion Action & the Doubling Problem	39
3.2.2	Staggered Quarks	40
3.2.3	Highly Improved Staggered Quarks	42
3.3	Heavy Quarks on the Lattice	44
3.3.1	Heavy-HISQ	46
3.3.2	Lattice NRQCD	47
4	Lattice Calculations	49
4.1	Evaluation of Lattice Correlation Functions	49
4.1.1	Generation of Gauge Ensembles	51
4.1.2	Dirac Operator Inversion	54
4.1.3	Staggered Correlation Functions	57

4.2	Analysis of Correlation Functions	60
4.2.1	Fitting Correlation Functions	60
4.2.2	Signal Degradation	64
5	$b \rightarrow c$ Transitions with Lattice NRQCD	67
5.1	NRQCD-HISQ Currents	67
5.2	Relativistic Normalization of the $b \rightarrow c$ Temporal Axial Current . . .	69
5.2.1	Calculation Details	69
5.2.2	Kinetic Mass	71
5.2.3	Decay Amplitude Ratios	72
5.3	$B_{(s)} \rightarrow D_{(s)}\ell\nu$ Form Factors	73
5.3.1	Lattice Setup	74
5.3.2	Correlator Fits	76
5.3.3	Form Factors	77
5.3.4	Results	78
5.3.5	Form Factors from V_0 and S	82
5.4	Non-Perturbative Renormalization Using $B_c \rightarrow \eta_c$ Data	85
5.4.1	Z_S	86
5.4.2	Z_{V_0}	87
5.4.3	Z_{V_k}	88
5.5	Conclusion of NRQCD Work	91
6	$B_s \rightarrow D_s^*\ell\nu$ Axial Form Factor at Zero Recoil from Heavy-HISQ	93
6.1	Motivation	94
6.2	Calculation Details	95
6.2.1	Lattice Setup	95
6.2.2	Correlator Fits	98
6.2.3	Normalization of the Axial Current	102
6.2.4	Extrapolation to the Physical Point	105
6.3	Results	110
6.3.1	$h_{A_1}^s(1)$	110
6.3.2	Implications for $B \rightarrow D^*$	113
6.3.3	HQET Low Energy Constants	115
6.3.4	Extrapolation Stability	116
6.3.5	H_s and D_s^* Masses	119
6.4	Conclusions	121

7	$B_s \rightarrow D_s \ell \nu$ Form Factors at All Physical q^2 from Heavy-HISQ	123
7.1	Motivation	124
7.2	Calculation Details	126
7.2.1	Lattice Setup	126
7.2.2	Analysis of Correlation Functions	130
7.2.3	Vector Current Renormalization	133
7.2.4	Extrapolation to the Physical Point	135
7.3	Results	140
7.3.1	Ratio Method	140
7.3.2	Direct Method	143
7.3.3	Unitarity Test	150
7.3.4	R_{D_s}	151
7.4	Conclusions	152
7.5	Reconstructing Form Factors	153
7.6	Numerical Values for Lattice Results	154
8	Conclusions	159

Introduction

As the LHC continuously refuses to supply exciting new resonances, the high energy physics community places their hope in the intensity frontier to finally break the Standard Model. Subtle differences between experimental measurements and Standard Model predictions are the new rock and roll. As collider experiments collect more data and measurements become more precise, theorists must keep up the pace and improve predictions.

This thesis focuses on the study of calculating form factors for semileptonic $b \rightarrow c$ transitions. These transitions occur between hadrons, bound together by Quantum Chromodynamics (QCD). At the confinement scale ($\sim 1\text{GeV}$), perturbation theory breaks down due to confinement, and the only sensible option is to compute the path integral directly, i.e., via Lattice QCD.

The b quark is difficult to deal with on the lattice, due to its mass being beyond the momentum cutoff imposed by most computationally feasible lattice spacings. Quark masses roughly equal or greater than the momentum cutoff mean discretization effects become too large to control. We calculated $b \rightarrow c$ form factors using two approaches to dealing with the heavy b quark, one employing a non-relativistic action for the b (*NRQCD*), and the other using a relativistic action with masses between the c and the b mass and extrapolating upwards to the b mass (*Heavy-HISQ*). The main take-home from this thesis is the following: **when it comes to semileptonic form factors; NRQCD is on shaky ground, and Heavy-HISQ is an excellent way to live.** If it can be computationally afforded, heavy-HISQ is the superior of the two approaches.

Using NRQCD, we attempted to compute form factors for the $B_{(s)} \rightarrow D_{(s)}l\nu$ decays. In NRQCD, flavour-changing current operators are made of an infinite series of terms in powers of the b -quark velocity v , each requiring their own

normalisation via perturbative matching to continuum QCD. It was discovered during this work that subleading terms in this series that contribute away from zero recoil in fact have a large contribution. Since the perturbative matching calculations for these terms have not been performed, this makes it very difficult to obtain $b \rightarrow c$ form factors at competitive precision using the NRQCD approach (existing results are at the few-percent level).

The NRQCD approach could in principle be saved by finding non-perturbative normalizations of these large subleading terms in the current. I investigated a way of achieving this by comparing NRQCD lattice data to pre-existing and more reliable Heavy-HISQ lattice data, with limited success.

To sidestep the problems with NRQCD, we focused instead on the Heavy-HISQ approach. With this, we successfully calculated the $B_s \rightarrow D_s^* l \nu$ axial form factor at zero recoil. This demonstrated the power of heavy-HISQ and laid the groundwork for a study of both $B_s \rightarrow D_s^* l \nu$ and $B \rightarrow D^* l \nu$ form factors away from zero recoil, which are now underway. We also calculated $B_s \rightarrow D_s l \nu$ form factors throughout the full physical range of momentum transfer. These studies, when combined with future experimental data of the $B_s \rightarrow D_s l \nu$ and $B_s \rightarrow D_s^* l \nu$ decays, will supply new tests of the Standard Model, and new channels to determining the CKM parameter $|V_{cb}|$.

All work reported in this thesis was performed using gluon field ensembles courtesy of the MILC collaboration, accounting for dynamical up, down, strange and charm quarks in the sea [7, 8]. We computed correlation functions using a combination of the MILC code, and HPQCD's NRQCD code.

Motivation & Tools from the Continuum

In this chapter, I lay out the physics context of this work and some theoretical machinery that was useful for this work. This section consists of a definition and the empirical status of the Standard Model. Then, I will expand on the details of the specific sector we are interested in - the flavor sector and the CKM matrix.

I will also summarize some physics machinery useful for this work, namely QCD, chiral symmetry, and effective field theories for heavy quarks.

2.1 Testing the Standard Model

The Standard Model of Particle Physics (SM) [9–11] is, so far, the most successful theory for describing fundamental particles and their interactions. It is an effective Yang-Mills quantum field theory. It is most succinctly defined by listing its symmetries, field content, and the irreducible representations (irreps) of the symmetries that those fields transform under. Below I follow the discussion in [12].

The symmetries are the following. The Lorentz group $SO(3,1)$, the group of coordinate transformations that leave the Minkowski metric invariant, which can be decomposed into $SU(2)_l \times SU(2)_r$ (*left-handed* and *right-handed*). We denote an irrep as (a, b) where a is the σ^z eigenvalue under $SU(2)_l$ transforms, and b is that of $SU(2)_r$. Then there are internal local gauge symmetries:

$$SU(3)_C \times SU(2)_L \times U(1)_Y, \quad (2.1)$$

irreps of which we denote with (x, y, z) , where x, y label the $SU(3)_C$ and $SU(2)_L$ irreps and z is the charge under $U(1)_Y$. The subscript C stands for color (corresponding to the *strong nuclear force*), L denotes the *weak nuclear force*, and Y denotes *hypercharge*.

The field content is: gauge bosons for each of the above gauge symmetries, each transforming in the adjoint of their corresponding symmetry and in the $(1/2, 1/2)$

irrep of the Lorentz group, denoted B_μ , W_μ , G_μ respectively. There are 6 $SU(2)_L$ doublets in the $(1/2, 0)$ Lorentz irrep; the *left-handed fermions*:

$$Q_{1,2,3} = \begin{pmatrix} u_L \\ d_L \end{pmatrix}, \begin{pmatrix} c_L \\ s_L \end{pmatrix}, \begin{pmatrix} t_L \\ b_L \end{pmatrix}, \quad (\mathbf{3}, \mathbf{2}, 1/6) \quad (2.2)$$

$$L_{1,2,3} = \begin{pmatrix} \nu_{e,L} \\ e_L \end{pmatrix}, \begin{pmatrix} \nu_{\mu,L} \\ \mu_L \end{pmatrix}, \begin{pmatrix} \nu_{\tau,L} \\ \tau_L \end{pmatrix}, \quad (\mathbf{1}, \mathbf{2}, -1) \quad (2.3)$$

and 9 $SU(2)_L$ singlets in the $(0, 1/2)$ Lorentz irrep; the *right-handed fermions*:

$$u_{1,2,3}^R = u_R, c_R, t_R, \quad (\bar{\mathbf{3}}, \mathbf{1}, 2/3) \quad (2.4)$$

$$d_{1,2,3}^R = d_R, s_R, b_R, \quad (\bar{\mathbf{3}}, \mathbf{1}, -1/3) \quad (2.5)$$

$$e_{1,2,3}^R = e_R, \mu_R, \tau_R, \quad (\mathbf{1}, \mathbf{1}, -1). \quad (2.6)$$

We have also listed the SM gauge irreps next to each definition. There is also in principle a further set of right-handed $SU(2)_L$ singlets, $\nu_{1,2,3}^R = (\nu_{e,R}, \nu_{\mu,R}, \nu_{\tau,R})$, but these are singlets of the entire SM gauge group so in a phenomenological sense are very much ‘not there’. There is also a Lorentz scalar $SU(2)_L$ doublet, the Higgs H , in gauge irrep $(\mathbf{1}, \mathbf{2}, 1/2)$ [13–15]. H obtains a vacuum expectation value under $\sim 200\text{GeV}$ and causes a breaking of the above gauge group to $SU(3)_C \times U(1)_E$, where $U(1)_E$ is the electromagnetic gauge group mediated by the photon.

There is at present no confirmed evidence of physics beyond the SM (or *new physics*), besides the presence of neutrino (ν) masses [16]. However, there are a number of problems with the SM that heavily imply that there must be new physics. Among the most famous sources of concern are:

- **Dark Matter & Dark Energy** - an estimated 96% of the content of the universe is dark matter [17] and dark energy [18, 19], that does not interact with the SM gauge group (only via gravity), so cannot be explained by the SM.
- **Matter/Antimatter Asymmetry** - the SM requires there to be an equal amount of matter and antimatter in the universe, however, we observe a massive dominance of matter over antimatter [20].
- **Neutrino Oscillations** - different species of neutrinos oscillate into each other over time implying neutrino masses. Neutrino masses cannot be naturally included in the SM.

- **The Hierarchy Problem** - the SM is ‘finely tuned’, the chances of the Higgs taking its current vacuum expectation value is estimated to be one in $\sim 10^{32}$ [21–23]. A more natural value for the vacuum expectation value would be at the same energy scale as the Planck mass.

In a sense, the central goal of particle physics is currently to pin down evidence against the SM. Only once we have detailed knowledge of how it breaks down will we be able to uniquely determine a new theory of fundamental physics.

There are many promising approaches to achieve this. They are traditionally separated into

- **The Energy Frontier** - explore the highest possible energies reachable with accelerators, directly looking for new physics via the production and identification of new states of matter.
- **The Cosmic Frontier** - use the universe as an experimental laboratory and observatory, taking advantage of naturally occurring events to observe indications of new interactions.
- **The Intensity Frontier** - use intense sources of particles from accelerators, reactors, the sun and the atmosphere to make ultra-precise measurements and find subtle deviations from SM predictions.

The work in this thesis contributes to the third approach.

2.2 Flavor-Changing Charged Currents

The SM tests relevant to this work are on quark flavor-changing interactions. Here I will detail the parts of the SM relevant to these interactions, following the discussion in [12].

The $SU(2)_L$ gauge symmetry of the SM is mediated by the vector boson $W = W^1\tau_1 + W^2\tau_2 + W^3\tau_3$, where τ_i are the three $SU(2)$ generators acting on the $SU(2)_L$ doublets defined in the last section. It is convenient to redefine the fields $W = W^+(\tau_0 + i\tau_1) + W^-(\tau_0 - i\tau_1) + W^3\tau_3$. W^\pm, W^3 are the stationary states at low energies due to electroweak symmetry breaking.

The part of the SM Lagrangian that describes the coupling of W^\pm to fermions is given by

$$\mathcal{L}_{\text{FCCC}} = \frac{e}{\sqrt{2}\sin\theta_W} \left(\bar{u}_L^i W^+ d_L^i + \bar{d}_L^i W^- u_L^i + \bar{\nu}_L^i W^+ e_L^i + \bar{e}_L^i W^- \nu_L^i \right), \quad (2.7)$$

where e is the electron charge, θ_W is the Weinberg angle (a parameter of the SM), and $\not{W} = \gamma^\mu W_\mu$ where γ^μ are members of the Clifford algebra acting on fermion spin components. The indices i, j label quark flavor. To understand the interactions these terms cause we must also consider the mass terms for the fermions:

$$\mathcal{L}_{\text{mass}} = y_{ij}^u \left(\frac{v}{\sqrt{2}} \right) \bar{u}_L^i u_R^j + y_{ij}^d \left(\frac{v}{\sqrt{2}} \right) \bar{d}_L^i d_R^j + y_{ij}^e \left(\frac{v}{\sqrt{2}} \right) \bar{e}_L^i e_R^j + (\text{h.c.}). \quad (2.8)$$

These terms come from the coupling of the fermions to the Higgs field, where the Higgs has taken a vacuum expectation value v at low energies. $y_{ij}^{u,d,e}$ are the Yukawa matrices, parameterising the coupling of the fermions to the Higgs. The absence of right-handed neutrinos forbids an analogous term for neutrinos.

Due to the nondiagonal mass terms, i.e. terms in Eq. (2.8) that couple different flavors, the fundamental fermion fields are not stationary states. To obtain more useful field definitions, one rotates the fields to diagonalise these terms

$$\psi_i^L \rightarrow L_{ij}^\psi \psi_j^L, \quad \psi_R^i \rightarrow R_{ij}^\psi \psi_R^j, \quad (2.9)$$

where $\psi = u, d$ or e , and we choose L_{ij}^ψ, R_{ij}^ψ according to

$$L^{\psi\dagger} y^\psi R^\psi \left(\frac{v}{\sqrt{2}} \right) = M^\psi, \quad (2.10)$$

where M^ψ is diagonal. This results in diagonal mass terms. However, this also has the effect of making couplings in $\mathcal{L}_{\text{FCCC}}$ non-diagonal:

$$\mathcal{L}_{\text{FCCC}} = \frac{e}{\sqrt{2} \sin \theta_W} \left(V_{ij} \bar{u}_L^i \not{W}^+ d_L^j + V_{ij}^* \bar{d}_L^i \not{W}^- u_L^j + \bar{\nu}_L^i \not{W}^+ e_L^i + \bar{e}_L^i \not{W}^- \nu_L^i \right). \quad (2.11)$$

V is the famous Cabibbo–Kobayashi–Maskawa (CKM) matrix [24, 25], consisting of parameters that must be fixed by experiment. $V = L^{u\dagger} L^d$ is by construction a unitary matrix ($V^\dagger V = (L^{d\dagger} L^u)(L^{u\dagger} L^d) = L^d L^{d\dagger} = 1$).

There is no non-diagonal flavor structure in the last two terms because we have redefined the neutrino fields: $\nu_L \rightarrow L^{e\dagger} \nu_L$, absorbing the rotation of the e_L fields. This can be done with impunity due to the lack of neutrino mass terms. While the SM does not include neutrino mass terms, it has in fact been experimentally confirmed that neutrinos have mass. It is however known that these masses are extremely small in comparison to the scales of the SM ($m_\nu \lesssim 0.05\text{eV}$) [16]. Any effect this could in principle have via lepton flavor-changing would be much smaller than the current sensitivity of any experiment.

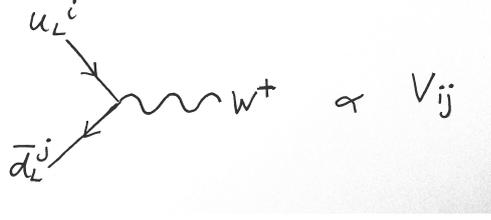


Figure 2.1: The flavor-changing charged current vertex.

Another useful redefinition is to collect the left-handed and right-handed fermion fields into Dirac spinors ψ :

$$\psi = \psi_L + \psi_R, \quad \psi_L = \frac{1}{2}(1 - \gamma^5)\psi, \quad \psi_R = \frac{1}{2}(1 + \gamma^5)\psi. \quad (2.12)$$

In terms of Dirac spinors, $\mathcal{L}_{\text{FCCC}}$ can be written as

$$\begin{aligned} \mathcal{L}_{\text{FCCC}} &= \frac{e}{\sqrt{2}\sin\theta_W} \left(V_{ij} J_\mu^{ij} W^{+\mu} + V_{ij}^* J_\mu^{ij\dagger} W^{-\mu} + L_\mu^{ii} W^{+\mu} + L_\mu^{ii\dagger} W^{-\mu} \right), \quad (2.13) \\ L_\mu^{ij} &= \frac{1}{2} (\bar{\nu}^i \gamma_\mu e^j - \bar{\nu}^i \gamma_5 \gamma_\mu e^j), \\ J_\mu^{ij} &= \frac{1}{2} (\bar{u}^i \gamma_\mu d^j - \bar{u}^i \gamma_5 \gamma_\mu d^j) \equiv V_\mu^{ij} - A_\mu^{ij}. \end{aligned}$$

J_μ^{ij} is known as the Flavor-Changing Charged Current (FCCC). It is often broken up into the *vector* and *axial-vector* components, V_μ and A_μ respectively. These two components can be categorised according to their transformations under the Lorentz group. V_μ is labelled 1^- , where the 1 represents its total spin, and the $-$ represents its negative parity $P: V_\mu \rightarrow -V_\mu$. A_μ is instead labelled 1^+ , due to its positive parity $P: A_\mu \rightarrow +A_\mu$.

We can now turn to the physical consequences of $\mathcal{L}_{\text{FCCC}}$. The interactions given in this part of the Lagrangian describe a quark changing flavor while emitting a W^\pm boson (Fig. 2.1). The propensity for flavor i to decay into another flavor j is governed in part by energy constraints and in part by the associated CKM element V_{ij} . These quark-level interactions mediate meson decays, namely leptonic and semileptonic decays, described in Sec. 2.2.2.

The deviation of V_{ij} from a unit matrix breaks some of the symmetries of the SM. $\mathcal{L}_{\text{SM}} - \mathcal{L}_{\text{FCCC}}$ has the property that one can independently rephase each of the quark fields, $q_i \rightarrow e^{i\theta_i} q_i$, a global $U(1)$ symmetry for each quark flavor. This implies, via Noether's theorem, that the number of quarks of each flavor, N_i , is conserved. However, $\mathcal{L}_{\text{FCCC}}$ breaks this symmetry $U(1)^6 \rightarrow U(1)$. There is only a remnant symmetry of transforming all flavors by the same phase. Individual quark flavor number is no longer conserved, but overall quark number is.

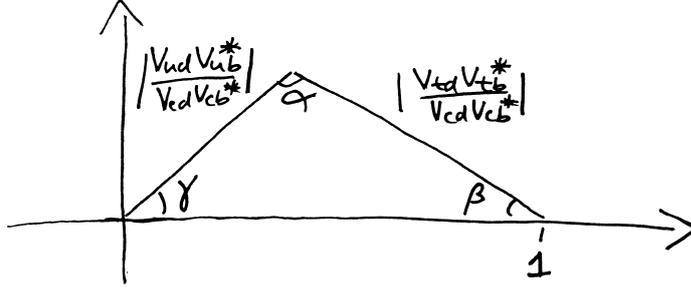


Figure 2.2: A sketch of the unitarity triangle.

Since there is no off-diagonal flavor structure for the leptons, the equivalent global $U(1)^6$ symmetry for the leptons survives in the SM. Individual lepton flavor number is conserved. This property of the SM is referred to as lepton flavor universality.

2.2.1 CKM Matrix Unitarity

The exact values of the CKM matrix elements are of interest in the search for new physics. The CKM matrix is unitary by construction. However, we may discover that the values we measure experimentally do not combine to produce a unitary matrix. This would be evidence that the elements we are measuring, in fact, compose a submatrix of a unitary matrix larger than 3×3 . This would imply the presence of further, heavier quark generations. Below once again I follow [12].

The assumption of unitarity in V :

$$V_{ji}^* V_{jk} = \delta_{ik}, \quad (2.14)$$

imposes 9 constraints on the CKM elements. Each of these constraints gives a test of the SM. If one of these constraints is found to be violated, this would represent evidence of new physics. The most studied constraint is given by taking $i = 3, k = 1$:

$$\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} + \frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*} + 1 = 0. \quad (2.15)$$

This can be visualized as a triangle (known as the *unitarity triangle*) on the complex plane, as shown in Fig. 2.2.

For unitarity, the triangle must close, in other words, $\alpha + \beta + \gamma = \pi$. Hence one test of CKM unitarity is to measure these angles

$$\alpha = \arg \left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*} \right), \quad \beta = \arg \left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*} \right), \quad \gamma = \arg \left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} \right). \quad (2.16)$$

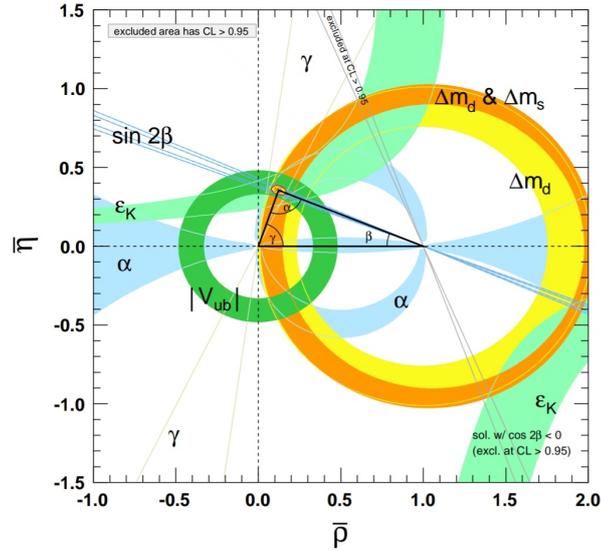


Figure 2.3: Exclusion regions for the vertices of the CKM triangle from various measurements, courtesy of the most recent PDG update [3].

These angles can be constructed from measured CKM elements. They can also be measured directly from certain processes, for example γ can be measured by studying $B \rightarrow DK$ decays [26,27]. The unitarity triangle also contains information about CP-violation from flavor-changing charged currents. The Jarlskog invariant, a measure of CP-violation, is proportional to the area enclosed by the triangle.

The most recent PDG update [3] reports the following averages for the measurements of CKM elements:

$$|V| = \begin{pmatrix} 0.97446 \pm 0.00010 & 0.22452 \pm 0.00044 & 0.00365 \pm 0.00012 \\ 0.22438 \pm 0.00044 & 0.97359^{+0.00010}_{-0.00011} & 0.04214 \pm 0.00076 \\ 0.00896^{+0.00024}_{-0.00023} & 0.04133 \pm 0.00074 & 0.999105 \pm 0.000032 \end{pmatrix}. \quad (2.17)$$

The averages given here are consistent with unitarity in all available tests. The element we are most interested in in this thesis is $|V_{cb}| = 0.04214 \pm 0.00076$, this is the second least precise of the determinations at present. The angles of the unitarity triangle currently satisfy $\alpha + \beta + \gamma = (180 \pm 7)^\circ$. Increasing the precision of CKM determinations is necessary to provide more stringent tests of CKM unitarity.

2.2.2 Weak Decays

I now move on to the methods of determining CKM elements, following discussion in [28]. At the confinement scale ($\sim 1\text{GeV}$ and below), quarks are confined by QCD

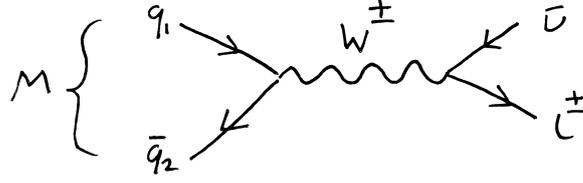


Figure 2.4: Leptonic decay of meson M at tree level in the electroweak coupling.

in hadrons. At these energies, the dynamics of quarks are only experimentally accessible by probing the dynamics of hadrons. CKM matrix elements are determined by studying hadron decays.

First some definitions of hadron classification. Hadrons are categorized into mesons (charged with one valence quark and one valence antiquark) and baryons (three valence quarks). The entirety of this thesis is concerned with mesons. Mesons are categorized in terms of the flavors they are charged under and their representations under the Lorentz group. We use the same notation as for the quantum numbers of the weak currents; L^\pm where L denotes spin and \pm denotes parity. In this thesis, we are concerned mostly with pseudoscalar (0^-) and vector (1^-) mesons.

Weak decays of mesons are categorized according to the final products:

- **Leptonic:** $meson \rightarrow leptons$.
- **Semileptonic:** $meson \rightarrow meson + leptons$.
- **Hadronic:** $meson \rightarrow mesons$.
- **Oscillation:** $meson \rightarrow meson$.

All of these types of decay are dependent on CKM elements so can in principle be used for studying them. We are most interested in the first two, leptonic and semileptonic, so will give detail of such decays here.

Fig. 2.4 shows a generic leptonic decay at tree level in electroweak coupling (virtual quark and gluon lines are implicit). The corresponding amplitude is given by

$$\mathcal{M} = \left(\frac{ie}{\sqrt{2} \sin \theta_W} \right) V_{q_1 q_2} \langle l \bar{\nu} | L_\mu^l D_W^{\mu\nu} J_\nu^{q_1 q_2} | M \rangle, \quad (2.18)$$

where D_W is a free W^\pm propagator, $|M\rangle$ is the ground state of the meson M , and $|l \bar{\nu}\rangle$ is a lepton-antineutrino state. We are using the notation $L_\mu^l = L_\mu^{kk}$, where l

indexes the k th charged lepton. If the momentum of the meson, p^2 , is much smaller than the W mass squared, one can integrate out the dynamics of the W resulting in Fermi effective theory:

$$\begin{aligned} \left(\frac{ie}{\sqrt{2}\sin\theta_W}\right)^2 D_W^{\mu\nu}(p^2) &= \left(\frac{ie}{\sqrt{2}\sin\theta_W}\right)^2 \left(\frac{-ig^{\mu\nu}}{p^2 - M_W^2}\right) \\ &= \frac{i}{M_W^2} \underbrace{\left(\frac{ie}{\sqrt{2}\sin\theta_W}\right)^2}_{\equiv -2\sqrt{2}G_F} g^{\mu\nu} + \mathcal{O}\left(\frac{p^2}{M_W^4}\right). \end{aligned} \quad (2.19)$$

Then \mathcal{M} can be factorised;

$$\mathcal{M} \simeq -2\sqrt{2}G_F V_{q_1 q_2} \langle l\bar{\nu} | L_\mu^l | \Omega \rangle \langle \Omega | J^{q_1 q_2 \mu} | M \rangle. \quad (2.20)$$

$\langle \Omega | J_\mu^{q_1 q_2} | M \rangle$ is a non-perturbative quantity, since it concerns the transitions of a strongly coupled bound state (QCD at the confinement scale). We know that it has a Lorentz index μ , and the only Lorentz vector in the system is the meson's 4-momentum p_μ . So we define

$$\langle \Omega | J_{q_1 q_2}^\mu | M \rangle = p^\mu f_M, \quad (2.21)$$

where f_M is a Lorentz invariant known as the *decay constant* of the meson M , and encodes all non-perturbative information in the amplitude.

By taking the modulus squared of \mathcal{M} and integrating over all allowed momenta of the final state, one finds the decay rate of the process;

$$\Gamma(M \rightarrow l\bar{\nu}) = \frac{G_F^2}{8\pi} f_M^2 m_\ell^2 M_M \left(1 - \frac{m_\ell}{M_M}\right)^2 |V_{q_1 q_2}|^2. \quad (2.22)$$

m_ℓ here is the mass of the final state charged lepton. In order to find $|V_{q_1 q_2}|$, one requires both a measurement of $\Gamma(M \rightarrow l\bar{\nu})$, and a value for f_M . f_M can be computed in a lattice QCD calculation.

A similar story accompanies semileptonic decays. A typical semileptonic decay (at tree level in the electroweak coupling) is depicted in Fig. 2.5. The amplitude is given by

$$\begin{aligned} \mathcal{M} &= \left(\frac{ie}{\sqrt{2}\sin\theta_W}\right) V_{q_1 q_2} \langle M', l\bar{\nu} | J_\mu^{q_1 q_2} D_W^{\mu\nu} L_\nu^l | M \rangle, \\ &\simeq -2\sqrt{2}G_F V_{q_1 q_2} \langle M', l\bar{\nu} | J_\mu^{q_1 q_2} L^{l\mu} | M \rangle, \\ &\simeq -2\sqrt{2}G_F V_{q_1 q_2} \langle l\bar{\nu} | L^{l\mu} | \Omega \rangle \langle M' | J_\mu^{q_1 q_2} | M \rangle, \end{aligned} \quad (2.23)$$

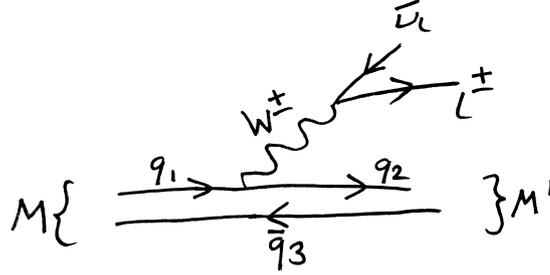


Figure 2.5: Semileptonic decay, $M \rightarrow M' l \bar{\nu}$, at tree level in electroweak coupling.

where on the second line we have integrated out the W propagator by using the same expansion as in the leptonic case, and on the third line we have factorised the QCD part from the electroweak part. The matrix element $\langle M' | J_\mu^{q_1 q_2} | M \rangle$ is a non-perturbative quantity. Unlike in the previous case, there are a number of ways one can choose to parameterise this matrix element, and appropriate choices vary depending on the quantum numbers of M and M' . Of interest to us are the cases where M is a pseudoscalar meson 0^- , and M' is either pseudoscalar or vector 1^- .

In the **pseudoscalar** \rightarrow **pseudoscalar** case, only the vector component of the current survives in the matrix element, $\langle M' | J_\mu^{q_1 q_2} | M \rangle = \langle M' | V_\mu^{q_1 q_2} | M \rangle$. $\langle M' | A_\mu^{q_1 q_2} | M \rangle$ vanishes since this does not respect the parity invariance of QCD (it has negative parity). The most popular parameterisation of $\langle M' | V_\mu^{q_1 q_2} | M \rangle$ is

$$\langle M' | V_\mu^{q_1 q_2} | M \rangle = f_+(q^2) \left[P_\mu + p_\mu - \frac{M^2 - m^2}{q^2} q_\mu \right] + f_0(q^2) \frac{M^2 - m^2}{q^2} q_\mu. \quad (2.24)$$

M, P_μ are the M -meson mass and momentum and m, p_μ are the M' -meson mass and momentum. $f_0(q^2)$ and $f_+(q^2)$, known as the scalar and vector form factors, encoding all non-perturbative information. We now have non-perturbative functions of q^2 rather than a single number. $q^2 = (P - p)^2$, the momentum carried away from the meson by the W , has an allowed range of values if the final states are on-shell (obey the classical dispersion relation):

$$m_\ell^2 \leq q^2 \leq (M - m)^2. \quad (2.25)$$

By integrating $|\mathcal{M}|^2$ over all final lepton and neutrino momenta, one finds a differential decay rate,

$$\frac{d\Gamma}{dq^2}(M \rightarrow M' l \bar{\nu}) = \eta_{\text{EW}} \frac{G_F^2 |V_{q_1 q_2}|^2}{24\pi^3 M^2} \left(1 - \frac{m_\ell^2}{q^2}\right)^2 |\mathbf{p}| \times \quad (2.26)$$

$$\left[\left(1 + \frac{m_\ell^2}{2q^2}\right) M^2 |\mathbf{p}|^2 f_+^2(q^2) + \frac{3m_\ell^2}{8q^2} (M^2 - m^2)^2 f_0^2(q^2) \right].$$

η_{EW} accounts for electroweak corrections due to diagrams where photons or Z s are exchanged in addition to a W^- , as well as the Coulomb attraction of the final-state charged particles [29–31]. \mathbf{p} is spatial momentum of the M' state. Once again, to deduce $|V_{q_1 q_2}|$, one requires both the decay rates $d\Gamma/dq^2$, and the form factors. To precisely determine the form factors requires a Lattice QCD calculation, since this is the only available approach to calculating non-perturbative observables from first principles with a systematic understanding of uncertainties (another approach is *QCD sum rules*, for example see [32]).

In the **pseudoscalar**→**vector** case, both the vector and axial-vector components of the current survive in the matrix element. A common choice of parameterisation is

$$\langle M'(\epsilon) | V_{q_1 q_2}^\mu | M \rangle = i\sqrt{Mm} h_V^s(w) \epsilon_{\mu\nu\alpha\beta} \epsilon^{*\nu} v'^\alpha v^\beta, \quad (2.27)$$

$$\begin{aligned} \langle M'(\epsilon) | A_{q_1 q_2}^\mu | M \rangle &= \sqrt{Mm} [h_{A_1}^s(w)(w+1)\epsilon_\mu^* - \\ &h_{A_2}^s(w)\epsilon^* \cdot v v_\mu - h_{A_3}^s(w)\epsilon^* \cdot v v'^\mu]. \end{aligned} \quad (2.28)$$

$v = P/M$ and $v' = p/m$ are the 4-velocities of M and M' respectively. ϵ is the polarization of the vector meson M' . $w = v \cdot v'$ is known as the recoil parameter, this is an alternative to q^2 often used in heavy quark effective theory. $h_V(w)$, $h_{A_0}(w)$, $h_{A_1}(w)$, and $h_{A_2}(w)$ are the form factors accounting for the non-perturbative physics. The decay rate is given by

$$\frac{d\Gamma}{dw}(M \rightarrow M' l \bar{\nu}) = \frac{G_F^2 m^3 |\eta_{\text{EW}} V_{q_1 q_2}|^2}{4\pi^3} (M-m)^2 \sqrt{w^2-1} \chi(w) |\mathcal{F}(w)|^2, \quad (2.29)$$

where $\mathcal{F}(w)$ is a linear combination of the form factors and $\chi(w)$ is a known function of w (both given in e.g. appendix G of [33]).

At the zero recoil point, where q^2 is maximized at $q_{\text{max}}^2 = (M-m)^2$, (corresponding to $w = 1$), a single form factor contributes:

$$\mathcal{F}(1) = h_{A_1}(1). \quad (2.30)$$

However the differential decay rate vanishes at $w = 1$. A common approach to determine $|V_{q_1 q_2}|$, for example used to find $|V_{cb}|$ via the $B \rightarrow D^* l \bar{\nu}$ decay, is to find $|\mathcal{F}(1) V_{cb}|^2$ at zero recoil by extrapolating from experimental data at non-zero recoil, and combining this with a lattice QCD determination of $h_{A_1}(1)$. This method is used since lattice results for the form factors have only been available at zero recoil, lattice calculations become more complicated away from zero recoil.

2.2.3 $b \rightarrow c$ Transitions and $|V_{cb}|$

The family of weak decays that have attracted the most attention are decays of B -mesons (pseudoscalar mesons containing a valence b and u, d, s or c quark), due to their rich variety of decay products.

The b can decay into either a c or a u quark via the flavor changing charged current. In this thesis we are interested in the $b \rightarrow c$ transition, with an amplitude proportional to the CKM element $|V_{cb}|$. In this section, I'll give a brief overview of how this is calculated and the value's current status.

B meson decays can be measured in a number of experiments. There are two so-called b -factories, the Belle (II) experiment at the KEKB collider in Japan, and the BaBar experiment at the PEP-II collider at SLAC in the US. These are e^+e^- colliders, that collide with an energy tuned to the mass of the $\Upsilon(4s)$, an excited state of the Υ meson (a 1^- state with $\bar{b}b$ valence quarks). The $\Upsilon(4s)$ has a large branching fraction into a $B\bar{B}$ pair, the decays of these can be measured with large statistics. B decays can also be measured in proton colliders, like at the LHCb experiment at CERN. Measurements from LHCb have poorer statistics but cover a larger range of the phase space of final states, due to the variance of momenta in the initial state protons.

So far 3 approaches to determining $|V_{cb}|$ have been carried out:

- $B \rightarrow D^*l\bar{\nu}$ decay rate measurements are extrapolated to zero recoil to determine $|V_{cb}h_{A_1}(1)|$. Then dividing out $h_{A_1}(1)$ from a Lattice calculation, one finds $|V_{cb}|$.
- $B \rightarrow Dl\bar{\nu}$ decay rates are measured throughout q^2 , and combined with $f_0(q^2)$ and $f_+(q^2)$ from lattice calculations.
- $B \rightarrow X_c l\bar{\nu}$ decay rates are measured (where X_c is all possible charmed final state mesons), this is used to constrain elements in the operator product expansion, a method first devised in [34, 35].

The first two are referred to as *exclusive* and the third *inclusive*. A selection of the most accurate examples of each method of determination is given in Fig. 2.6.

This figure tells a story of the recent history of $|V_{cb}|$. Determinations from $B \rightarrow Dl\nu$ have been consistent with, but not as precise as, the other two methods. Until recently, there was a $\sim 3\sigma$ tension between determinations from the $B \rightarrow D^*l\nu$ decay and inclusive decays. A possible explanation of this tension appeared when concern

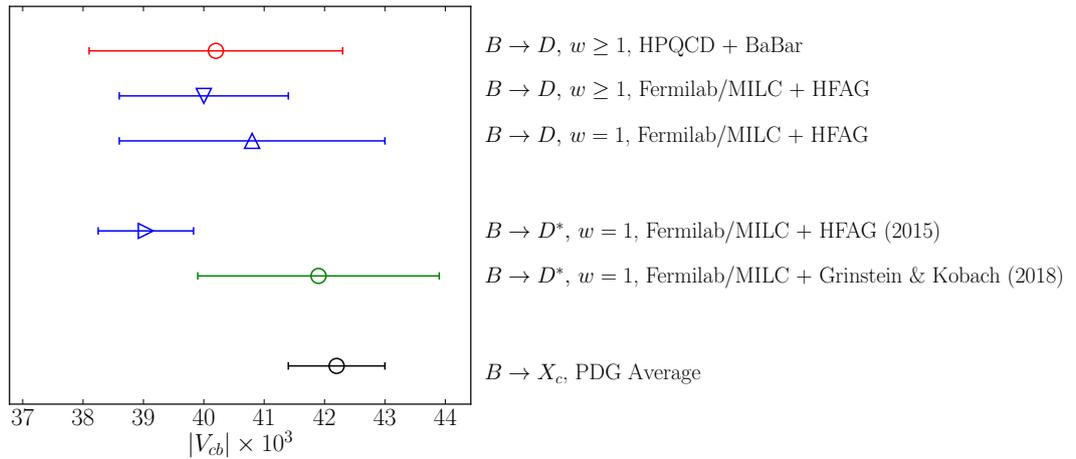


Figure 2.6: Different determinations of $|V_{cb}|$. Points labelled $w = 1$ are determinations from extrapolating measurements of decay rates to the zero recoil point and combining them with a lattice determination of the form factor at zero recoil. Points labelled $w \geq 1$ are results from using a combination of both branching fractions and lattice form factors through some range of w . The first name mentioned in the labels give the source of the lattice form factors, and the second gives the source of the experimental data (e.g. the HPQCD+BaBar point used form factors from the HPQCD collaboration and data from the BaBar experiment). The highest point in red is from [36], the second and third highest from [37], fourth from [38], fifth from [39]. The bottom point is from the PDG [3], using data from the ALPEPH [40], Belle [41], BaBar [42, 43], and CLEO [44] experiments.

was raised about the method of extrapolating experimental data for $B \rightarrow D^* l \bar{\nu}$ decay rates to the zero recoil point ($w = 1$) [39, 45, 46].

The Heavy Flavor Averaging Group HFAG (Now HFLAV) determination of $|V_{cb} h_{A_1}(1)|$ in 2015 parameterized the form factors in the extrapolation using the CLN parameterisation [47]. It has become clear that the constraints imposed on the form factors in the CLN parameterisation are not justified. In [39, 46], the results of an extrapolation using the CLN parameterisation were compared to results from a more general, model-independent parameterisation, the BGL parameterisation [48]. It was found that they differed by 3.5σ . Since BGL is model independent, one may consider this the more reliable result.

The $|V_{cb}|$ result using BGL to extrapolate the decay rates is given in the green point on Fig. 2.6. Hence, if this work is to be trusted, the long-standing $|V_{cb}|$ tension has been resolved.

There are however a number of other reasons to be interested in studying $|V_{cb}|$. It constrains one side of the unitarity triangle via the ratio $|V_{ub}|/|V_{cb}|$, so it is one of the bottlenecks for precise tests of CKM unitarity. It is also a dominant uncertainty in the determination of the CP -violation parameter ϵ_K (that is currently at tension between the SM and experiment, see for example [49] where a 4σ tension is reported).

2.2.4 Flavor Anomalies & Lepton Flavor Violation

The SM can be tested by studying semileptonic decays more directly, without any consideration of CKM elements. CKM-independent observables can be constructed by taking ratios of branching fractions for decays with common CKM dependence. Then, form factors from lattice QCD can be used to form pure SM predictions of these ratios and compared to purely experimental measurements. Such comparisons have uncovered a number of tensions between the SM and experiment.

The ratios are defined by

$$R_{X_q} = \frac{\Gamma(B_q \rightarrow X_q \tau \nu_\tau)}{\frac{1}{2} [\Gamma(B_q \rightarrow X_q e \nu_e) + \Gamma(B_q \rightarrow X_q \mu \nu_\mu)]}, \quad (2.31)$$

where X_q is any meson with valence quark content of $x\bar{q}$. The numerator and denominator will have a common factor of $|V_{bx}|$, so cancel in the ratio.

There is currently tension between SM and experiment in R_D and R_{D^*} :

$$R_{D^*}|_{\text{exp}} = 0.306(13)_{\text{stat}}(07)_{\text{sys}}, \quad R_{D^*}|_{\text{SM}} = 0.258(5), \quad (2.32)$$

$$R_D|_{\text{exp}} = 0.407(39)_{\text{stat}}(24)_{\text{sys}}, \quad R_D|_{\text{SM}} = 0.299(3). \quad (2.33)$$

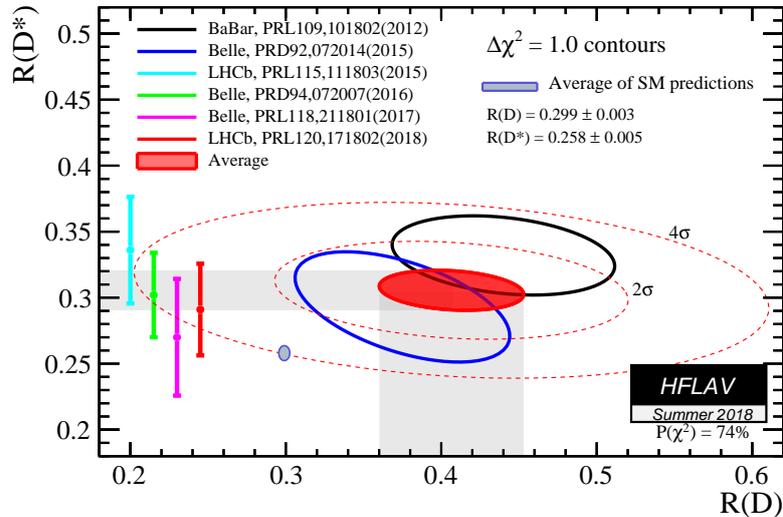


Figure 2.7: $R(D^{(*)})$ determinations from SM and measurement [4]. One can see from this that the SM prediction (the small blue ellipse) is inconsistent with the experimental average (the filled red ellipse).

The experimental values are the HFLAV averages, from BaBar [50,51], Belle [52–55], and LHCb [56–58] data. The value for $R_{D^*}|_{\text{SM}}$ is the average of results from [45,59,60], using a combination of light-cone sum rules and form factor constraints due to heavy quark symmetry. The $R_D|_{\text{SM}}$ value is the average of results from [45,60,61], which used lattice form factors from [36,37].

A joint analysis of R_D and R_{D^*} by HFLAV shows the combined tension to have a significance of 4.0σ (see Fig. 2.7). Clearly more precise experimental results are necessary to either confirm or dismiss this anomaly. While the SM values are currently much more precise than the experimental ones, further work on the theoretical results is necessary. More independent calculations are required to make the SM numbers more robust, such that if this tension ever hits 5σ , we can be confident that it is due to new physics and not some underestimated SM systematic.

There are also tensions in the quantities [62]

$$R_{K^{(*)}} = \frac{\Gamma(B \rightarrow K^{(*)}\mu^+\mu^-)}{\Gamma(B \rightarrow K^{(*)}e^+e^-)}. \quad (2.34)$$

LHCb measured R_K between 1 and 6 GeV, and found a disagreement with the SM value [63,64] of 2.6σ [65]. LHCb also measured R_{K^*} in 2 bins ($0.045 < q^2 < 1.1 \text{ GeV}^2$ and $1.1 < q^2 < 1.6 \text{ GeV}^2$), and reported disagreement with the SM prediction [66–73]

of $2.1\text{-}2.3\sigma$ and $2.4\text{-}2.5\sigma$ respectively [74].

Each of these anomalies points to one potential new physics scenario: lepton flavor violation (LFV). This is a breakdown of the lepton flavor universality in the SM discussed earlier in this section. A consequence of LFV would be that the different lepton generations would no longer have the same coupling to gauge fields. For example, imagine couplings like $U_{ij}\bar{e}_L^i W^+ \nu_L^j$, where U_{ij} is unitary but non-diagonal, then the different lepton generations would have different couplings to W . This can lead to a modification of the $B \rightarrow D^{(*)}\ell\nu$ and $B \rightarrow K^{(*)}\bar{l}l$ decays rates by different amounts depending on the lepton flavors in the final state, resulting in the ratios $R_{D^{(*)}}$, $R_{K^{(*)}}$ deviating from the SM prediction.

There are broadly speaking two ways one can explain LFV. The first is to posit that there are in fact right-handed neutrinos, ν_R , and neutrinos have Dirac mass terms $m\bar{\nu}_L\nu_R$ (from a coupling to the Higgs). Then, the argument preventing the presence of non-trivial lepton flavor structure in $\mathcal{L}_{\text{FCCC}}$ breaks down, we obtain an equivalent of the CKM matrix for leptons (the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix). Lepton flavor violation is then mediated by the W . Neutrinos have in fact already been shown to have mass, the PMNS matrix exists, and its elements have been measured. However, as mentioned already, these effects would be extremely small due to the extremely small mass of the neutrinos. Experiments have looked for evidence of W -mediated LFV processes, $\tau \rightarrow \mu\gamma$ and $\mu \rightarrow e\gamma$, and they found upper bounds for their branching fractions of 4.2×10^{-13} [75] and 3.1×10^{-7} [76] respectively.

Besides there being no evidence for W -mediated LFV, this picture of neutrino masses is not very aesthetically satisfying. It requires unnaturally small Yukawa couplings between the Higgs and the neutrinos. The second, much more popular approach to explaining both LFV and neutrino masses is the existence of new particles.

In the face of evidence against the SM, the most general way to parameterise the space of possible new physics models is to study the Standard Model Effective Theory (SMEFT) [77]. In this approach, one introduces higher dimension, non-renormalisable operators to the SM (the SM has only dimension 4 operators), and impose a hard momentum cutoff Λ . Then the SMEFT is

$$\mathcal{L}_{\text{SMEFT}} = \mathcal{L}_{\text{SM}} + \sum_i \frac{c_i^{(5)}}{\Lambda} \mathcal{O}_i^{(5)} + \sum_i \frac{c_i^{(6)}}{\Lambda^2} \mathcal{O}_i^{(6)} + \dots \quad (2.35)$$

where $\{\mathcal{O}_i^{(d)}\}$ is the set of dimension- d operators that satisfy the symmetries of

the SM, and $\{c_i^{(d)}\}$ are coefficients to be measured, known as Wilson coefficients. Wilson coefficients differing from the SM expectation can be evidence that the SM must be augmented with new fields at energies above Λ . The quantum numbers of the associated operators gives information about the quantum numbers of the new fields.

One can fit the available $B \rightarrow D^{(*)}l\bar{\nu}$ and $B \rightarrow K^{(*)}l\bar{l}$ data to predictions from SMEFT, in order to infer the Wilson coefficients necessary to explain the anomalies. In [78] it was found that $R_{D^{(*)}}$ can be explained with the $d = 6$ operators:

$$\begin{aligned} & (\bar{c}\gamma_\mu P_L b)(\bar{\tau}\gamma^\mu P_L \nu_\tau), \quad (\bar{c}\sigma^{\mu\nu} P_L b)(\bar{\tau}\sigma_{\mu\nu} P_L \nu_\tau), \quad (\bar{\tau}P_L c^c)(\bar{b}^c P_L \nu_\tau), \\ & (\bar{\tau}\gamma_\mu P_R b)(\bar{c}\gamma^\mu P_L \nu_\tau), \quad (\bar{\tau}\gamma_\mu P_L b)(\bar{c}\gamma^\mu P_L \nu_\tau), \quad (\bar{\tau}P_R c^c)(\bar{b}^c \gamma^\mu P_L \nu), \end{aligned} \quad (2.36)$$

where $P_{L/R} = (1 \pm \gamma_5)/2$, $\psi^c = -i(\bar{\psi}\gamma^0\gamma^2)^T$ and $\bar{\psi}^c = -i(\gamma^0\gamma^2\psi)^T$. In [62], a similar process found the operators necessary to explain $R_{K^{(*)}}$:

$$\begin{aligned} & (\bar{s}\gamma_\mu P_L b)(\bar{e}\gamma^\mu e), \quad (\bar{s}\gamma_\mu P_L b)(\bar{\mu}\gamma^\mu \mu) \\ & (\bar{s}\gamma_\mu P_L b)(\bar{e}\gamma^\mu \gamma_5 e), \quad (\bar{s}\gamma_\mu P_L b)(\bar{\mu}\gamma^\mu \gamma_5 \mu) \end{aligned} \quad (2.37)$$

This information, along with constraints from other measurements, strongly reduces the space of possible new physics models that could produce these anomalies. Hot topics include Leptoquarks, Z' models, and partial compositeness [62, 78–80].

2.3 Strong Interaction Physics

The work of this thesis is essentially quantifying the effect that the strong interaction has on branching fractions for semileptonic decays. The strong interaction and the observed pattern of hadrons can be explained with QCD. In this section, I review the fundamental theory and the force's physical features.

2.3.1 Quantum Chromodynamics

QCD is an $SU(3)$ Yang-Mills gauge theory. The Lagrangian is derived by requiring:

- N_f fermion fields transforming in the fundamental representation of the $SU(3)$ gauge group.
- Invariance under that gauge group.
- Renormalizability of all interactions.

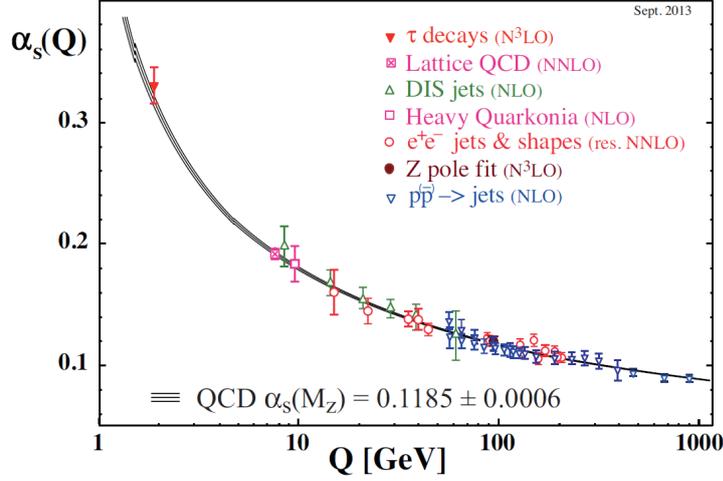


Figure 2.8: The relationship between scale Q and the strong coupling constant α_s , from the PDG [3].

From these we find [12]:

$$\mathcal{L}_{\text{QCD}} = \sum_i \bar{q}_i (i\not{D} - m_i) q_i - \frac{1}{4} \text{Tr} G_{\mu\nu} G^{\mu\nu} - g \frac{\bar{\theta}}{64\pi^2} \epsilon^{\mu\nu\rho\sigma} \text{Tr} G_{\mu\nu} G_{\rho\sigma} \quad (2.38)$$

$$D_\mu = \partial_\mu - igG_\mu, \quad G_{\mu\nu} = [D_\mu, D_\nu].$$

$q_i = (q_{i,r}, q_{i,b}, q_{i,g})$ are the N_f fermions, vectors in *color space*, transforming under

$$q_i(x) \rightarrow \Lambda(x) q_i(x), \quad \bar{q}_i(x) \rightarrow \bar{q}_i(x) \Lambda^\dagger(x), \quad (2.39)$$

where $\Lambda(x)$ is an $SU(3)$ matrix acting on the color space. G_μ are the $\mathfrak{su}(3)$ -valued gluon fields, transforming under the gauge group like

$$G_\mu(x) \rightarrow \Lambda(x) G_\mu(x) \Lambda^\dagger(x) - \frac{i}{g} [\partial_\mu \Lambda(x)] \Lambda^\dagger(x). \quad (2.40)$$

g is the coupling constant of the theory, often expressed instead as $\alpha_s = (g/4\pi)^2$. $\bar{\theta}$ has strong experimental bounds on its size, to the extent that for our purposes that term can be neglected [81].

The most notable feature of QCD is due to the running of α_s [82]. Unlike in quantum electrodynamics where the coupling increases with energy scale, the coupling of QCD decreases as energy scales increase. This is referred to as *asymptotic freedom*. A corollary is that at low energies the coupling becomes strong. At energies around or below $\Lambda_{\text{QCD}} \sim 0.5\text{GeV}$, α_s becomes too large to be a good expansion parameter, and perturbation theory becomes unreliable for making predictions.

At large α_s , quarks and gluons become strongly interacting. This is believed to be the source of confinement, the mechanism that binds quarks together into hadrons.

Broadly speaking there are two approaches to making predictions in QCD at low energies:

1. Chiral perturbation theory - an effective theory of hadrons with the same symmetry properties as QCD.
2. Lattice QCD calculations - solve the path integral by brute force, eliminating the need for an expansion in α_s . This is covered in chapters 3 and 4.

2.3.2 Chiral Symmetry

Here I follow the discussion in [83]. In the limit of $m_i \rightarrow 0 \forall i$, QCD develops two new global symmetries between the flavors:

$$q_i \rightarrow \exp(i\theta_a^V \lambda_a^{ij}) q_j, \quad (2.41)$$

$$q_i \rightarrow \exp(i\gamma_5 \theta_a^A \lambda_a^{ij}) q_j, \quad (2.42)$$

where λ_a are $U(N_f)$ generators. They are labelled $U(N_f)_V$ and $U(N_f)_A$ respectively, standing for vector and axial-vector.

From Noether's theorem, each generator of these symmetries implies a current that is conserved in the massless limit:

$$V_\mu^a = \bar{q} \gamma_\mu \lambda_a q, \quad A_\mu^a = \bar{q} \gamma_\mu \gamma_5 \lambda_a q. \quad (2.43)$$

The (partial) conservation of these currents in quantum mechanics is captured by the *Ward identities*. There is an infinite number of possible Ward identities, but for the purpose of this work, we only need to consider the most simple of them.

Consider the partition function for QCD:

$$\mathcal{Z} = \int [d\psi d\bar{\psi} dA] e^{iS[\psi, \bar{\psi}, A]}, \quad (2.44)$$

where $[d\psi d\bar{\psi} dA]$ represents the functional integral over quark, antiquark and gauge fields. Consider performing a shift of the integration variables of the form (2.41), and allow the parameters θ_a to be local, $\theta_a = \theta_a(x)$. The partition function becomes

$$\mathcal{Z} = \int \mathcal{J}[d\psi d\bar{\psi} dA] (1 + i\delta S) e^{iS[\psi, \bar{\psi}, A]}. \quad (2.45)$$

\mathcal{J} is the Jacobian of the measure $[d\psi d\bar{\psi} dA]$ under the coordinate transform (2.41). In many cases, \mathcal{J} will be non-trivial, due to either regularization schemes that don't respect the symmetry or quantum anomalies. The symmetries we are concerned with here are anomaly free, so $\mathcal{J} = 1$.

The effect of the local version of Eq. (2.41) on the action is

$$\delta S = \int d^4x \theta_a(x) [\partial_\mu V_a^\mu(x) - i\bar{q}(x)[\lambda_a, M]q(x)], \quad (2.46)$$

where $M = \text{diag}(m_u, m_d, m_s, \dots)$ acts on flavor. Removing \mathcal{Z} from each side of (2.45), setting the arbitrary functions $\theta_a(x)$ to 1, and removing the spacetime integral results in

$$\partial_\mu \langle V_a^\mu \rangle = i \langle \bar{q}[\lambda_a, M]q \rangle, \quad (2.47)$$

where $\langle \rangle$ represents a quantum expectation value, the state the expectation value is taken in need not be specified since the above derivation does not assume any particular state. Repeating the above steps with the vector chiral transform replaced with the axial-vector chiral transform, one finds

$$\partial_\mu \langle A_a^\mu \rangle = i \langle \bar{q}\{\lambda_a, M\}q \rangle. \quad (2.48)$$

(2.47) and (2.48) are examples of Ward identities, they describe the (partial) conservation of the chiral currents. (2.47) is often referred to as the Partially Conserved Vector Current (PCVC) relation, and (2.48) the Partially Conserved Axial Current (PCAC) relation.

A useful theorem [84] is that partially conserved currents (currents that become conserved when some parameter in the theory vanishes, like V_a^μ and A_a^μ) require no renormalisation in any regularisation scheme. The conserved or partially conserved current J_a^μ has a corresponding charge $Q_a(t) = \int d^3x J_a^0(\underline{x}, t)$ that is the generator of its corresponding symmetry transform on Hilbert space. In this case, these charges are members of the Lie algebra of the symmetry group:

$$[Q_a(t), Q_b(t)] = i f_{abc} Q_c(t), \quad (2.49)$$

where f_{abc} are the structure constants of the algebra. Under some regularization, change in regularization scheme, or running of scale, each operator in the theory may require multiplicative renormalization $Q_a \rightarrow Z_Q Q_a$. Eq. (2.49) demands that $Z_Q = 1$ in all cases. So J^0 obtains no renormalisation, and if the regularization is Lorentz invariant, this carries on to J^μ .

Since one can transform any flavor into any other flavor via the chiral $U(N_f)$ generators, one can build currents charged with any combination of flavors from linear combinations of V_a^μ and A_a^μ :

$$V_{ij}^\mu = \bar{q}_i \gamma^\mu q_j, \quad \partial_\mu \langle V_{ij}^\mu \rangle = i(m_i - m_j) \langle S_{ij} \rangle \quad (2.50)$$

$$A_{ij}^\mu = \bar{q}_i \gamma^\mu \gamma^5 q_j, \quad \partial_\mu \langle A_{ij}^\mu \rangle = i(m_i + m_j) \langle P_{ij} \rangle \quad (2.51)$$

where we have defined $S_{ij} = \bar{q}_i q_j$ and $P_{ij} = \bar{q}_i \gamma^5 q_j$, the scalar and pseudoscalar densities. The non-renormalisation of V_a^μ and A_a^μ carry on to V_{ij}^μ and A_{ij}^μ , and onto the operators $(m_i - m_j)S_{ij}$, $(m_i + m_j)P_{ij}$ via the Ward identities.

The partially conserved currents V_μ^{ij} and A_μ^{ij} are the same currents that feature in the matrix element of leptonic and semileptonic decays in Sec. 2.2, and their expectation values appear in amplitudes for leptonic and semileptonic decays. Hence, the fact that these can be related to alternative expectation values via ward identities, and that they obtain no renormalisation, is very useful in the calculation of these amplitudes.

2.4 Heavy Quark Physics

Quarks with mass $m_Q \gg \Lambda_{\text{QCD}}$ are referred to as heavy quarks. Charm and bottom quarks are considered heavy: $\Lambda_{\text{QCD}}/m_c \sim 1/4$, $\Lambda_{\text{QCD}}/m_b \sim 1/14$. This separation of scales can come in very useful. They mean one can integrate out some degrees of freedom at m_Q , and still have a good description of the dynamics at Λ_{QCD} . This philosophy gives rise to Heavy Quark Effective Theory (HQET) and Non-Relativistic QCD (NRQCD). Below I will summarise the aspects of this theory most relevant to our work.

2.4.1 HQET

HQET [85] is an effective field theory with the cutoff at the heavy quark mass m_Q , and operators organized in a series in Λ_{QCD}/m_Q . Since at the b (and c) mass QCD is perturbative ($\alpha_s(m_Q) \ll 1$), one can match HQET to perturbative QCD at m_Q , then run the couplings of HQET down to produce useful predictions at the confinement scale.

HQET Lagrangian

Here is the derivation of HQET for a single heavy quark interacting with gluons, following [86] (the generalization to many flavors is straightforward). The fermion part of the Lagrangian is

$$\mathcal{L}_{\text{QCD}} = \bar{Q}(i\not{D} - m_Q)Q, \quad (2.52)$$

where Q is the heavy quark field and D is the covariant derivative (Eq. (2.38)). Define the heavy quark velocity v according to $v = p_Q/m_Q$. Split Q into “heavy” (H) and “light” (h) components:

$$Q = e^{im_Q v \cdot x}(h + H) \quad : \quad h = \frac{1}{2}e^{-im_Q v \cdot x}(1 + \not{v})Q, \quad (2.53)$$

$$H = \frac{1}{2}e^{-im_Q v \cdot x}(1 - \not{v})Q, \quad (2.54)$$

with the important property

$$\not{v}h = h \quad \not{v}H = -H. \quad (2.55)$$

In terms of these new fields the Lagrangian becomes

$$\mathcal{L}_{\text{QCD}} = i\bar{h}(v \cdot D)h - \bar{H}(i(v \cdot D) + 2m_Q)H + i\bar{h}\not{D}^\perp H + i\bar{H}\not{D}^\perp h, \quad (2.56)$$

where $D^\perp = D - v(v \cdot D)$ are the components of D perpendicular to v . A physical interpretation of the definition of h in Eq. (2.53) can be seen by acting a spatial derivative on the definition of h , and by recognising $\partial Q = -ip_Q$, $\partial h = -ip_h$, we find that

$$p_Q = m_Q v + p_h. \quad (2.57)$$

Since $p_h \ll p_Q$, we see that the quark’s momentum is dominated by its mass (the quark is close to on-shell), and the h field represents perturbations around on-shell due to interactions with the lighter degrees of freedom at Λ_{QCD} .

From the Lagrangian (2.56), we see that h is a massless field and H has a mass of $2m_Q$. From this Lagrangian we can derive an equation of motion for H :

$$(i(v \cdot D) + 2m_Q)H = i\not{D}^\perp h, \quad (2.58)$$

with the solution

$$H = \frac{1}{i(v \cdot D) + 2m_Q} i\not{D}^\perp h = \frac{1}{2m_Q} \sum_{n=0}^{\infty} \frac{(-i(v \cdot D))^n}{2m_Q} \not{D}^\perp h. \quad (2.59)$$

By substituting this into the Lagrangian (2.56) we arrive at

$$\mathcal{L}_{\text{HQET}} = i\bar{h}(v \cdot D)h - \bar{h}\not{D}^{\perp} \frac{1}{2m_Q} \sum_{n=0}^{\infty} \frac{(-i(v \cdot D))^n}{2m_Q} \not{D}^{\perp} h. \quad (2.60)$$

Since we expect $v \cdot D \sim \Lambda_{\text{QCD}}$, we can interpret the infinite sum as a series in Λ_{QCD}/m_Q , and truncate it at some order.

Leading order HQET exhibits new symmetries not present in full QCD, known as the heavy quark symmetries. Since m_Q is not present in the leading order Lagrangian, there is a flavor symmetry - a set of N heavy quarks with the same v can be mixed via an $SU(N)$ symmetry. Similarly, due to the absence of spin-mixing matrices, a heavy quark has an $SU(2)$ spin symmetry. At leading order a heavy quark in a meson behaves like a static color charge, the dynamics at Λ_{QCD} are not affected by its mass or spin.

I will now use HQET to derive a useful theorem used in our work, following the proof given in [87].

Luke's Theorem

Luke's theorem [88], which can be derived from the Ademollo-Gatto (AG) theorem [89], tells us the leading order heavy quark mass dependence of form factors. First I will derive the AG theorem.

Consider the transition amplitude

$$\langle \alpha | Q_a | \beta \rangle, \quad (2.61)$$

where Q_a is a conserved charge associated with some global symmetry \mathcal{G} , and $|\alpha\rangle$ and $|\beta\rangle$ belong to some irrep of \mathcal{G} , $\mathcal{R}(\mathcal{G})$. Imagine explicitly breaking the symmetry with a term like $\mathcal{L}_{\text{break}} = \lambda \mathcal{O}_{\text{break}}$. The states $|\alpha\rangle, |\beta\rangle$ are asymptotic states of the complete lagrangian including the symmetry breaking. The breaking causes the states to mix with states in other irreps of \mathcal{G} :

$$|\beta\rangle = c_{\beta\beta}|\beta'\rangle + \sum_m c_{\beta m}|m'\rangle \quad (2.62)$$

$$\langle \alpha | = c_{\alpha\alpha}^* \langle \alpha' | + \sum_n c_{\alpha n}^* \langle n' |. \quad (2.63)$$

$|\alpha'\rangle, |\beta'\rangle$ are pure $\mathcal{R}(\mathcal{G})$ states, and $|m'\rangle, |n'\rangle$ are purely in some other irrep $\mathcal{R}'(\mathcal{G})$.

The transition amplitude becomes

$$\begin{aligned}
\langle \alpha | Q_a | \beta \rangle &= c_{\alpha\alpha}^* c_{\beta\beta} \langle \alpha' | Q_a | \beta' \rangle \\
&+ \sum_m c_{\alpha\alpha}^* c_{\beta m} \langle \alpha' | Q_a | m' \rangle \\
&+ \sum_n c_{\alpha n}^* c_{\beta\beta} \langle n' | Q_a | \beta \rangle \\
&+ \sum_m \sum_n c_{\alpha n}^* c_{\beta m} \langle n' | Q_a | m' \rangle.
\end{aligned} \tag{2.64}$$

Since $|m'\rangle$ is purely in a different irrep to $|\alpha'\rangle$, $Q_a|m'\rangle$ has no overlap with $|\alpha'\rangle$. Similarly for $|n'\rangle$ and $|\beta\rangle$. Hence the second and third terms in Eq. 2.64 vanish. Now consider the order of the coefficients c_{nm} . We can assume that $c_{nm} = \mathcal{O}(\lambda)$ for arbitrary $n, m \neq \alpha, \beta$, since switching off the symmetry breaking by setting $\lambda = 0$ should cause $|\alpha\rangle$ and $|\alpha'\rangle$ to coincide. Then, using the normalization of the states $\sum_n |c_{\alpha n}|^2 = 1$, we find $c_{\alpha\alpha} = \sqrt{1 - \mathcal{O}(\lambda)^2} = 1 + \mathcal{O}(\lambda^2)$, and similarly for $c_{\beta\beta}$. Applying this to the two surviving terms in Eq. (2.64), we end up with

$$\langle \alpha | Q_a | \beta \rangle = c + \mathcal{O}(\lambda^2), \tag{2.65}$$

where $c \neq c(\lambda)$. This is the AG theorem: if the current Q_a and the symmetry breaking term $\mathcal{O}_{\text{break}}$ act orthogonally on the states, the transition amplitude can have at most a second order correction in the symmetry breaking parameter.

Now we can apply this to HQET to produce Luke's theorem. Consider a transition including two heavy quarks (b and c). The theory is now HQET with two heavy quark flavors, $h = (b, c)$. Luke's theorem applies the AG theorem to the breaking of the flavor and spin symmetry at leading order HQET. First consider the flavor breaking. By unpacking the $1/m$ -order terms in the Lagrangian (2.60), we find the leading order flavor breaking term to be

$$\left(\frac{1}{2m_b} - \frac{1}{2m_c} \right) \frac{1}{2} \bar{h} \sigma_z \not{D}^{\perp 2} h, \tag{2.66}$$

where σ_z is the third pauli matrix acting on flavor. These terms cause states like that of a B -meson, $|B\rangle$, to mix with other states $|n'\rangle$ in different irreps of the flavor symmetry. Consider for example the $B \rightarrow D$ decay at zero recoil. Since this is mediated by a generator of the flavor symmetry, the AG theorem leads to

$$\frac{\langle D | \bar{c} \gamma_\mu b | B \rangle}{\sqrt{M_B M_{D^*}}} = \xi + \mathcal{O} \left(\left(\frac{1}{2m_b} - \frac{1}{2m_c} \right)^2 \right), \tag{2.67}$$

where ξ is some b - and c -mass independent number. This motivates a new parameterisation of pseudoscalar-pseudoscalar transition amplitudes alternative to (2.24):

$$\frac{\langle M' | V_\mu^{q_1 q_2} | M \rangle}{\sqrt{Mm}} = h_+(w)(v + v')_\mu + h_-(w)(v - v')_\mu, \quad (2.68)$$

since Eq. (2.67) implies a form of $h_+(1)$ [90]:

$$h_+(1) = \eta_V \left(1 - l_P \left(\frac{1}{2m_b} - \frac{1}{2m_c} \right)^2 \right) + \mathcal{O} \left(\frac{1}{m_c^n m_h^m}, n + m \geq 3 \right). \quad (2.69)$$

η_V is a matching factor between QCD and HQET, and can contain logarithms of heavy masses. The factor l_P is a free non-perturbative parameter that must be fixed by some non-perturbative calculation e.g. a lattice QCD calculation.

Another decay that the AG theorem can be applied to is the $B \rightarrow D^*$ decay. At zero recoil, the amplitude of this decay is $\langle D^* | \bar{c} \gamma_\mu \gamma_5 b | B \rangle$. The operator here is not only a generator of the flavor symmetry but also of the spin symmetry, so we must also take spin breaking into account. The leading order spin breaking terms are

$$\frac{1}{2m_c} \bar{h}_c \gamma^\mu \gamma^\nu G_{\mu\nu} h_c + \frac{1}{2m_b} \bar{h}_b \gamma^\mu \gamma^\nu G_{\mu\nu} h_b. \quad (2.70)$$

So by an analogous argument to that of the $B \rightarrow D$ case, we end up with

$$\frac{\langle D^* | \bar{c} \gamma_\mu \gamma_5 b | B \rangle}{\sqrt{M_B M_{D^*}}} = \xi + \mathcal{O} \left(\left(\frac{1}{2m_b} - \frac{1}{2m_c} \right)^2 \right) + \mathcal{O} \left(\left(\frac{1}{2m_c} \right)^2 \right) + \mathcal{O} \left(\left(\frac{1}{2m_b} \right)^2 \right). \quad (2.71)$$

This carries onto the form factor h_{A_1} at zero recoil [90]:

$$h_{A_1}(1) = \eta_A \left(1 + \frac{l_V}{(2m_c)^2} + \frac{l_A}{2m_b m_c} - \frac{l_P}{(2m_b)^2} \right) + \mathcal{O} \left(\frac{1}{m_c^n m_h^m}, n + m \geq 3 \right). \quad (2.72)$$

where η_A is again a matching factor between HQET and QCD, and $l_{V,A,P}$ are non-perturbative quantities.

2.4.2 NRQCD

An effective field theory closely related to HQET is Non-Relativistic QCD (NRQCD) [91, 92]. This differs from HQET only by the power counting; instead of organizing terms in the Lagrangian according to their order in Λ_{QCD}/m , the terms are organized in terms of powers of the heavy quark's spatial velocity $v \sim |\mathbf{p}|/m$. NRQCD is derived with the following process [93]:

- Separate the quark and antiquark components of the heavy quark. Since a non-relativistic fermion is decoupled from its antiparticle, our action only requires to describe the top two components of a Dirac spinor. Define the antiquark-free 2-component spinor h via the Foldy-Wouthuysen transformation $\psi \rightarrow h = e^{\gamma \cdot \mathbf{D}/2m} \psi$ [94]. This acts to remove the $\gamma \cdot \mathbf{D}$ term from the Dirac part of the Lagrangian, which is the only part that couples the fermion to the anti-fermion (at leading order in $1/m$).
- Define power-counting by considering the expected expectation values of operators for heavy mesons [95]. The three relevant scales concerning the heavy meson are $M, p \sim Mv$ and $E_K \sim Mv^2$, where M is the meson mass, p the spatial momentum and E_K the kinetic energy. By relating operators to these three scales, we can deduce their order in v . Start with the normalization of a scalar current:

$$\langle M | \int d^3x h^\dagger(x) h(x) | M \rangle \sim 1, \quad (2.73)$$

where $|M\rangle$ is some heavy meson state. Since we expect the meson state to be localized in a region of size $1/p$, we can assert that

$$\int d^3x \sim \frac{1}{p^3}. \quad (2.74)$$

From this and (2.73), we find $h \sim p^{3/2} \sim v^{3/2}$. The order of the derivative operator can be deduced from

$$E_K = \langle M | \int d^3x h^\dagger(x) \frac{\mathbf{D}^2}{2M} h(x) | M \rangle, \quad (2.75)$$

to be $D \sim v$. Following such a chain of arguments, we can deduce the order in v of any operator.

- The Lagrangian to $\mathcal{O}(v^n)$ is then simply all of the operators satisfying the symmetries of QCD of order below v^n , with some Wilson coefficients [95]. To

$\mathcal{O}(v^6)$ [93]:

$$\begin{aligned}
\mathcal{L}_{\text{NRQCD}} = & h^\dagger \left(iD_0 + \frac{\mathbf{D}^2}{2m} + c_1 \frac{\mathbf{D}^4}{m^3} + c_2 g \frac{\mathbf{D} \cdot \mathbf{E} - \mathbf{E} \cdot \mathbf{D}}{m^2} \right. \\
& + c_3 ig \frac{\boldsymbol{\sigma} \cdot (\mathbf{D} \times \mathbf{E} - \mathbf{E} \times \mathbf{D})}{m^2} + c_4 g \frac{\boldsymbol{\sigma} \cdot \mathbf{B}}{m} \\
& + f_1 g \frac{\{\mathbf{D}^2, \boldsymbol{\sigma} \cdot \mathbf{B}\}}{m^3} + f_2 ig \frac{\{\mathbf{D}^2, \boldsymbol{\sigma} \cdot (\mathbf{D} \times \mathbf{E} - \mathbf{E} \times \mathbf{D})\}}{m^4} + f_3 ig^2 \frac{\boldsymbol{\sigma} \cdot \mathbf{E} \times \mathbf{E}}{m^3} \left. \right) h \\
& + d_1 \frac{(h^\dagger H)(H^\dagger h)}{m^2} + d_2 \frac{(h^\dagger \boldsymbol{\sigma} H) \cdot (H^\dagger \boldsymbol{\sigma} h)}{m^2} \\
& + d_3 \sum_a \frac{(h^\dagger T^a H)(H^\dagger T^a h)}{m^2} + d_4 \sum_a \frac{(h^\dagger T^a \boldsymbol{\sigma} H) \cdot (H^\dagger T^a \boldsymbol{\sigma} h)}{m^2}. \tag{2.76}
\end{aligned}$$

\mathbf{E} and \mathbf{B} are the chromoelectric and chromomagnetic fields, T^a are fundamental representation of the $SU(3)$ color generators, and H is the antiquark components of the heavy quark. $c_{1,2,3,4}, f_{1,2,3}, d_{1,2,3,4}$ are Wilson coefficients, that can be fixed by perturbative matching to full QCD at the cutoff (the heavy quark mass, where QCD is perturbative).

Lattice Quantum Chromodynamics

At low energies QCD becomes non-perturbative. In other words, the coupling α_s becomes $\mathcal{O}(1)$, and an expansion in α_s (as in perturbation theory) will not be dominated by the leading orders. In order to calculate observables of low energy QCD (like hadronic form factors), we require an alternative to perturbation theory.

The expectation value of an observable \mathcal{O} in QCD can be expressed as a path integral [96]:

$$\langle \mathcal{O} \rangle = \frac{1}{Z} \int [dG d\psi d\bar{\psi}] \mathcal{O} e^{iS[G, \psi, \bar{\psi}]}, \quad (3.1)$$

where G is the gauge field, $\bar{\psi}(\psi)$ are the (anti)fermion fields, S is the classical action, and $[dG d\psi d\bar{\psi}]$ denotes integration over all configurations of the gauge and fermion fields. Z is the partition function. In the perturbative approach, we would expand $\exp(-\text{interacting part of } S)$ resulting in a power series in the gauge coupling populated by Feynman diagrams.

We must instead carry out the integral directly by numerical brute force. Since it is not numerically feasible to carry out an infinite number of integrals, one must approximate spacetime as a discrete 4-dimensional lattice with spacing a between lattice sites, finite spatial volume $L_x^3 = (aN_x)^3$ and finite temporal extent $L_t = aN_t$ ($N_{x,t} \in \mathbb{N}$). The functional integral can be replaced with [97]

$$\int [dG d\psi d\bar{\psi}] = \prod_n \int dU(x_n) d\psi(x_n) d\bar{\psi}(x_n), \quad (3.2)$$

where n is a 4-vector with integer components labelling the sites, and $x_n^\mu = an^\mu$ where $n_\mu \in \mathbb{N}$. This has a second benefit which is to naturally regularize the theory with a momentum cutoff $\Lambda \sim \pi/a$. The gauge field has been replaced with the *gauge link* U , to be defined in the following section.

Typically one uses lattices that have periodic boundary conditions in the temporal direction, i.e $\psi(x + aN_t \hat{t}) = \psi(x)$. This reduces unwanted effects in expectation

values of operators due to the finite temporal extent. In all the work in this thesis, we use periodic temporal boundary conditions.

To avoid having to integrate over imaginary numbers (equivalently to avoid the scourge of the *sign problem* [98]), one also performs a *Wick rotation* [99]. This is the redefinition $t \rightarrow it$, which changes the metric from Minkowski to Euclidean, and changes the weight of the integral $\exp(iS) \rightarrow \exp(-S)$. This is valid since one can rotate the contour of an integral over t from the real line to the imaginary line without crossing any poles. This has the advantage that it turns the quantum path integral into simply an average in statistical mechanics, this means we can apply all of the machinery of statistical mechanics to computing expectation values.

To obtain the ‘real world’ result for some expectation value, where real world means $a = 0$ and volume $= \infty$, one must perform the path integral at a number of different a values, and then extrapolate the results to $a = 0$.

One must choose a discretized version of the QCD action, one that becomes continuum QCD in the $a \rightarrow 0$ limit. This is far from a trivial step. There is an infinite number of choices of lattice actions that become QCD in the continuum limit. There therefore is a huge literature of different choices of discrete lattice actions.

This chapter is dedicated to motivating and detailing the choices of discretized action used in the work of this thesis.

3.1 Lattice Gauge Fields

The discussion of this section follows chapter 5 of [97]. Imagine attempting a naive discretization of the QCD action. Derivatives can be replaced with, for example,

$$\partial_\mu f(x) \rightarrow \frac{1}{2a} (f(x + a\hat{\mu}) - f(x - a\hat{\mu})) , \quad (3.3)$$

where $\hat{\mu}$ is the unit vector in the μ direction. The quark kinetic part of the QCD action, $\bar{q}\not{D}q$, becomes

$$\frac{1}{2a}\bar{q}(x)\gamma_\mu q(x + a\hat{\mu}) - \frac{1}{2a}\bar{q}(x)\gamma_\mu q(x - a\hat{\mu}) - ig\bar{q}(x)G_\mu(x)\gamma_\mu q(x). \quad (3.4)$$

This is no longer invariant under gauge transforms (2.39), for example the first term would become $\bar{q}(x)\Lambda(x)^\dagger\gamma_\mu\Lambda(x + a\hat{\mu})q(x + a\hat{\mu})$. The finite distance between lattice sites force us to think more carefully about the interpretation of gauge symmetry on a lattice.

Formally speaking, a gauge field is a connection on a fibre bundle. We will unpack what this means. At each spacetime point x , there is a space of possible colour vectors that a quark field $q(x)$ could be, call it V_x . V_x is a *fibre*. Spacetime is called the *base space* in this context, there is a fibre at each point in the base space.

The problem with our non-gauge-invariant terms above is that we are trying to compare colour vectors in different fibres. To compare colour vectors at two different fibres, one must *parallel transport* the vector from one point to another, according to some rule of how it should change. Such a rule is called a *connection*. In our case the parallel transport is a Wilson line:

$$\begin{aligned} W(x, y) &: V_y \rightarrow V_x, \\ W(x, y) &= P e^{ig \int_{dc} G}, \end{aligned} \quad (3.5)$$

where c is some curve between x and y , and P orders the operation of the gauge field G on the fibres, i.e. it operates at x first and y last. A Wilson line transforms under the gauge group like $W(x, y) \rightarrow \Lambda(x)W(x, y)\Lambda^\dagger(y)$. This means that operators like $\bar{q}(x)W(x, y)q(y)$ are gauge-invariant, reflecting the fact that the color vector $q(y)$ has been parallel transported into the same fibre as $\bar{q}(x)$.

On a lattice, the natural degrees of freedom are no longer the elements of the Lie algebra, G_μ , but Wilson lines connecting adjacent lattice sites, also known as *links*:

$$U_\mu(x) \in SU(3) : V_x \rightarrow V_{x+a\hat{\mu}}, \quad (3.6)$$

that gauge transform like

$$U_\mu(x) \rightarrow \Lambda(x)U_\mu(x)\Lambda^\dagger(x+a\hat{\mu}). \quad (3.7)$$

Then, a bilinear of color vectors at any two points can be made to be gauge invariant by including a path between them made of links. For example;

$$\begin{aligned} \bar{q}(x)U_\mu(x)q(x+a\hat{\mu}) &\rightarrow [\bar{q}(x)\Lambda^\dagger(x)](\Lambda(x)U_\mu(x)\Lambda^\dagger(x+a\hat{\mu}))[\Lambda(x+a\hat{\mu})q(x+a\hat{\mu})] \\ &= \bar{q}(x)U_\mu(x)q(x+a\hat{\mu}). \end{aligned} \quad (3.8)$$

The $\bar{q}\not{D}q$ term in the QCD Lagrangian can then be represented on the lattice in a gauge invariant way by

$$\frac{1}{2a}\bar{q}(x)\gamma_\mu(U_\mu(x)q(x+a\hat{\mu}) + U_\mu^\dagger(x-a\hat{\mu})q(x-a\hat{\mu})). \quad (3.9)$$

If one defines the links in terms of the the continuum gauge fields G_μ via

$$U_\mu(x) = \exp\left(igaG_\mu\left(x + \frac{a\hat{\mu}}{2}\right)\right), \quad (3.10)$$

then (3.9) takes the correct form in the continuum limit, i.e. it becomes $\bar{q}\not{D}q + \mathcal{O}(a^2)$.

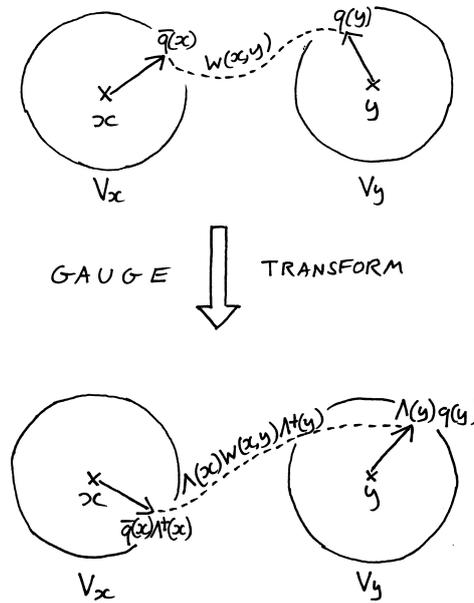


Figure 3.1: Depiction of color spaces (fibres) at two points in spacetime (base space) with the value of the quark field q represented at each point as a color vector, and the connection $W(x,y)$ needed to compare the two color vectors. A gauge transform changes the two vectors in different ways, so for the comparison to be gauge independent the connection must also transform appropriately.

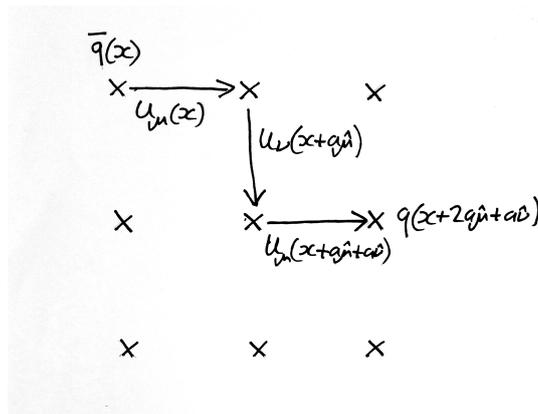


Figure 3.2: Depiction of a gauge invariant quark bilinear, connected by a Wilson line made of gauge links.

3.1.1 The Gauge Action

We must design a pure gauge part of the action in terms of link variables. It is clear that the only gauge invariant operator that depends only on the link variables are

closed loops of links, as in Fig. 3.3.

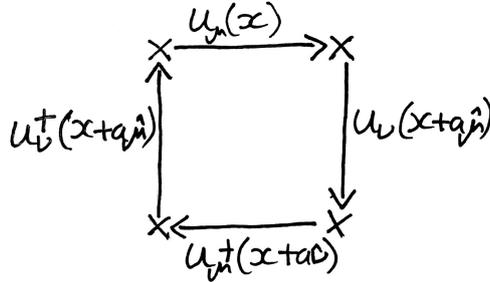


Figure 3.3: The elementary plaquette.

This brings us basically all the way to a legitimate lattice gauge action. The simplest lattice discretisation of the Yang-Mills action is the real part of the smallest possible closed loop of gauge links:

$$S_G = -\frac{1}{g^2} \sum_x \sum_{\mu \neq \nu} \text{Re} \text{Tr}(1 - \square_{\mu\nu}(x)), \quad (3.11)$$

$$\square_{\mu\nu} = U_\mu(x)U_\nu(x + a\hat{\mu})U_\mu^\dagger(x + a\hat{\nu})U_\nu^\dagger(x). \quad (3.12)$$

$\square_{\mu\nu}$ is called the *elementary plaquette*. In the continuum limit this action reduces to

$$S_G = \frac{1}{4} \int d^4x \text{Tr} G_{\mu\nu} G^{\mu\nu} + \mathcal{O}(a^2), \quad (3.13)$$

as required.

In fact, any closed loop reduces to the Yang-Mills action in the continuum. This can be seen intuitively, taking the continuum limit means shrinking any closed loop into an infinitesimally small point. We can choose a gauge action made of any combination of closed loops, so what is the optimal choice?

3.1.2 Symanzik Improvements of the Gauge Action

Any lattice action is admissible for a calculation as long as it reduces to the QCD action in the continuum. This gives us a lot of freedom in how we chose our lattice action.

This freedom can be exploited in order to push expectation values of observables on the lattice closer to their continuum values (reduce the ‘discretisation effects’). This program is known as *Symanzik improvement* [100].

In general, a sensible lattice action can be written as [101]

$$S = \sum_i c_i \mathcal{O}_{\text{lat}}^i = z_0(\{c_i\}) S_{\text{cont}} + a^2 \sum_{n=1} z_n(\{c_i\}) S_n, \quad (3.14)$$

where S_{cont} is the continuum action. We are free to choose any $\{c_i\}$ such that $z_0(\{c_i\}) = 1$. In every example we are concerned with, $\mathcal{O}(a)$ terms are absent, so we ignore them here (the arguments presented here carry straightforwardly to situations with $\mathcal{O}(a)$ corrections). A fundamental postulate of the Symanzik approach is that improvement of one observable (removal of discretisation effects) results in improvement of all other observables. With this in mind, a reasonable approach is:

- Choose some set of lattice operators $\{\mathcal{O}_{\text{lat}}^i\}$ of mass dimension matching the order of a you want to remove. The number of operators required, N , is the number of allowed irrelevant operators in the continuum theory at that dimension. This is because, formally speaking, one needs N tunable c_i values in order to tune N $z_n(\{c_i\})$ values to zero.
- Inspect the continuum limit of the lattice action to find $z_0(\{c_i\})$, enforce $z_0(\{c_i\}) = 1$.
- Choose some observable \mathcal{O} that can be calculated in both the lattice and continuum theory. Use the remaining freedom in $\{c_i\}$ to remove the leading a dependence in $\langle \mathcal{O} \rangle$ order by order in perturbation theory. i.e., if we write the expectation value as

$$\langle \mathcal{O} \rangle = \sum_{n,m} a^{2n} g^{2m} \langle \mathcal{O}_{n,m}(\{c_i\}) \rangle, \quad (3.15)$$

then this amounts to demanding that $\langle \mathcal{O}_{1,m}(\{c_i\}) \rangle = 0$, for as many m 's as possible.

Applying this to pure QCD, this procedure results in the Lüscher-Weisz action [102]. First consider the number of operators required. In continuum pure QCD, the only dimension 4 operator is $\text{Tr} G_{\mu\nu} G^{\mu\nu}$. There are no dimension 5 operators, hence there can be no $\mathcal{O}(a)$ contribution to the continuum limit of a lattice action. There are three independent dimension 6 operators:

$$\begin{aligned} \text{Tr} J_{\mu\nu\rho} J_{\mu\nu\rho}, \quad \text{Tr} J_{\mu\mu\rho} J_{\nu\nu\rho}, \quad \text{Tr} J_{\mu\mu\nu} J_{\mu\mu\nu}, \\ J_{\mu\nu\rho} = [D_\mu, G_{\nu\rho}]. \end{aligned} \quad (3.16)$$

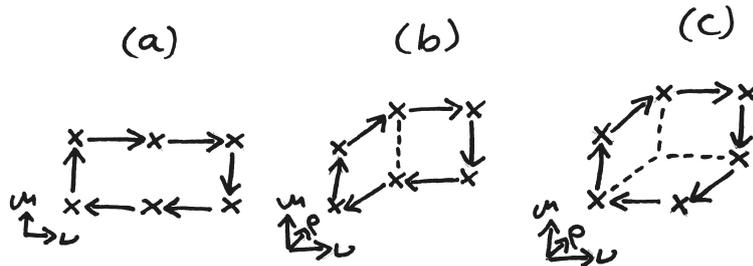


Figure 3.4: Terms additional to the elementary plaquette in the improved pure QCD action.

Hence we require 3 extra operators in the lattice action to be tuned in order to remove the three contributions from the a^2 terms in Eq. (3.14). The simplest choice is to take the plaquette action (3.12), and add all possible Wilson loops containing 6 links. This set consists of three families related by hypercubic invariance, *rectangles* (a), *parallelograms* (b) and *chairs* (c), depicted in Fig. 3.4.

So the new lattice action is

$$S_G = -\frac{1}{g^2} \sum_x \sum_{\mu \neq \nu} (c_0 \text{Re Tr}(1 - \square_{\mu\nu}(x)) + c_1 \text{Re Tr}(1 - \square_{\mu\nu}^a(x)) + \sum_{\rho \neq \mu, \nu} (c_2 \text{Re Tr}(1 - \square_{\mu\nu\rho}^b(x)) + c_3 \text{Re Tr}(1 - \square_{\mu\nu\rho}^c(x)))) \quad (3.17)$$

where $\square_{\mu\nu(\rho)}^{a,b,c}$ are the Wilson loops in fig. 3.4. Expanding this in small a , one finds the function $z_0(\{c_i\})$, setting this to one we find the condition [101]:

$$c_0 + 8(c_1 + c_2) + 16c_3 = 1. \quad (3.18)$$

The rest of the freedom must be fixed by comparing observables in the lattice and continuum theories. In [103] for example, by matching the gluon propagator between the two theories, one constrains the coefficients further to find

$$c_1 = -\frac{1}{12}, \quad c_0 - 8c_3 = \frac{5}{3}. \quad (3.19)$$

These are tree-level relations, so will only prevent lattice artifacts up to $\mathcal{O}(\alpha_s)$. For better improvement, one must compare observables that are sensitive to loop corrections. A popular choice of observable is the static quark potential $V(L)$, this is the potential energy between two static color charges, as a function of separation L between them.

This procedure is affected by the presence of fermions, so it has been performed a number of times to accommodate different fermion discretizations. In this thesis we report results using the Lüscher-Weisz action for gauge fields and Highly Improved Staggered Quarks (defined in Sec. 3.2). The coefficients $\{c_i\}$ were fixed at one-loop in [104] in this context to be

$$c_0 = \frac{5}{3} + (0.237088(46) - 0.1008(34)N_f)\alpha_s + \mathcal{O}(\alpha_s^2), \quad (3.20)$$

$$c_1 = -\frac{1}{12} + (-0.025218(4) + 0.0110(3)N_f)\alpha_s + \mathcal{O}(\alpha_s^2), \quad (3.21)$$

$$c_2 = 0 + (-0.04418(4) + 0.0016(3)N_f)\alpha_s + \mathcal{O}(\alpha_s^2), \quad (3.22)$$

$$c_3 = 0. \quad (3.23)$$

Since these have been tuned to remove a^2 effects up to α_s , lattice artifacts in observables computed using this action will be of size $\mathcal{O}(a^2\alpha_s^2)$, so we say this action is $\mathcal{O}(a^2\alpha_s)$ -improved.

3.2 Lattice Fermions

Putting fermions on the lattice create a much larger host of complications than gauge fields do. There exist a diverse array of approaches to dealing with fermions on the lattice adopted by different collaborations. Different actions are suited to different types of applications. The plethora of fermion actions are necessitated by the famous **doubling problem**, which I will describe below, following [105].

Before beginning the discussion of fermion discretisations, I will define some common notation used for gamma matrices in this context. The Euclidian gamma matrices are defined to obey

$$\{\gamma_\mu, \gamma_\nu\} = 2\delta_{\mu\nu}. \quad (3.24)$$

These have the useful property $\gamma_\mu^2 = 1$. The full set of spin-mixing matrices can be labelled according to

$$\gamma_n = \prod_{\mu} (\gamma_{\mu})^{n_{\mu}}, \quad n_{\mu} = \mathbb{Z}_2. \quad (3.25)$$

We implicitly understand the product to be ordered such that $\mu = 0$ is the rightmost factor and $\mu = 3$ is the leftmost factor. There are 16 such matrices representing corners of the hypercube. One can also use a general site vector x_{μ} to label the matrix, then $\gamma_x = \gamma_n$ where $n_{\mu} = (x_{\mu}/a) \bmod 2$. It is straightforward to show that for any n ; $\gamma_n^{\dagger}\gamma_n = 1$. We also define $\gamma_{5\mu} = i\gamma_5\gamma_{\mu}$, and $\gamma_{5n} = \prod_{\mu}(\gamma_{5\mu})^n$.

3.2.1 The Naive Fermion Action & the Doubling Problem

The interacting Dirac action is most naively discretised with

$$S_F = \sum_{x,\mu} \bar{\psi}(x) \gamma_\mu \nabla_\mu \psi(x) + m \sum_x \bar{\psi}(x) \psi(x), \quad (3.26)$$

where ∇_μ is the gauge covariant finite difference operator,

$$\nabla_\mu \psi(x) = \frac{1}{2a} \left(U_\mu(x) \psi(x + a\hat{\mu}) - U_\mu^\dagger(x - a\hat{\mu}) \psi(x - a\hat{\mu}) \right). \quad (3.27)$$

S_F is invariant under a so-called *doubling symmetry*, which is generated by

$$\psi(x) \rightarrow \mathcal{B}_\mu \psi(x) \equiv (-1)^{x_\mu/a} \gamma_{5\mu} \psi(x), \quad (3.28)$$

$$\bar{\psi}(x) \rightarrow \bar{\psi}(x) \mathcal{B}_\mu^\dagger \equiv (-1)^{x_\mu/a} \bar{\psi}(x) \gamma_{5\mu}^\dagger. \quad (3.29)$$

The product space of these form a group of 16 elements $\{\mathcal{B}_\zeta\}$, labeled by vectors ζ with $\zeta_\mu \in \mathbb{Z}_2$ (e.g. the element $\mathcal{B}_0 \mathcal{B}_1$ is labeled by $\zeta = (1, 1, 0, 0)$).

The physical significance of this symmetry can be seen when we study its effect on the action. First, notice that

$$\mathcal{B}_\mu \psi(x) = \gamma_{5\mu} \sum_k \tilde{\psi}(k) e^{i(k + \frac{\pi}{a}\hat{\mu}) \cdot x} \quad (3.30)$$

$$= \gamma_{5\mu} \sum_k \tilde{\psi} \left(k - \frac{\pi}{a}\hat{\mu} \right) e^{ik \cdot x}, \quad (3.31)$$

where $\{k\}$ is a discrete set of 4-momenta, with $k_\mu = \pi/an_\mu, n_\mu \in [1, N_\mu]$. The action in momentum space can be written as

$$S = \sum_k \bar{\tilde{\psi}}(k) M(k) \tilde{\psi}(k). \quad (3.32)$$

After the operation of \mathcal{B}_μ it becomes

$$S \rightarrow \sum_k \bar{\tilde{\psi}}(k) \gamma_{5\mu} M \left(k + \frac{\pi}{a}\hat{\mu} \right) \gamma_{5\mu} \tilde{\psi}(k). \quad (3.33)$$

Since we know S is invariant under this transformation, it must be true that $\gamma_{5\mu} M \left(k + \frac{\pi}{a}\hat{\mu} \right) \gamma_{5\mu} = M(k)$, and therefore

$$\boxed{M^{-1} \left(k + \frac{\pi}{a}\hat{\mu} \right) = \gamma_{5\mu} M^{-1}(k) \gamma_{5\mu}.} \quad (3.34)$$

This is the doubling problem. M^{-1} is the momentum space propagator for the fermion field, so Eq. (3.34) shows that the spectrum of the fermion is periodic, with

a period of π/a . We expect a pole in $M^{-1}(k)$ where $k \sim m$, where m is the pole mass of the fermion. But due to (3.34) there will now be a second pole at $m + \pi/a$.

Generalizing this argument to all elements of the doubling symmetry, we see that

$$M^{-1}\left(k + \frac{\pi}{a}\zeta\right) = \gamma_{5\zeta}M^{-1}(k)\gamma_{5\zeta}. \quad (3.35)$$

This leads to 16 poles in the fermion spectrum, one for each ζ choice, therefore 16 distinct excitations. We call these excitations *tastes*.

One can isolate a single taste by a block-scaling procedure. Define

$$\psi^{(\zeta)}(x_B) = \frac{1}{16} \sum_{\delta x_\mu \in \mathbb{Z}_2} \mathcal{B}_\zeta(x_B + \delta x)\psi(x_B + \delta x). \quad (3.36)$$

To understand why this only contains one of the tastes, first consider the $\zeta = 0$ case. This would only contain the original non-doubler taste since all other poles at $|k| \sim \pi/a$ have been integrated out. For $\zeta \neq 0$, the \mathcal{B}_ζ operator pushes the ζ doubler to where the $\zeta = 0$ taste originally was in k space, then the blocking procedure integrates out the rest.

3.2.2 Staggered Quarks

There are a number of solutions to the doubling problem. The most straightforward is to modify the action to push the mass of the unwanted tastes above the momentum cutoff, preventing it from influencing the dynamics. These are called *Wilson-type fermions* [106]. However, actions of this type explicitly break Chiral symmetry. Among other issues, this causes additive renormalization of the fermion mass, immensely complicating renormalization procedures.

Another approach, known as *staggered fermions* [107], partially resolves the doubling issue while retaining a remnant chiral symmetry. The work presented in this thesis makes extensive use of the staggered formalism.

Staggered fermions are defined via the following. Redefine the fields according to

$$\psi(x) = \Omega(x)\chi(x), \quad (3.37)$$

where $\Omega(x) = \gamma_x$. In terms of the new spinor variables $\chi(x)$, the naive action given in Eq. (3.26) becomes

$$S_F = \sum_{x,\mu} \bar{\chi}(x)(\alpha_\mu(x)\nabla_\mu + m)\chi(x) \quad (3.38)$$

where $\alpha_\mu(x) = (-1)^{\sum_{\nu < \mu} x^\nu/a}$. The action is now diagonal in spin, leading to 4 grassman variables with identical actions and identical coupling to the gauge field. As a result, χ propagators (on fixed gauge backgrounds) are spin-diagonal:

$$M_\chi^{-1}(x, y) = g(x, y) \mathbf{1}_{\text{spin}}, \quad (3.39)$$

where $g(x, y)$ is a scalar in spin. One need only to include a single component of χ in a simulation (i.e. fix $\chi = (\chi_1, 0, 0, 0)$), then they can compute $M_\chi^{-1}(x, y)[U]$ to obtain $g(x, y)$. Then, using the inverse of Eq. (3.37), $g(x, y)$ can be transformed to a propagator of the original spinors:

$$M_\psi^{-1}(x, y) = g(x, y) \Omega(x) \Omega^\dagger(y). \quad (3.40)$$

This is clearly computationally beneficial since one only needs to simulate one spinor component. But also, by only having one spinor component, one reduces the number propagating degrees of freedom by a factor of 4, cutting the number of tastes from 16 down to 4.

I will show more explicitly how this happens. Consider rewriting an isolated taste (as in Eq. (3.36)) in the staggered formalism, i.e., in terms of χ :

$$\psi^{(\zeta)}(x_B) = \frac{1}{16} \sum_{\delta x_\mu \in \mathbb{Z}_2} \Omega(\delta x) \mathcal{B}_\zeta(0) \chi(x + \delta x). \quad (3.41)$$

Recall we set $\chi(x) = (\chi_1(x), 0, 0, 0)$. The product $\Omega(\delta x) \mathcal{B}_\zeta(0)$ is simply a product of gamma matrices, so can only serve to “scramble” the elements of χ . Then, in the staggered formalism, all 16 tastes $\psi^{(\zeta)}$ amount to only 4 distinguishable fermions: $(\chi_1, 0, 0, 0)$, $(0, \chi_1, 0, 0)$, $(0, 0, \chi_1, 0)$, $(0, 0, 0, \chi_1)$ (with factors of (-1) and i).

To obtain a useful new notation for staggered quarks, we can rewrite Eq. (3.41) as

$$\psi^{\alpha a}(x_B) = \frac{1}{8} \sum_{\eta_\mu \in \mathbb{Z}_2} \gamma_\eta^{\alpha a} \chi(x_B + a\eta). \quad (3.42)$$

$\psi^{\alpha a}$ has spin α and taste a . Define the *spin-taste* notation for operators on $\psi^{\alpha a}$ as $(\gamma_n \otimes \gamma_m)$, where γ_n acts on the spin component α and γ_m acts on the taste component a .

The first operator in the spin-taste notation corresponds to regular spin in the

continuum. This can be seen by writing the free fermion action in terms of ψ^{aa} :

$$S = \sum_{x_B, \mu} \bar{\psi}(x_B) [(\gamma_\mu \otimes 1) \nabla_\mu + \frac{m}{4} (1 \otimes 1)] \psi(x_B) + \mathcal{O}(a^2), \quad (3.43)$$

$$\nabla_\mu \psi(x_B) = \frac{1}{4a} \left(U_\mu(x_B) U_\mu(x_B + a\hat{\mu}) \psi(x_B + 2a\hat{\mu}) - U_\mu(x_B^\dagger - a\hat{\mu}) U_\mu^\dagger(x_B - 2a\hat{\mu}) \psi(x_B - 2a\hat{\mu}) \right). \quad (3.44)$$

If we interpret $(\gamma_\mu \otimes 1)$ as a gamma matrix acting on spin in the continuum, we obtain the continuum Dirac action in the $a \rightarrow 0$ limit.

Hence, to reproduce some current $\bar{\psi} \gamma_n \psi$ in the continuum, one can use $\bar{\psi}(\gamma_n \otimes \gamma_m) \psi$ on the lattice, where we have the freedom to choose any γ_m . In terms of χ fields, these look like

$$\bar{\psi}(x_B) (\gamma_n \otimes \gamma_m) \psi(x_B) = \sum_{\eta, \eta'} \text{Tr}(\gamma_\eta \gamma_n \gamma_{\eta'} \gamma_m) \chi^\dagger(x_B + a\eta) \chi(x_B + a\eta'). \quad (3.45)$$

The $n = m$ case results in local operators in terms of χ , since the trace will vanish unless $\eta = \eta'$. To build the case with $n \neq m$, one must use 'point-split' operators, i.e. $\chi^\dagger(x) \chi(x + \delta x)$. Each choice of δx corresponds to a different *meson taste*, i.e. a different combination of tastes in the two valence quarks of the meson.

In practice in lattice calculations, the remaining 4-fold multiplicity of tastes is tackled in 3 steps:

1. Ensure only one meson taste is created and destroyed at the source and sink of correlation functions.
2. Minimize the interaction between tastes by a modification of the action.
3. Remove contributions of extra tastes in the fermion sea by taking $\det M \rightarrow \sqrt[4]{\det M}$ (the context required to understand this step is elucidated in Sec. 4.1.1).

3.2.3 Highly Improved Staggered Quarks

Step 2 above is the guiding principle for the action we use in much of this work, the Highly Improved Staggered Quark (HISQ) action [105].

The interaction between different tastes ("taste mixing") is dominated by the process in Fig. 3.5, the exchange of single gluons carrying momenta close to $\zeta\pi/a$. In HISQ, this is suppressed by modifying the gauge fields in such a way as to

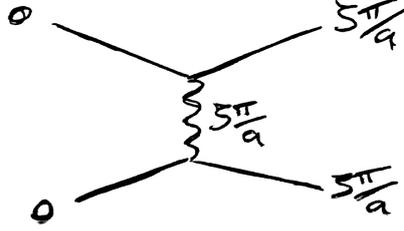


Figure 3.5: Taste mixing at tree level.

minimize the coupling between a gluon with momentum $\zeta\pi/a$ and the fermions, in other words, minimize the vertices in Fig. 3.5.

To this end, one can change the action so that fermions only couple to *smeared* gauge links, in which high-frequency excitations have been removed. Define the first and second covariant derivative operators:

$$\begin{aligned} \delta_\rho U_\mu(x) \equiv & \frac{1}{a} (U_\rho(x) U_\mu(x + a\hat{\rho}) U_\rho^\dagger(x + a\hat{\mu}) \\ & - U_\rho^\dagger(x - a\hat{\rho}) U_\mu(x - a\hat{\rho}) U_\rho(x - a\hat{\rho} + a\hat{\mu})), \end{aligned} \quad (3.46)$$

$$\begin{aligned} \delta_\rho^{(2)} U_\mu(x) \equiv & \frac{1}{a^2} (U_\rho(x) U_\mu(x + a\hat{\rho}) U_\rho^\dagger(x + a\hat{\mu}) \\ & - 2U_\mu(x) \\ & + U_\rho^\dagger(x - a\hat{\rho}) U_\mu(x - a\hat{\rho}) U_\rho(x - a\hat{\rho} + a\hat{\mu})). \end{aligned} \quad (3.47)$$

With this we can define the smearing operator:

$$\mathcal{F}_\mu = \prod_{\rho \neq \mu} \left(1 + \frac{a^2 \delta_\rho^{(2)}}{4} \right). \quad (3.48)$$

HISQ uses two different smeared gauge fields defined by

$$X_\mu(x) \equiv \mathcal{U} \mathcal{F}_\mu U_\mu(x), \quad (3.49)$$

$$W_\mu(x) \equiv \left(\mathcal{F}_\mu - \sum_{\rho \neq \mu} \frac{a^2 (\delta_\rho)^2}{2} \right) \mathcal{U} \mathcal{F}_\mu U_\mu(x), \quad (3.50)$$

where \mathcal{U} is a re-unitarization operator, that acts on a matrix A like $\mathcal{U}A = A/\sqrt{A^\dagger A}$.

The HISQ action can then be written as:

$$S_{\text{HISQ}} = \sum_x \bar{\psi}(x) \left(\sum_{\mu} \gamma_{\mu} \left(\nabla_{\mu}(W) - \frac{a^2}{6} (1 + \epsilon_{\text{Naik}}) \nabla_{\mu}^3(X) \right) + m \right) \psi(x), \quad (3.51)$$

where $\nabla_{\mu}(Z)$ is the covariant derivative (3.27) with gauge links replaced with Z . $\psi(x)$ are the naive form of the fermions, i.e. $\psi(x) = \Omega(x)\chi(x)$. This action in fact not only removes tree level interactions like Fig. 3.5, but also all taste mixing interactions at 1-loop.

The ∇_{μ}^3 term is a Symanzik improvement, it reduces the size of discretisation effects of observables computed using this action. The value of ϵ_{Naik} is fixed according to the constraint

$$\lim_{\mathbf{p} \rightarrow 0} \frac{E^2(\mathbf{p}) - m^2}{\mathbf{p}^2} = 1. \quad (3.52)$$

where $E(\mathbf{p})$ obeys the tree-level dispersion relation from the HISQ action. This means HISQ fermions obey $E^2 = \mathbf{p}^2 + m^2$ at tree level. Tuning ϵ_{Naik} according to this constraint gives us the expression [108]

$$\epsilon_{\text{Naik}} = \frac{4 - \sqrt{4 + 12 \frac{m_{\text{tree}}}{\cosh(m_{\text{tree}}) \sinh(m_{\text{tree}})}}}{\sinh^2(m_{\text{tree}}) - 1}, \quad (3.53)$$

where m_{tree} is the tree-level pole mass given by the expansion [105]

$$m_{\text{tree}} = m \left(1 - \frac{3}{80} m^4 + \frac{23}{2240} m^6 + \frac{1783}{537600} m^8 \right. \quad (3.54)$$

$$\left. - \frac{76943}{23654400} m^{10} \right) + \mathcal{O}(m^{12}). \quad (3.55)$$

3.3 Heavy Quarks on the Lattice

The large hierarchy of different quark masses in nature present a number of further complications to lattice calculations. u and d quarks cause huge problems due to how light they are, this will be addressed in Sec. 4.1.2. s quarks are easy (in the sense that their mass doesn't cause complications).

As quarks get heavier, we begin to encounter another problem. Discretisation effects will generally grow like the largest scale in the theory. If the observable being computed on the lattice is sensitive to the dynamics of a heavy quark of mass

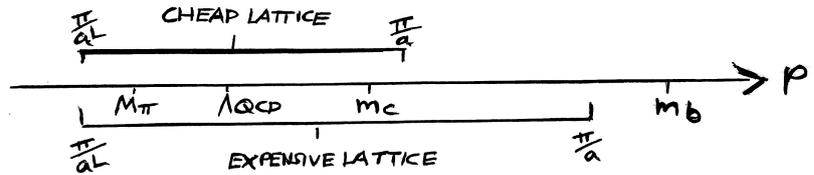


Figure 3.6: Different scales relevant to non-perturbative physics, and brackets showing the range of scales that typical lattices can resolve. The larger the range of scales resolved, the more computationally expensive the calculation.

m_h , this will contain discretisation effects of size $(am_h)^n$ (where n depends on how improved the action is). This is essentially due to the de Broglie wavelength of the heavy quark excitations being close to the lattice spacing, the excitations 'hide' in-between lattice sites.

How heavy we can go is limited by two factors: the improvement of the lattice action and the lattice spacing. How fine we can get the lattice spacing is limited by computational cost. The physical size of the lattice must always be at least large enough to fit the lightest degrees of freedom in the system, namely, it must be larger than the wavelength of pions. This means to get smaller lattice spacing requires increasing the number of sites on the lattice, hence increasing the computational costs (details in chapter 4).

In the past, c quarks resulted in uncontrollable discretisation effects but now armed with highly improved actions like HISQ, and very fine lattices, c physics has been conquered on the lattice [105, 109–111]. Other approaches besides HISQ are also capable of charm physics, see for example [112].

Physics of the b quark is less well developed since the b mass is so much heavier than the c . The b can only be resolved by the very finest of lattice spacings available, and using such fine lattices can be prohibitively costly. Putting a physical b quark on coarser lattices will create uncontrollable discretization effects.

The work in this thesis concerns the decays of mesons containing b quarks. We approach the issue of the heavy b in two different ways, the *heavy-HISQ* approach, and the *Lattice NRQCD* approach. Since the main results of this thesis come from our heavy-HISQ studies, I will not go into too much detail in describing lattice NRQCD.

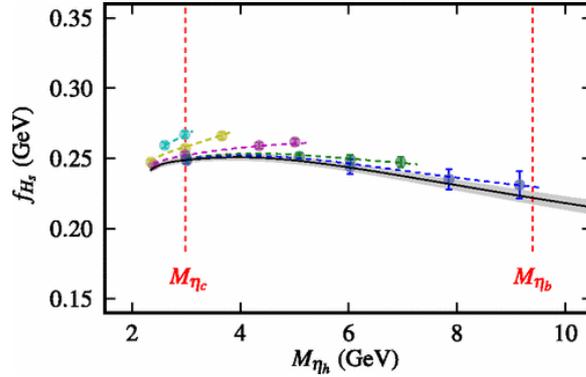


Figure 3.7: An extrapolation to $m_h = m_b$ of the H_s decay constant (where H_s is a pseudoscalar $\bar{h}s$ meson) [5]. The colorful points are measurements of f_{H_s} on the lattice, the color denotes lattice spacing. The grey band shows the result of a continuum extrapolation of this lattice data. The x axis, M_{η_h} , is a proxy for the h -quark mass.

3.3.1 Heavy-HISQ

The heavy-HISQ approach is essentially to model the b with the HISQ action, but to perform the calculation at a number of unphysically light b masses (that we refer to generically as heavy h quarks), and extrapolate the results to the physical b mass. Typically the h masses span most of the region between the c mass and the b mass.

Luckily there exists an effective field theory for understanding how to perform such an extrapolation - HQET. HQET gives a framework to describe how observables depend on masses of heavy quarks, so one can use HQET to derive fit forms of such an extrapolation.

Heavy-HISQ has so far been used for computing b decay constants and masses [5, 113, 114]. In [115], the approach was used to determine c and b quark masses and produce a new determination of α_s . A number of heavy-HISQ calculations of semileptonic form factors are currently underway. The work presented in chapters 6 and 7 adopt the heavy-HISQ approach for computing $B_s \rightarrow D_s^* \ell \nu$ and $B_s \rightarrow D_s \ell \nu$ form factors. Besides these, there are also currently ongoing calculations of form factors for the $B_c \rightarrow \eta_c \ell \nu$, $B_c \rightarrow J/\psi \ell \nu$ [116], $B_c \rightarrow B_s \ell \nu$, $B_s \rightarrow \eta_s \ell \nu$, and $B \rightarrow D^* \ell \nu$ decays.

3.3.2 Lattice NRQCD

The root of the problem of heavy quarks on the lattice is in the rest mass of the quark. Consider the expansion in momentum \mathbf{p}^2 of the continuum relativistic dispersion relation:

$$\omega = \sqrt{\mathbf{p}^2 + m^2} \simeq m + \frac{\mathbf{p}^2}{2m} - \frac{\mathbf{p}^4}{4m^3} + \dots \quad (3.56)$$

The rest mass in the first term is the source of the issue, when $m > \pi/a$ the first term pushes the frequency of excitations ω close to or over π/a .

One could replace the relativistic fermion action e.g. HISQ, with a lattice version of NRQCD [95]. In NRQCD the b has no rest mass, so b excitations will have frequencies much smaller than π/a .

Another benefit of NRQCD is that it does not suffer from a doubling problem since the doubling problem is a purely relativistic issue (the doubling symmetry requires 4 component spinors for γ matrices to act on).

The lattice calculations we perform require us to compute propagators for b quarks on fixed gauge backgrounds. The form of the action allows propagators $G_b(\mathbf{x}, t)[U]$ to be computed using a simple recursion relation

$$G_b(\mathbf{x}, t+1)[U] = e^{-aH}[U]G_b(\mathbf{x}, t)[U], \quad (3.57)$$

which is numerically very fast in comparison to how HISQ propagators are computed (see Sec. 4.1.2). H is the NRQCD Hamiltonian. In the interest of numerical stability, the time evolution operator is re-cast as [95]

$$e^{-aH} = \left(1 - \frac{a\delta H}{2}\right) \left(1 - \frac{aH_0}{2n}\right)^n U_0^\dagger(\mathbf{x}, t) \left(1 - \frac{aH_0}{2n}\right)^n \left(1 - \frac{a\delta H}{2}\right), \quad (3.58)$$

where n is an arbitrary integer (chosen in our studies to be $n = 4$), and the Hamiltonian has been broken up into a leading part H_0 and correction δH . G_b here are propagators for the 2-component spinor fields used in NRQCD. We use the $\mathcal{O}(\alpha_s v^4)$

corrected NRQCD Hamiltonian [95]:

$$\begin{aligned}
aH_0 &= -\frac{\nabla^{(2)}}{2am_b}, & (3.59) \\
a\delta H &= -c_1 \frac{(\nabla^{(2)})^2}{8(am_b)^3} + c_2 \frac{i}{8(am_b)^2} (\nabla \cdot \tilde{\mathbf{E}} - \tilde{\mathbf{E}} \cdot \nabla) \\
&\quad - c_3 \frac{1}{8(am_b)^2} \sigma \cdot (\nabla \times \tilde{\mathbf{E}} - \tilde{\mathbf{E}} \times \nabla) \\
&\quad - c_4 \frac{1}{2am_b} \sigma \cdot \tilde{\mathbf{B}} + c_5 \frac{\nabla^{(4)}}{24am_b} \\
&\quad - c_6 \frac{(\nabla^{(2)})^2}{16n(am_b)^2}, & (3.60)
\end{aligned}$$

where $\nabla^{(2,4)}$ are the second and fourth lattice derivatives, σ are $SU(2)$ matrices acting on spin, and $\tilde{\mathbf{E}}$ and $\tilde{\mathbf{B}}$ are the (Symanzik improved) chromoelectric and chromomagnetic fields. The form of $\nabla^{(2,4)}$, $\tilde{\mathbf{E}}$, $\tilde{\mathbf{B}}$ were defined in Sec. 4.2 of [95] and improved upon in [117].

The coefficients $\{c_i\}$ have been fixed via various calculations adopting a number of methods. The coefficients of the kinetic terms, $c_{1,5,6}$, were most recently fixed by comparing the lattice NRQCD dispersion relation to that of the continuum in perturbation theory [118]. c_2 is a spin-independent term which can affect radial and orbital excitation energies, this is not expected to have as large an effect as the kinetic terms, so is set to its tree-level value of 1. The result of varying c_2 on relevant observables was investigated in Sec. IIIC of [119], and the effects were very small. c_3 and c_4 are spin-dependent terms, which would have a small effect on spin-averaged observables (i.e. all observables computed in this work). c_3 is set to 1, and c_4 is tuned non-perturbatively, by matching predictions of the fine structure of the Υ spectrum from lattice NRQCD to experiment [119].

Another Symanzik improvement is introduced in this context by multiplying the gauge links by the so-called *tadpole factor* $u_0 = \sum_{\mu,\nu} \langle \text{Tr} \square_{\mu\nu} / 4 \rangle$. This removes the tadpole diagrams proportional to a^2 that appear in gluon propagators.

There is a limit to how small a lattice spacing can be used while employing the Lattice NRQCD action. Lattice NRQCD relies on convergent series in inverse b -mass, in lattice units this becomes $1/(am_b)$. It is therefore necessary for am_b to be suitably larger than one, or equivalently $a > 1/m_b$.

Lattice Calculations

The previous chapter focused on how to discretize the QCD action. This chapter is focused on the practical side of lattice QCD - given a lattice action, how does one perform the functional integral to determine expectation values?

4.1 Evaluation of Lattice Correlation Functions

All physics of a quantum field theory can be extracted from correlation functions. So a typical lattice calculation involves computing a correlation function (or just *correlator*) on the lattice, then extracting physical quantities from it. A typical correlator that is computed on the lattice is a 2-point meson correlator, i.e. $\langle \Phi(x)\Phi^\dagger(y) \rangle$ where Φ is a meson creation operator and Φ^\dagger is an annihilation operator. This is a good working example for showing the steps in a lattice calculation, the generalization to N -point correlators is reasonably natural.

A creation/annihilation operator for a meson in this context can be any operator containing the same quantum numbers as the meson one is studying. For example, the neutral B meson is a pseudoscalar charged with a b and \bar{d} quark, so a suitable operator is $\Phi(x) = \bar{b}(x)\gamma_5 d(x)$. The corresponding functional can then be written as

$$C(x, y) = \langle \Phi(x)\Phi^\dagger(y) \rangle = \int [d\psi d\bar{\psi} dU] (\bar{b}(x)\gamma_5 d(x)\bar{d}(y)\gamma_5 b(y)) \times \exp\left(-S_G[U] - \sum_{w,z,i} \bar{q}_i(w)M_{q_i}(w,z)[U]q_i(z)\right), \quad (4.1)$$

where we have broken the action up into a gauge part $S_G[U]$, and a fermion part. $M_{q_i}(x, y)[U]$ is the Dirac operator for flavour i , and can be seen as a matrix in lattice site, color and spin.

The integral over fermions can be performed analytically since the fermion fields

are Grassmann valued. In our example, the result is [96]:

$$C(x, y) = \int [dU] \text{Tr} (M_b^{-1}(y, x)[U] \gamma_5 M_d^{-1}(x, y)[U] \gamma_5) \\ \times e^{-S_G[U]} \prod_i \det(M_{q_i}[U]). \quad (4.2)$$

Performing this integral is equivalent to a Wick contraction between the quark sources and sinks. In the case of two degenerate flavours (e.g. for a neutral pion, an $\bar{u}u$ pseudoscalar) a second term would be present. This so-called ‘disconnected’ term would connect the source to itself and the sink to itself rather than connecting the source to sink, i.e. the trace would be replaced with $\text{Tr}(M_u^{-1}(x, x)[U] \gamma_5) \text{Tr}(M_u^{-1}(y, y)[U] \gamma_5)$. In this thesis we will not be concerned with disconnected contributions to correlators.

The quantities $M_{q_i}^{-1}(x, y)[U]$ are propagators of a quark of flavour q on a fixed gauge background U . For clarity: here U denotes a configuration of angles comprising an $SU(3)$ matrix for each element of the set of all links on the lattice $\{U_\mu(x) | \forall \mu, x\}$. The trace is over color and spin. The integration over gauge fields is generally carried out by an importance sampling method. A finite *ensemble* of gauge configurations $\{U_n\}$ is generated by a Monte Carlo Markov Chain (MCMC), where the probability of a gauge configuration U_n being added to the ensemble is proportional to

$$p(U_n) = e^{-S_G[U_n]} \prod_i \det(M_{q_i}[U_n]). \quad (4.3)$$

Once the ensemble is created, the path integral can be approximated by simply

$$C(x, y) \simeq \frac{1}{N} \sum_{n=1}^N \text{Tr} [M_b^{-1}(y, x)[U_n] \gamma_5 M_d^{-1}(x, y)[U_n] \gamma_5], \quad (4.4)$$

where N is the number of configurations in the ensemble. This introduces a statistical error that scales like $1/\sqrt{N}$. The calculation of the correlation function then is split into 3 steps:

1. Generate an ensemble of gauge configurations $\{U_i\}$ by MCMC (Sec. 4.1.1).
2. Compute $M_{q_i}^{-1}(x, y)[U]$ by inverting the Dirac operator on each gauge configuration (Sec. 4.1.2).
3. Construct the trace in Eq. (4.4), and average over the ensemble. This step is dealt with in the context of staggered quarks in Sec. 4.1.3.

4.1.1 Generation of Gauge Ensembles

The calculation requires a number of samples of gauge configurations $\{U_n\}$ sampled from the distribution $p(U)$ defined in Eq. (4.3). In this section (Sec. 4.1.1) I largely follow the discussion given in [97].

The physical interpretation of the determinant in (4.3) is that it accounts for virtual quark loops in gluon propagators. In the early days of lattice calculations, this determinant was approximated to 1, since its evaluation was an insurmountable computational cost, and it was expected that sea quarks had small effects on observables (this is known as the *quenched approximation*). However, this introduced large systematic effects that could not be well controlled. These days, our computational ability has improved and sophisticated approaches to computing the determinant have been developed (e.g. [120]), so we can include it in our calculations.

We will roughly follow the history of gauge ensemble generation, by first ignoring the determinant, and then showing how it is eventually included in the process.

Quenched MCMC

Gauge ensembles are generated via an MCMC, inspired by statistical mechanics. The distribution $\exp(-S_G[U])$ is suggestive of something like a Boltzmann distribution for a gas of particles, each with some state U_i , in thermal equilibrium. The ergodic hypothesis states that a single particle in this gas will jump between possible states over time such that, at any given time, its probability of being in state U_i is given by $\exp(-S_G[U_i])$. In MCMC, one starts with some random state U_0 , then repeatedly updates the state according to some update rule or ‘hopping rate’ $p(U_i \rightarrow U_j)$.

The hopping rate must be designed to bring the chain into thermal equilibrium with the correct distribution. A sufficient condition for thermal equilibrium is known as *detailed balance*, where the probability of jumps between any pair of states i and j is equal:

$$p(U_i)p(U_i \rightarrow U_j) = p(U_j)p(U_j \rightarrow U_i). \quad (4.5)$$

Hence $p(U_i \rightarrow U_j)$ must be designed according to the rule

$$\frac{p(U_i \rightarrow U_j)}{p(U_j \rightarrow U_i)} = \exp(-(S_G[U_i] - S_G[U_j])). \quad (4.6)$$

There are a number of possible choices of how to design $p(U_i \rightarrow U_j)$. One approach, called **molecular dynamics** [121,122] is to model the chain as the trajectory $U(\tau)$

of a system with Hamiltonian

$$H(\pi, U) = \frac{\pi^2}{2} + S_G[U], \quad (4.7)$$

where π is a fictitious momentum conjugate to U . It can be demonstrated that such a trajectory obeys (4.6) [121]. The trajectory is computed via Runge-Kutta numerical integration. One may worry about the possibility of fixed points, limit cycles etc. in the dynamics, which would prevent ergodicity. To avoid this one can introduce a periodic **refreshing** step, where π assigned a new value from normally distributed noise [123, 124].

Another problem that can occur in molecular dynamics is when errors in Runge-Kutta iterations accumulate over time. Diversion from the dynamics enforced by $H(\pi, U)$ can ruin the ergodicity of the trajectory. To fix this, one can add a **Metropolis** step at regular intervals $\delta\tau$ throughout the evolution [125]. In this step, one either accepts (continues onto the next stage of molecular dynamics) or rejects (refreshes π and re-calculates the $\delta\tau$ worth of molecular dynamics), according to the criterion

- If $S_G[U(\tau + \delta\tau)] < S_G[U(\tau)]$, always accept.
- Otherwise, accept if $\exp(S_G[U(\tau + \delta\tau)] - S_G[U(\tau)]) > \lambda$, where λ is randomly chosen from the interval $[0, 1]$.

The metropolis step ensures detailed balance (Eq. (4.6)) is satisfied even in the presence of Runge-Kutta errors.

The combination of molecular dynamics, refreshing steps and Metropolis steps is referred to as **Hybrid Monte Carlo** [126], and is the basic method of how the ensembles we use in this thesis were generated. I now address how the determinant $\det M$ is included.

Unquenched MCMC

Simply evaluating $\det M[U]$ directly, given a configuration U , is prohibitively expensive due to the non-local nature of the determinant. Recall $M[U]$ is a matrix in spin, colour, and lattice site, in modern calculations this will have a dimension of order 10^8 . Even holding that much information in memory is not feasible. A solution to this is to use the Φ -algorithm [120].

First, we replace $\det M$ with $\det M^\dagger M$. If we were only including u and d quarks in the sea, this would be fine since we can approximate u and d to be two degenerate

flavours, then $\prod_q \det M_q = \det M \det M = \det M^\dagger M$. In the case of an arbitrary set of flavours, this requires a correction that will be addressed later. The Φ -algorithm involves introducing new artificial scalar fields $\Phi(x)$ and $\Phi^\dagger(x)$ via

$$\det M^\dagger M = \int [d\Phi^\dagger d\Phi] \exp(-\Phi^\dagger (M^\dagger M)^{-1} \Phi). \quad (4.8)$$

then one can add $\Phi^\dagger (M^\dagger M)^{-1} \Phi$ to S_G in the Hybrid Monte Carlo algorithm. The extra functional integral over Φ, Φ^\dagger is easily evaluated, by sampling a vector η from a normal distribution $\exp(-\eta^\dagger \eta)$, then transforming it to $\Phi = M^\dagger \eta$.

The Rooting Trick

We will now address how to correct for the fact that we have replaced $\det M$ with $\det M^\dagger M$ in the presence of arbitrary non-degenerate flavours. We have explicitly doubled the fermions to two degenerate flavours per physical flavour. In the case of staggered quarks, this is not a huge marginal complication since we already have four degenerate tastes which we have to deal with anyway. In order to cut down the number of tastes in the sea, the solution is to take the fourth-root of $\det M$. When using the Φ -algorithm, this becomes the 8th root of $\det M^\dagger M$.

$$\begin{aligned} (\det M^\dagger M)^{1/8} &= \left(\prod_i \lambda_i^2 \right)^{1/8} = \left(\prod_i \lambda_i \right)^{1/4} \\ &\stackrel{?}{=} \left(\prod_i \lambda_i'^4 \right)^{1/4} = \prod_i \lambda_i' \quad (a \rightarrow 0). \end{aligned} \quad (4.9)$$

where λ_i are eigenvalues of M . On the second line, we have assumed that the matrix M can be decomposed into four matrices, one for each of the four tastes, with eigenvalues λ_i' which are degenerate in the continuum limit.

This assumption is not rigorously justified in field theory, so the fourth-root trick is a source of controversy. Much has been said about the problems this may cause in lattice results [127–129], however, these concerns have been refuted [130, 131]. It has been demonstrated that the eigenvalues smoothly become degenerate as one approaches the continuum limit [132, 133]. There has so far emerged no evidence that the rooting trick is harmful, observables computed using unquenched staggered quarks have always agreed with experiment, analytical approaches (e.g. [134]), and other lattice discretisations.

Introducing the 1/2 or 1/8th root to the determinant requires a modification of the Φ -algorithm, we can no longer simply sample Φ using $\Phi = M^\dagger \eta$. The effective

action is now $S_G + \Phi^\dagger (M^\dagger M)^{-1/8} \Phi$. The root is dealt with by replacing it with a partial fraction representation [135]:

$$(M^\dagger M)^{-1/8} \simeq a_0 + \sum_{n=1}^N \frac{a_n}{M^\dagger M + b_n}. \quad (4.10)$$

This can only be evaluated by some variation of a conjugate gradient algorithm (specifically a multishift solver [136, 137]). Conjugate gradient will be described in Sec. 4.1.2. This approach is called the **Rational Hybrid Monte Carlo** (RHMC) algorithm.

The $N_f = 2 + 1 + 1$ MILC Ensembles

In this work, we use ensembles of gauge configurations generated by the MILC collaboration [7, 8]. The ingredients of these configurations are

- Gauge fields obeying the one-loop Symanzik improved Lüscher-Weisz action described in Sec. 3.1.2.
- Four flavours of quark in the sea, u, d, s and c (with $m_u = m_d \equiv m_l$), hence the notation $N_f = 2 + 1 + 1$, obeying the HISQ action, described in Sec. 3.2.3.
- Ensemble generated (mostly) using the RHMC algorithm as described earlier in this section. Some configurations on set 3 were instead generated using an RHMD algorithm - similar to RHMC except with the Metropolis accept/reject step omitted.

Table 4.1 gives the details of the MILC ensembles that were used in this work. One may notice that for the majority of ensembles here, the light quarks are much heavier than in reality. The necessity for this is explained in the next section.

4.1.2 Dirac Operator Inversion

Once the ensemble $\{U_i\}$ has been generated, to compute the 2-point correlator (4.4) one must compute $M^{-1}[U_i]$ for each U_i . We have already seen how this can be done in the case of the flavour in question being governed by the NRQCD action, one can use the recursion relation (3.57). In the case of relativistic actions like HISQ, there is no equivalent recursion relation.

M is large but sparse. It technically has $\mathcal{O}(\text{Vol}^2)$ elements, but for suitably local actions (like HISQ) it has only $\mathcal{O}(\text{Vol})$ non-zero elements. This means it is

set	name	w_0/a	$N_x^3 \times N_t$	am_{l0}	am_{s0}	am_{c0}
0	very coarse	1.1119(10)	$16^3 \times 48$	0.013	0.067	0.838
1	coarse	1.3826(11)	$24^3 \times 64$	0.0102	0.0509	0.635
2	fine	1.9006(20)	$32^3 \times 96$	0.0074	0.037	0.440
3	fine-physical	1.9518(7)	$64^3 \times 96$	0.0012	0.0363	0.432
4	superfine	2.896(6)	$48^3 \times 144$	0.0048	0.024	0.286
5	ultrafine	3.892(12)	$64^3 \times 192$	0.00316	0.0158	0.188

Table 4.1: Parameters for the MILC gluon ensembles [7,8]. a is the lattice spacing, determined from the Wilson flow parameter w_0 . Values for w_0/a are from: sets 0,1,2 [138], sets 3 and 4 [139], set 5 [140]. The physical value of w_0 was determined to be $w_0 = 0.1715(9)\text{fm}$ in [141]. Columns 5-7 give the masses used in the action for light, strange and charm quarks in the sea.

well-suited to the **conjugate gradient** (CG) algorithm [142] (and its variants), which has become the most successful approach to computing M^{-1} . However, CG requires the matrix being inverted to be hermitian and positive definite, which is not necessarily the case for M . We instead invert $M^\dagger M$, which *is* hermitian and positive definite, then we can recover M^{-1} by acting M^\dagger on $(M^\dagger M)^{-1}$.

The design of CG requires a lot of explanation that I will not go into here. I will instead briefly describe the philosophy behind it, and state the algorithm. For a nice review with lots of detail see [143]. The goal is, given some vector b and matrix A , to find x where

$$Ax = b. \quad (4.11)$$

In our case $A = M^\dagger M$ and b is a suitably chosen 'source' for the propagator (see Sec. 4.1.3). This is equivalent to finding the $x = x^*$ that minimizes

$$f(x) = \frac{1}{2}x^T Ax - b^T x. \quad (4.12)$$

A reasonable solution to this problem is something like a *steepest descent* approach, where one starts at a random x_0 , then moves some distance α_0 in the direction $r_0 = -f'(x_0) = b - Ax_0$ to $x_1 = x_0 + \alpha_0 r_0$. α_0 is chosen to minimize $x^* - x_1$. And then repeat. This approach has the property that each new step $\alpha_n r_n$ is orthogonal

to every other step, this means the algorithm takes a sub-optimal zig-zag path towards the solution.

CG is designed to take a more direct path, by imposing the condition that the direction of each step $d_n = (x_n - x_{n-1})/\alpha_n$ is orthogonal with respect to the metric A , i.e. $d_n^T A d_m = 0$ for $n \neq m$. The CG algorithm is

$$\begin{aligned}
 x_{n+1} &= x_n + \alpha_n d_n, \text{ where} \\
 \alpha_n &= \frac{r_n^T r_n}{d_n^T A d_n}, \\
 d_n &= \begin{cases} r_0, & n = 0 \\ r_n + \beta_n d_{n-1}, & n > 0, \end{cases} \\
 r_n &= b - A x_n, \\
 \beta_n &= \frac{r_n^T r_n}{r_{n-1}^T r_{n-1}}. \tag{4.13}
 \end{aligned}$$

One terminates the algorithm when some stopping condition is achieved, namely when $r_n < \epsilon$ where ϵ is some small number referred to as the error tolerance, or when some maximum number of iterations has been reached.

The complexity of the CG algorithm is $\mathcal{O}(c)$ where $c = \lambda_{\max}/\lambda_{\min}$ is the condition number of the matrix A . $\lambda_{\max/\min}$ are the largest and smallest eigenvalues of A . The condition number quantifies the size of rounding errors that accumulate in iterative processes like CG. In our case where $A = M^\dagger M \sim (-i\mathcal{D} + m)(i\mathcal{D} + m)$, the condition number is proportional to m^{-2} . Hence, propagators for lighter quarks are quadratically more expensive to compute than heavier ones. This affects the computation of correlation functions including light valence quarks via M_l^{-1} . It also affects any unquenched calculation with rooting since in that case we must also perform an inversion to evaluate (4.10).

For this reason, lattice calculations are often computed with unphysically heavy u/d quarks. Modern lattice calculations have computed observables for a number of light quark masses and extrapolated downwards to the physical light mass, using chiral perturbation theory as a guide. In the MILC ensembles we use in this work, summarized in Table 4.1, all but one have a light mass at around $m_l/m_s \simeq 1/5$, while set 3 (fine-physical) has roughly physical light quarks at $m_l/m_s \simeq 1/30$.

4.1.3 Staggered Correlation Functions

We now turn to how to evaluate traces of quark propagators, as in Eq. (4.2), in the staggered formalism.

Recall from Sec. 3.2.2, propagators for naive quarks M^{-1} are related to staggered propagators g by

$$M^{-1}(x, y) = \Omega(x)\Omega^\dagger(y)g(x, y). \quad (4.14)$$

Throughout this section we will keep the gauge field dependence of M^{-1} and g implicit. By conjugating both sides and using the property of the naive propagator $(M^{-1})^\dagger(x, y) = \gamma_5 M^{-1}(y, x)\gamma_5$ one can show that M^{-1} can also be written as

$$M^{-1}(y, x) = \phi_5(y)\phi_5(x)\Omega(y)\Omega^\dagger(x)g^\dagger(x, y), \quad (4.15)$$

where $\phi_5(x) = (-1)^{\sum_\mu x_\mu/a}$.

In this section we will only treat ‘connected’ correlators (i.e. ignoring disconnected contributions in the case of degenerate flavors mentioned in Sec. 4.1). The generalization to disconnected contributions from the below discussion is straightforward.

2-point Correlation Functions

Consider the generic 2-point correlator; involving two valence flavours a and b , and spin structure γ_X and γ_Y at the source and sink:

$$\begin{aligned} C(x, y) &= \langle \Phi_X^\dagger(x)\Phi_Y(y) \rangle_{\psi, U} \quad , \quad \Phi_X(x) = \frac{1}{4}\bar{\psi}_a(x)\gamma_X\psi_b(x) \\ &= \frac{1}{16}\langle \text{Tr}_{c,s}\gamma_X M_a^{-1}(x, y)\gamma_Y M_b^{-1}(y, x) \rangle_U \\ &= \frac{1}{16}\phi_5(x)\phi_5(y)\text{Tr}_s \left(\Omega^\dagger(x)\gamma_X\Omega(x)\Omega^\dagger(y)\gamma_Y\Omega(y) \right) \langle \text{Tr}_c \left(g_a(x, y)g_b^\dagger(x, y) \right) \rangle_U. \end{aligned} \quad (4.16)$$

Tr_s is a trace over spin and Tr_c is over color. We have applied Eq. (4.15) to the b propagator in the last line. To deal with the spin trace, define the family of phases $\{\phi_X(x)\}$ according to

$$\Omega^\dagger(x)\gamma_X\Omega(x) = \phi_X(x)\gamma_X. \quad (4.17)$$

For example, if $X = 5$, then $\gamma_x^\dagger\gamma_5\gamma_x = (-1)^{\sum_\mu x_\mu}\gamma_x^\dagger\gamma_x\gamma_5 = \phi_5(x)\gamma_5$. The map from X to ϕ_X is structure preserving, i.e. if $\gamma_X = \gamma_A\gamma_B$, then $\phi_X(x) = \phi_A(x)\phi_B(x)$. The

spin trace becomes $\phi_X(x)\phi_Y(y)\text{Tr}_s(\gamma_X\gamma_Y)$. The remaining trace will vanish unless $Y = X$, and is 4 otherwise. We end up with

$$C(x, y) = \frac{1}{4}\phi_{5X}(x)\phi_{5Y}(y)\langle \text{Tr}_c g_a(x, y)g_b^\dagger(x, y) \rangle_U. \quad (4.18)$$

We are usually interested instead in the correlation function of a meson in a momentum eigenstate with spatial momentum \mathbf{p} . This will take the form

$$\begin{aligned} C_{\mathbf{p}}(t_0, t) &= \frac{1}{N_x^3} \sum_{\mathbf{x}, \mathbf{y}} e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})} C(\mathbf{x}, t_0; \mathbf{y}, t) \\ &= \frac{1}{4N_x^3} \sum_{\mathbf{x}, \mathbf{y}} e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})} \phi_{5X}(x)\phi_{5Y}(y)\langle \text{Tr}_c g_a(x, y)g_b^\dagger(x, y) \rangle_U, \end{aligned} \quad (4.19)$$

where it is understood that $x_0 = t_0$ and $y_0 = t$. In order to evaluate this function, we must perform inversions to create $g_{a/b}(x, y)$ for each x and y , so $2\cdot\text{Vol}^2$ calculations. This is prohibitively expensive. The number of inversions can be reduced by using *random wall sources*. Define

$$P_{a,\mathbf{p},X}^{t_0}(y) \equiv \frac{1}{\sqrt{N_x^3}} \sum_{\mathbf{x}} e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})} \phi_{5X}(\mathbf{x}, t_0)\xi(\mathbf{x})g_a(\mathbf{x}, t_0; y), \quad (4.20)$$

where $\xi(\mathbf{x})$ is a random field of colour vectors, a different field for each gauge configuration. This has the property

$$\langle f(\mathbf{x}, \mathbf{x}')\xi^*(\mathbf{x}')\xi(\mathbf{x}) \rangle_U = \delta_{\mathbf{x}, \mathbf{x}'} \langle f(\mathbf{x}, \mathbf{x}') \rangle_U. \quad (4.21)$$

Using this property the correlator can be built instead according to

$$C_{\mathbf{p}}(t_0, t) = \frac{1}{4} \sum_{\mathbf{y}} \phi_{5Y}(y)\langle \text{Tr}_c P_{a,\mathbf{p},X}^{t_0}(\mathbf{y}, t)P_{b,0,5}^{t_0\dagger}(\mathbf{y}, t) \rangle_U. \quad (4.22)$$

Now all one has to compute is $P_{a/b}^{t_0}(y)$ for general y , so $2\cdot(\text{Vol})$ calculations, a reduction by a factor of (Vol) .

3-point Correlation Functions

The above discussion can be generalized to 3-(or N -)point correlation functions. Consider a 3-point correlation function, for example encoding an $X \rightarrow Z$ semileptonic decay via a current $J(y)$:

$$\begin{aligned} C(x, y, z) &= \langle \Phi_X^\dagger(x)J(y)\Phi_Z(z) \rangle_{\psi,U}, & \Phi_X(x) &= \frac{1}{4}\bar{\psi}_b(x)\gamma_X\psi_s(x) \\ & & J(y) &= \bar{\psi}_b(y)\gamma_J\psi_a(y) \\ & & \Phi_Z(z) &= \frac{1}{4}\bar{\psi}_a(z)\gamma_Z\psi_s(z). \end{aligned} \quad (4.23)$$

We can reduce this in the same way as before

$$C(x, y, z) = \frac{1}{16} \text{Tr}_s \left(\Omega^\dagger(x) \gamma_X \Omega(x) \Omega^\dagger(y) \gamma_J \Omega(y) \Omega^\dagger(z) \gamma_Z \Omega(z) \right) \quad (4.24)$$

$$\begin{aligned} & \times \phi_5(x) \phi_5(z) \langle \text{Tr}_c g_b(x, y) g_a(y, z) g_s^\dagger(x, z) \rangle_U \\ & = \frac{1}{4} \phi_{5X}(x) \phi_J(y) \phi_{5Z}(z) \langle \text{Tr}_c g_b(x, y) g_a(y, z) g_s^\dagger(x, z) \rangle_U. \end{aligned} \quad (4.25)$$

We have assumed that $\text{Tr}_s \gamma_X \gamma_J \gamma_Z = 4$, requiring that each gamma matrix in this combination has a partner and therefore cancels.

Putting the X -meson into an eigenstate of zero momentum, and the Y -meson into an eigenstate of momentum \mathbf{p} , we get

$$\begin{aligned} C_{\mathbf{p}}(t_0, t, T) &= \frac{1}{4N_x^3} \sum_{\mathbf{x}, \mathbf{y}, \mathbf{z}} e^{i\mathbf{p} \cdot (\mathbf{y} - \mathbf{z})} \phi_{5X}(x) \phi_J(y) \phi_{5Z}(z) \\ & \times \langle \text{Tr}_c g_b(\mathbf{x}, t_0; \mathbf{y}, t) g_a(\mathbf{y}, t; \mathbf{z}, T) g_s^\dagger(\mathbf{x}, t_0; \mathbf{z}, T) \rangle_U. \end{aligned} \quad (4.26)$$

This can be built by first creating propagators for the b and s quarks - $P_{b,0,X}^{t_0}(y), P_{s,0,1}^{t_0}(z)$. Then, build the a propagator using a so-called *extended source*:

$$P_{a,\mathbf{p},ext}^T(y) = \sum_{\mathbf{z}} P_{s,0,5}^{t_0 \dagger}(\mathbf{z}, T) \phi_Z(\mathbf{z}, T) e^{i\mathbf{p} \cdot (\mathbf{y} - \mathbf{z})} g_a(y; \mathbf{z}, T). \quad (4.27)$$

We can build the 3-point correlator (4.26) using essentially the same 'tie together' as (4.22):

$$C_{\mathbf{p}}(t_0, t, T) = \frac{1}{4} \sum_{\mathbf{y}} \phi_J(y) \langle \text{Tr}_c P_{b,0,5X}^{t_0}(\mathbf{y}, t) P_{a,\mathbf{p},ext}^T(\mathbf{y}, t) \rangle_U. \quad (4.28)$$

I'll briefly connect the above discussion to the spin-taste notation introduced in Sec. 3.2.2. In the above, we have not used any point-split operators. Hence, we denote these operators in spin-taste notation as $(\gamma_n \otimes \gamma_n)$, where γ_n is the continuum spin structure we are aiming for. In the work in this thesis, we will not use any point-split operators, so the above discussion is sufficient for understanding the methods used.

Momentum Twist

The way in which spatial momentum is introduced into the correlation functions requires some explanation. The momentum space 2-point correlation function for an operator \mathcal{O} with momentum \mathbf{p} is given by

$$C_{\mathbf{p}}(0, t) = \sum_{\mathbf{x}} e^{i\mathbf{p} \cdot \mathbf{x}} \langle \mathcal{O}^\dagger(\mathbf{x}, t) \mathcal{O}(\mathbf{0}, 0) \rangle. \quad (4.29)$$

To introduce \mathbf{p} one can add an appropriate phase to the operators:

$$\mathcal{O}(\mathbf{x}, t) \rightarrow \mathcal{O}(\mathbf{x}, t)e^{-i\mathbf{p}\cdot\mathbf{x}} \quad (4.30)$$

$$\implies C(\mathbf{0}, t) \rightarrow C(\mathbf{p}, t). \quad (4.31)$$

This generalizes straightforwardly to n -point functions. One can assign the rephasing to any factor in \mathcal{O} , for example a fermion operator

$$\psi(\mathbf{x}, t) \rightarrow \psi(\mathbf{x}, t)e^{-i\mathbf{p}\cdot\mathbf{x}}. \quad (4.32)$$

Rephasing ψ is equivalent to introducing a *momentum twist* to the gauge links [144]. The action of Eq. (4.32) on any gauge invariant quantity is equivalent to

$$U_i \rightarrow U_i e^{iap_i} \quad (\text{no sum}). \quad (4.33)$$

For example, consider the effect this has on the following operator

$$\begin{aligned} & \psi^\dagger(x)U_\mu(x)\psi(x+a\hat{\mu}) \\ \rightarrow & \psi^\dagger(x)(e^{iap_\mu}U_\mu(x))\psi(x+a\hat{\mu}) \\ = & \psi^\dagger(x)e^{-i\mathbf{p}\cdot\mathbf{x}}U_\mu(x)e^{+i\mathbf{p}\cdot(\mathbf{x}+a\hat{\mu})}\psi(x+a\hat{\mu}). \end{aligned} \quad (4.34)$$

When computing a propagator $g_a(x, y)$, we add these phases to the gauge fields which have the effect of the flavour a carrying the spatial momentum. This is how momentum is included in the work of this thesis. We report momentum twist in units of π/N_x , e.g., a twist of θ in the k direction corresponds to a momentum of $ap_k = \pi\theta/N_k$.

4.2 Analysis of Correlation Functions

Once correlation functions like $C_{\mathbf{p}}(t_0, t)$ and $C_{\mathbf{p}}(t_0, t, T)$ have been computed on the lattice, how can we extract physics from them?

4.2.1 Fitting Correlation Functions

2-point correlators contain information about (amongst other things) masses and decay constants of the propagating meson. One performs a χ^2 -fit of the correlator to a theoretically motivated function of t . To derive such a function, we use a complete set of momentum \mathbf{p} states -

$$1 = \sum_{n=0}^{\infty} \frac{1}{2E_n} |\lambda_n\rangle\langle\lambda_n|, \quad (4.35)$$

where E_n are the energies of each state. Inserting this into the correlation function, and moving from the Heisenberg to Schrödinger picture [145]:

$$\begin{aligned}
C_{\mathbf{p}}(t) &= \langle \Omega | \Phi(\mathbf{p}, t) \Phi^\dagger(\mathbf{p}, 0) | \Omega \rangle \\
&= \sum_{n=0}^{N_{\text{exp}}} \frac{1}{2E_n} \langle \Omega | \left(e^{Ht} + e^{H(T_{\text{lat}}-t)} \right) \Phi(\mathbf{p}, 0) \left(e^{-Ht} + e^{-H(T_{\text{lat}}-t)} \right) | \lambda_n \rangle \\
&\quad \times \langle \lambda_n | \Phi^\dagger(\mathbf{p}, 0) | \Omega \rangle \\
&= \sum_{n=0}^{N_{\text{exp}}} \left(\frac{\langle \Omega | \Phi(\mathbf{p}, 0) | \lambda_n \rangle}{\sqrt{2E_n}} \right) \left(\frac{\langle \lambda_n | \Phi^\dagger(\mathbf{p}, 0) | \Omega \rangle}{\sqrt{2E_n}} \right) \left(e^{-\bar{E}_n t} + e^{-\bar{E}_n (T_{\text{lat}}-t)} \right) \\
&\equiv \sum_{n=0}^{N_{\text{exp}}} |a_n|^2 f(\bar{E}_n, t) \quad , \quad f(E, t) = \left(e^{-Et} + e^{-E(T_{\text{lat}}-t)} \right) , \tag{4.36}
\end{aligned}$$

where $T_{\text{lat}} = aN_t$ is the temporal extent of the lattice. I have here set $t_0 = 0$ for clarity. In practice, one would only use ‘late time’ data, $t \geq t_{\text{cut}}$ for some t_{cut} . In the late time data the correlator is dominated by the lowest- n terms, since higher n terms are suppressed by faster decaying exponentials $\exp(-\bar{E}_n t)$. Hence we can afford to truncate the sum at some finite number of terms, N_{exp} .

The fit results in a determination of the parameters a_n and \bar{E}_n . The sum over n will be populated only by states $|\lambda_n\rangle$ with the same quantum numbers as Φ , since $\langle \Omega | \Phi | \lambda_n \rangle$ vanishes in all other cases. We can then interpret $|\lambda_0\rangle$ to be the ground state of the meson we are studying.

We are maintaining a distinction between \bar{E}_n and E_n here, since these differ in calculations involving NRQCD quarks. In NRQCD \bar{E}_n is the non-relativistic energy with leading- v behaviour $\mathbf{p}^2/2m$.

In the HISQ case, one can safely interpret these as relativistic energies, $\bar{E}_n = E_n$. One can find the meson’s mass by computing the correlation function at zero momentum $C_{\mathbf{0}}(t)$, the fit parameter \bar{E}_0 will equal the mass M . a_n can be related to the meson’s decay constant. For example for a pseudoscalar meson, using the definition of a meson decay constant (2.21) and the PCAC relation in (2.51), we find

$$f_M = (m_a - m_b) \sqrt{\frac{2}{M^3}} \left(\frac{a_0}{a^{3/2}} \right) , \tag{4.37}$$

where a and b are the two flavours the meson is charged under.

The above discussion can be straightforwardly generalized to 3-point correlation functions, from which we are able to extract quantities like the hadronic transition

amplitudes $\langle M'|J|M\rangle$, from which we can determine semileptonic form factors. The generalization of the above for 3-point correlators is

$$\begin{aligned}
C(0, t, T) &= \langle \Omega | \Phi_{M'}(0) J(t) \Phi_M^\dagger(T) | \Omega \rangle \\
&= \sum_{n,m} \left(\frac{\langle \Omega | \Phi_{M'} | \lambda_n \rangle}{\sqrt{2E_{M',m}}} \right) \left(\frac{\langle \lambda_n | J | \lambda_m \rangle}{2\sqrt{E_{M,n}E_{M',m}}} \right) \left(\frac{\langle \lambda_m | \Phi_M^\dagger | \Omega \rangle}{\sqrt{2E_{M,n}}} \right) \\
&\quad \times f(\bar{E}_{M',m}, T-t) f(\bar{E}_{M,n}, t) \\
&\equiv \sum_{n,m} a_{M',n} J_{nm} a_{M,m}^* f(\bar{E}_{M',m}, T-t) f(\bar{E}_{M,n}, t). \tag{4.38}
\end{aligned}$$

I have suppressed spatial momentum dependence here for notational simplicity. $a_{M,n}$ will vanish for states $|\lambda_n\rangle$ that have different quantum numbers to Φ_M , similarly for $a_{M',m}$ and $\Phi_{M'}$. Non-zero $a_{M,n}$'s will match the analogous parameters extracted from fitting a 2-point function $\langle \Phi_M^\dagger \Phi_M \rangle$, similarly for $a_{M',m}$'s and $\Phi_{M'}$. This carries on to the energies; $\{\bar{E}_{M,n}\}$ is the spectrum for the M' meson, and $\{\bar{E}_{M',m}\}$ is the spectrum for the M . Therefore, we compute and fit the appropriate 2-point functions to deduce the parameters $\{a_{M^{(\prime)},n}\}, \{\bar{E}_{M^{(\prime)},n}\}$, then fitting $C(0, t, T)$ results in an accurate determination of the remaining free parameters, J_{nm} . This set contains the transition amplitude one is interested in $\langle M'|J|M\rangle$, recognising that

$$J_{00} = \frac{\langle M'|J|M\rangle}{2\sqrt{E_{M,0}E_{M',0}}}. \tag{4.39}$$

Oscillating States

In the case of staggered quarks, these fit functions must be modified to contain the effects of the *oscillating states*. The oscillating states are due to propagation of mesons in the correlator containing the $\zeta = (1, 0, 0, 0)$ taste of one of the valence quarks (in the language of Sec. 3.2.1). No other tastes contribute, since $\Phi(\mathbf{p}, t)$ has a 3-momentum fixed at \mathbf{p} , which we always take to be small relative to π/a . Hence $\Phi(\mathbf{p}, t)$ does not couple to the states at $k \sim (0, \pi/a, 0, 0)$, $k \sim (0, 0, \pi/a, 0)$ etc. However, $\Phi(\mathbf{p}, t)$ can couple to arbitrarily high energy states, so the pole at $k \sim (\pi/a, 0, 0, 0)$ contributes.

How this taste contribution can be seen using the doubling symmetry. I will use a B meson as an example. We can translate the $\zeta = (1, 0, 0, 0)$ pole in momentum space down to around \mathbf{p} by the transform $\psi_l \rightarrow (i\gamma_5\gamma_0)(-1)^{t/a}\psi_l$ on the offending flavour, say it is the light quark l . This causes $\Phi_B(\mathbf{p}, t)$ to become

$$\Phi_B(\mathbf{p}, t) = \bar{\psi}_b \gamma_5 \psi_l \rightarrow i(-1)^{t/a} \bar{\psi}_b \gamma_0 \psi_l. \tag{4.40}$$

The $\zeta = (1, 0, 0, 0)$ taste of the l -quark manifests itself as a scalar (0^+) meson with an oscillating phase $(-1)^{t/a}$.

Accounting for oscillating states modifies the fit functions to

$$C_{\mathbf{p}}(t)|_{\text{fit}} = \sum_{j=0}^{N_{\text{exp}}} |a_j|^2 f(\bar{E}_j, t) + (-1)^{t/a} |a_{j,o}|^2 f(\bar{E}_{j,o}, t) \quad (4.41)$$

$$\begin{aligned} C_{\mathbf{p}}(t, T)|_{\text{fit}} = & \sum_{j,k=0}^{N_{\text{exp}}, N_{\text{exp}}} \left(a_j^M J_{jk}^{nn} a_k^{M'} f(\bar{E}^M, t) f(\bar{E}_n^{M'}, T-t) \right. \\ & + a_j^{M,o} J_{jk}^{on} a_k^{M'} (-1)^{t/a} f(\bar{E}_n^{M,o}, t) f(\bar{E}^{M'}, T-t) \\ & + a_j^M J_{jk}^{no} a_k^{M',o} (-1)^{(T-t)/a} f(\bar{E}^M, t) f(\bar{E}_n^{M',o}, T-t) \\ & \left. + a_j^{M,o} J_{jk}^{oo} a_k^{M',o} (-1)^{T/a} f(\bar{E}_n^{M,o}, t) f(\bar{E}^{M',o}, T-t) \right). \end{aligned} \quad (4.42)$$

There is a special case where oscillating states do not contribute. If the two quarks in $\Phi(\mathbf{p}, t)$ are degenerate (have the same flavour, momentum etc.) then the doubling symmetry acts on both of the quark fields identically. If the meson is a pseudoscalar, then the effect of the doubling symmetry cancels, and no oscillating states contribute:

$$\bar{\psi} \gamma_5 \psi \rightarrow (-1)^{2 \times t/a} \bar{\psi} (i \gamma_5 \gamma_0) \gamma_5 (i \gamma_5 \gamma_0) \psi = \bar{\psi} \gamma_5 \psi. \quad (4.43)$$

Bayesian χ^2 Fitting

We use the `CorrFitter` package [146] for performing the χ^2 fitting. We adopt a Bayesian approach first introduced in [147]. Given a fit function $f_\rho(x)$ with parameters $\{\rho_\alpha\}$, a set of inputs $\{x_i\}$, and a set of corresponding observations $\{y_i\}$, with a covariance matrix σ_{ij}^y , the fitter minimizes

$$\chi^2 = \sum_{ij} \frac{(f_\rho(x_i) - y_i)(f_\rho(x_j) - y_j)}{(\sigma_{ij}^y)^2} + \sum_{\alpha} \left(\frac{\rho_\alpha - \rho_\alpha^{\text{prior}}}{\sigma_\alpha^{\text{prior}}} \right)^2. \quad (4.44)$$

$\rho_\alpha^{\text{prior}}$ and $\sigma_\alpha^{\text{prior}}$ are the mean and standard deviations of the prior distributions given to the fit parameters. In our case, x_i is the set of times t , y_i are the correlators $C_i(t)$, and ρ_α are a_n, E_n, J_{nm} . Using this χ^2 means we take into account all correlations between different timeslices t , and between different correlators.

The actual parameters ρ_α of these fits are slightly reparameterized from simply the amplitudes a_j^M , energies E_j^M and transition amplitudes J_{jk} . Instead of energies, the fit parameters are $\log(\delta E_j^M)$, where $\delta E_0 = E_0^M$ and $\delta E_j = E_j^M - E_{j-1}^M$ for

$j > 0$. These are the parameters ρ given to Eq. (4.44), this is equivalent to giving Gaussian prior distributions to these logs. This forbids the ground state energy to become negative or go arbitrarily close to zero in the fit. Similarly, for the excited-state differences $E_j^M - E_{j-1}^M$, setting Gaussian priors for $\log(E_j^M - E_{j-1}^M)$ enforces $E_j^M > E_{j-1}^M$, a large reduction in the space of possible solutions to the fit. Often the fit is also given log-amplitudes rather than amplitudes as fit parameters. This also prevents the amplitudes a_j^M from becoming negative or zero. This is only an option when both the source and sink of the correlators being fitted have the same operator, otherwise, a_j^M are not necessarily positive.

A common problem for large fits involving many correlators is that the data's covariance matrix can be somewhat singular (very large condition number) if there are strong correlations in the data. This makes the inversion of the covariance matrix (for constructing χ^2) susceptible to roundoff error. To address this we impose an *svd cut* c_{svd} . This replaces any eigenvalue of the covariance matrix smaller than $c_{\text{svd}}x$ with $c_{\text{svd}}x$, where x is the largest eigenvalue in the matrix. This makes the matrix less singular. It can be considered a conservative move when it comes to the uncertainty of the results since the only possible effect this can have is to inflate those uncertainties.

4.2.2 Signal Degradation

A large obstacle in the analysis of correlation functions is *signal degradation* [148, 149].

A random variable x has a mean and standard deviation

$$\hat{x} = \langle x \rangle, \quad \sigma^2 = \frac{1}{N}(\langle x^2 \rangle - \langle x \rangle^2), \quad (4.45)$$

where N is the size of the sample. So the (square of) the signal/noise ratio is

$$\frac{\hat{x}^2}{\sigma^2} = N \left(\frac{\langle x^2 \rangle}{\langle x \rangle^2} - 1 \right)^{-1}. \quad (4.46)$$

Consider 2-point correlators where $x = \Phi^\dagger(t)\Phi(0)$, and Φ is some meson creation operator.

In the $t \rightarrow T_{\text{lat}}/2$ limit, $\langle x^2 \rangle$ and $\langle x \rangle$ can be written as

$$\langle x \rangle = \sum_n \frac{1}{2E_n} \langle \Omega | \Phi^\dagger(t) | \lambda_n \rangle \langle \lambda_n | \Phi(0) | \Omega \rangle e^{-E_n t} \sim e^{-E_0 t}, \quad (4.47)$$

$$\langle x^2 \rangle = \sum_n \frac{1}{2E_n} \langle \Omega | \Phi^{\dagger 2}(t) | \lambda_n \rangle \langle \lambda_n | \Phi^2(0) | \Omega \rangle e^{-E_n t} \sim e^{-E_0' t}. \quad (4.48)$$

where we have assumed the ratio of matrix elements and energies are $\mathcal{O}(1)$. The operator Φ^2 will contain two quark and two antiquark operators, connected by some matrices in spin space. Φ^2 can create a combination of all possible 2 meson states where the mesons are made of the available flavours and quantum numbers. For example, for 2-point D -meson correlators ($c\bar{d}$ pseudoscalars), $E'_0 = (M_\pi + M_{\eta_c})$. Plugging $E_0 = M_D$ and $E'_0 = (M_\pi + M_{\eta_c})$ into Eqs. (4.47), (4.48) and (4.46), we see that D meson correlators have a signal/noise ratio that degrades like

$$\frac{\hat{x}^2}{\sigma^2} \propto e^{-(M_D - (M_\pi + M_{\eta_c})/2)t}. \quad (4.49)$$

In general, a meson with two valence quarks of very different masses will suffer from a signal degrading exponentially with t . B -mesons suffer more than D -mesons. Adding spatial momentum to one of the quarks in the meson would have the effect of replacing M_D in the above equation with some higher energy E_D , thus exacerbating the problem further.

Signal degradation strongly limits the types of calculations that can be performed in lattice QCD. In the context of semileptonic decays, it can limit the region of q^2 that form factors can be calculated.

I have now introduced all of the relevant machinery for understanding the work of this thesis. The following three chapters cover research performed over the period of my PhD.

$b \rightarrow c$ Transitions with Lattice NRQCD

This chapter outlines of a number of projects attempted using the NRQCD action for the b quark. Much of the discussion in this chapter will concern the NRQCD-HISQ representation of the vector and axial $b \rightarrow c$ currents, i.e, the current if one of the quarks obeys NRQCD and the other obeys HISQ. I show here a number of attempts to improve the normalization of these currents (sections 5.2 and 5.4) and an attempt at a calculation of the $B \rightarrow D\ell\nu$ and $B_s \rightarrow D_s\ell\nu$ form factors (Sec. 5.3).

None of the work in this chapter reached a particularly satisfying conclusion. The salient result of this work is that using NRQCD for $b \rightarrow c$ currents away from zero recoil has some considerable issues.

5.1 NRQCD-HISQ Currents

Here I will define some notation used to describe the NRQCD-HISQ currents. To construct such a current, both the HISQ c and NRQCD b must be transformed into 4-component spinors such that they can be contracted with one-another in the current. The staggered c -quark χ_c is simply related to the naive spinor ψ_c by $\psi_c(x) = \Omega(x)\chi_c(x)$. The NRQCD b , $\Psi_b = (\Psi_+, 0)$, is a 2-component spinor related to the 4-component spinor ψ_b via an inverse Fouldy-Wouthuysen transform $\psi_b = \exp(-\gamma \cdot \nabla/2am_b)\Psi_b$. (∇ is defined here by $\nabla_\mu\psi(x) = (U_\mu(x)\psi(x + a\hat{\mu}) - U_\mu^\dagger(x - a\hat{\mu})\psi(x - a\hat{\mu}))/2$.)

Due to the Fouldy-Wouthuysen transform, a current $\bar{\psi}_c\Gamma\psi_b$ (where Γ is some product of gamma matrices) will be made of an infinite sum of lattice currents in terms of Ψ_b , $\bar{\psi}_c\Gamma\psi_b \sim \sum_j(1/am_b^j)\sum_k\bar{\psi}_c\mathcal{O}^{j,k}\Psi_b$. However, this is only half the story - as additional to the contribution from the Fouldy-Wouthuysen expansion, matching the lattice NRQCD theory to continuum QCD gives radiative corrections

to this series. The result is the series being populated by all operators $\mathcal{O}^{j,k}$ of dimension $-j$ with the same Lorentz indices as Γ .

So a continuum current J_μ is constructed from a series of the form

$$J_\mu = \sum_{j,k} c_j(\alpha_s, am_b) \frac{1}{(2am_b)^j} \bar{\psi}_c \mathcal{O}_\mu^{j,k} \Psi_b, \quad (5.1)$$

where j sums over powers of inverse b -mass and k sums over all operators of dimension $-j$. The coefficients $c_j(\alpha_s, am_b)$ are fixed by matching appropriate transition amplitudes in 1-loop continuum QCD and the lattice NRQCD/HISQ theory. The vector and axial vector currents take the general form [150]:

$$\begin{aligned} J_\mu &= (1 + z_0^{J_\mu} \alpha_s) J_{\mu,\text{lat}}^{(0)} + (1 + z_1^{J_\mu} \alpha_s) J_{\mu,\text{lat}}^{(1)} \\ &\quad + \alpha_s \sum_{n=2}^4 z_n^{J_\mu} J_{\mu,\text{lat}}^{(n)} + \mathcal{O}(\alpha_s^2, (\Lambda_{\text{QCD}}/m_b)^2, (\mathbf{p}/m_b)^2), \quad (5.2) \\ J_{\mu,\text{lat}}^{(0)} &= \bar{\psi}_c \Gamma_\mu \Psi_b, \quad J_{\mu,\text{lat}}^{(1)} = -\frac{1}{2am_b} \bar{\psi}_c \Gamma_\mu \gamma \cdot \nabla \Psi_b, \\ J_{\mu,\text{lat}}^{(2)} &= -\frac{1}{2am_b} \bar{\psi}_c \gamma \cdot \overleftarrow{\nabla} \gamma_0 \Gamma_\mu \Psi_b, \quad J_{\mu,\text{lat}}^{(3)} = -\frac{1}{2am_b} \bar{\psi}_c \Gamma_0 \nabla_\mu \Psi_b, \\ J_{\mu,\text{lat}}^{(4)} &= \frac{1}{2am_b} \bar{\psi}_c \overleftarrow{\nabla}_\mu \Gamma_0 \Psi_b, \end{aligned}$$

where Γ_μ is the continuum spin structure (e.g. for A_μ ; $\Gamma_\mu = \gamma_5 \gamma_\mu$) and \mathbf{p} is the momentum spatial exchange $\mathbf{p}_b - \mathbf{p}_c$. The last two currents $J_{\mu,\text{lat}}^{(3)}$ and $J_{\mu,\text{lat}}^{(4)}$ do not appear in the temporal current J_0 , $z_{3,4}^{J_0} = 0$.

A subset of the matching factors $\{z^{J_\mu}\}$ have been calculated for V_μ and A_μ in [108]. In the case where the charm is replaced with an s, u or d quark (therefore has negligible mass), results for $z_{0,1,2}^{J_\mu}$ are available for both V_μ and A_μ . However, in the $b \rightarrow c$ case the c mass must be taken into account which complicates the calculation. In this case, only $z_0^{J_\mu}$ is available. To sidestep this in studies using these currents, an extra truncation in the ‘‘cross-terms’’ of the perturbative and NRQCD series, $\alpha_s \Lambda_{\text{QCD}}/m_b$ and $\alpha_s \mathbf{p}/m_b$, is added resulting in

$$\begin{aligned} J_\mu &= (1 + z_0^{J_\mu} \alpha_s) (J_{\mu,\text{lat}}^{(0)} + J_{\mu,\text{lat}}^{(1)}) \\ &\quad + \mathcal{O}(\alpha_s^2, (\Lambda_{\text{QCD}}/m_b)^2, (\mathbf{p}/m_b)^2, \alpha_s \Lambda_{\text{QCD}}/m_b, \alpha_s \mathbf{p}/m_b). \quad (5.3) \end{aligned}$$

In the work of this thesis, we also compute $\langle J_{\mu,\text{lat}}^{(2,3,4)} \rangle$ in the lattice calculation to check that their magnitude is suitably small such that they can be ignored.

There are a number of orders here to consider, however the main order in which we will be concerned with is $\alpha_s \mathbf{p}/m_b$. The normalization of NRQCD-HISQ currents

have been demonstrated to be robust at zero spatial momentum (see for example [151]). However it was not yet known (before this work) whether the $\alpha_s \mathbf{p}/m_b$ terms were negligible.

5.2 Relativistic Normalization of the $b \rightarrow c$ Temporal Axial Current

In this small project, we tested to see if a B_c meson containing a HISQ c quark and an NRQCD b quark obeys a relativistic dispersion relation. The goal of this was to

- Provide a consistency check for the NRQCD-HISQ current truncation and normalization $z_0^{A_0}$ for the temporal axial current A_0 .
- If possible, fix $z_{1,2}^{A_0}$ for the $b \rightarrow c$ case by demanding the relativistic dispersion relation is obeyed.

5.2.1 Calculation Details

To test this process we computed B_c (pseudoscalar meson charged with b and c valence quarks) 2-point correlation functions on the fine ensemble (set 2 on table 4.1). The Wilson coefficients in the NRQCD action, the tadpole improvement factor u_0 , and the bare valence quark masses are given in the bottom row of Table 5.1. The interpolating operators for creating/annihilating the momentum-space B_c meson take the form

$$\tilde{\Phi}_n^\alpha(\mathbf{p}, t) = \sum_{\mathbf{x}, \mathbf{x}'} e^{-i\mathbf{p} \cdot \mathbf{x}} \bar{\psi}_c(\mathbf{x}, t) \phi^\alpha(\mathbf{x} - \mathbf{x}') \mathcal{O}_n \Psi_b(\mathbf{x}', t). \quad (5.4)$$

We chose \mathcal{O}_n to produce the current operators in the NRQCD-HISQ $b \rightarrow c$ current (5.2): $\mathcal{O}_0 = \gamma_0 \gamma_5$, $\mathcal{O}_1 = -\gamma_0 \gamma_5 \boldsymbol{\gamma} \cdot \nabla / 2m_b$, $\mathcal{O}_2 = -\boldsymbol{\gamma} \cdot \overleftarrow{\nabla} \gamma_0 \gamma_5 / 2m_b$. These have the same quantum numbers as the B_c meson (pseudoscalar with flavor $\bar{b}c$) so serve as suitable interpolating operators, but also let us probe the individual pieces of the NRQCD-HISQ $b \rightarrow c$ axial current.

In the NRQCD formalism, we simulate the b at its physical mass. Using physical mass b quarks cause severe signal degradation (see Sec. 4.2.2). To improve statistics, we compute a number of correlation functions using a family of *smearing* functions $\phi^\alpha(\mathbf{x} - \mathbf{x}')$:

$$\phi^0(\mathbf{y}) = \delta_{\mathbf{y}0}, \quad \phi^{r>0}(\mathbf{y}) = e^{-|\mathbf{y}|/a_{\text{sm}}^r}, \quad (5.5)$$

where $a_{\text{sm}}^1 = 3a$ and $a_{\text{sm}}^2 = 6a$. The $r > 0$ smearing functions represent a stationary c quark with a wavefunction for the c that exponentially decays with the radius from the b (in practice we implement the reverse - a stationary c surrounded by a b wavefunction - but this has the same effect on the correlation function). Using these reduces statistical errors in two ways. Firstly, it means we have more samples. Secondly, the smearing functions increase the overlap with the B_c meson state $\langle \Omega | \tilde{\Phi}_n^\alpha | B_c \rangle$, which decreases the overlap with excited states, therefore decreasing the contribution of excited states to the correlation functions. One can then afford to use timeslices closer to the source, therefore increasing statistics further. Note that in order to use these smearing functions one must fix the gauge of the configurations to Coulomb gauge.

The NRQCD-HISQ correlation functions are then generated using

$$C_{nm}^{\alpha\beta}(\mathbf{p}, t) = \sum_{\mathbf{x}, \mathbf{x}'} \sum_{\mathbf{y}, \mathbf{y}'} \phi^\beta(\mathbf{x} - \mathbf{x}') \phi^\alpha(\mathbf{y} - \mathbf{y}') \left\langle \text{Tr}_c \left[g_c^{\theta_{\mathbf{p}}}(\mathbf{x}', t; \mathbf{y}', t_0) \text{Tr}_s \left(\gamma_5 \Omega(\mathbf{y}', t_0) \Omega^\dagger(\mathbf{x}', t) \gamma_5 \mathcal{O}_m G_b(\mathbf{x}, t; \mathbf{y}, t_0) \mathcal{O}_n \right) \right] \right\rangle. \quad (5.6)$$

Tr_s is a trace over spin and Tr_c is over color. $g_c^{\theta_{\mathbf{p}}}$ is a staggered propagator given momentum twist $\theta_{\mathbf{p}}$ corresponding to a momentum \mathbf{p} , and G_b is an NRQCD b propagator. $\langle \rangle$ denotes an average over gauge configurations. Eq. (5.6) can be arrived at from $\langle \tilde{\Phi}_n^\alpha(\mathbf{p}, t_0) \tilde{\Phi}_m^{\beta\dagger}(\mathbf{p}, t) \rangle$, applying the steps given in Sec. 4.1, and converting the charm propagator to a staggered propagator as in Sec. 4.1.3.

We generated these correlators on 500 configurations and 16 choices for t_0 evenly distributed across the temporal extent of the lattice. We obtained correlators at 3x different spatial momenta, $a\mathbf{p} = 0, 3\pi(1, 1, 1)/L, 5\pi(1, 1, 1)/L$ ($L = 32$), using momentum twists $\theta = 0, 3, 5$ in each direction. The resulting correlators are shown in Fig. 5.1.

These were then fitted to the fit functions

$$C_{nm}^{\alpha\beta}(t)|_{\text{fit}} = \sum_{j=0}^{N_{\text{exp}}} \left(a_j^{\alpha,n} a_j^{\beta,m} f(\bar{E}_j, t) + (-1)^{t/a} a_{j,o}^{\alpha,n} a_{j,o}^{\beta,m} f(\bar{E}_{j,o}, t) \right). \quad (5.7)$$

See Sec. 4.2.1 for definitions of f and \bar{E} . One can recognise that

$$a_0^{\alpha,n} = \frac{\langle \Omega | \tilde{\Phi}_n^\alpha | B_c \rangle}{\sqrt{2E_{B_c}}}. \quad (5.8)$$

In the $\alpha = 0$ case, the matrix elements become $\langle \Omega | \bar{\psi}_c \mathcal{O}_n \Psi_b | B_c \rangle$. Combining these as in Eq (5.2) should produce $\langle \Omega | A_0 | B_c \rangle$. I will show how these quantities are used to test the A_0 normalisation after a brief detour.

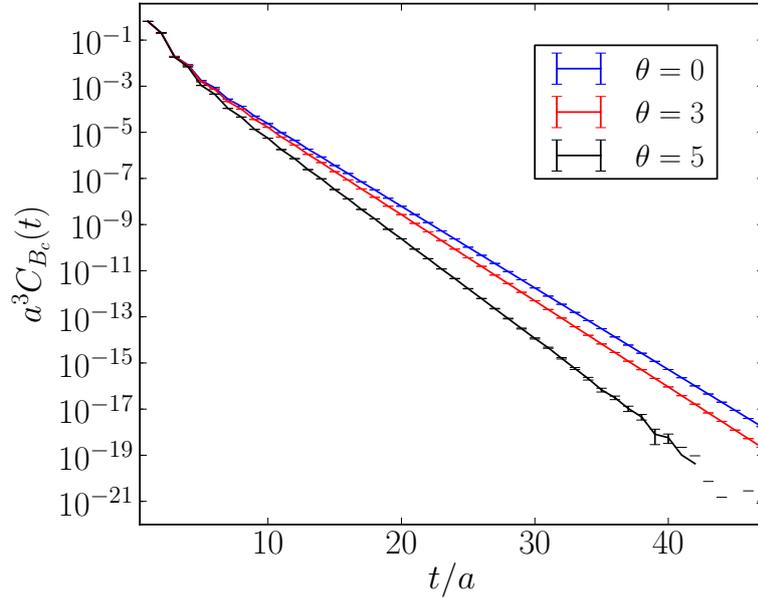


Figure 5.1: B_c 2-point correlation functions, with the valence charm given momentum twists $\theta = 0, 3$ and 5 .

5.2.2 Kinetic Mass

We require a determination of the mass of the meson in our simulation. If we were using a fully relativistic action, one could simply consider \bar{E}_0 (with $\mathbf{p} = 0$) to be the mass. However, in our case one would expect NRQCD to cause a shift in energy E_s due to the effective removal of the rest mass, so

$$\bar{E}_0(\mathbf{p}) \equiv E_s + \sqrt{\mathbf{p}^2 + M_{\text{kin}}^2}. \quad (5.9)$$

We can deduce M_{kin} in this case by taking the difference of energies at different momenta $\delta\bar{E}_0(\mathbf{p}) \equiv \bar{E}_0(\mathbf{p}) - \bar{E}_0(0)$ and rearranging to find

$$M_{\text{kin}} = \frac{\mathbf{p}^2 - \delta\bar{E}_0^2(\mathbf{p})}{2\delta\bar{E}_0(\mathbf{p})}, \quad (5.10)$$

which one would expect to be invariant of \mathbf{p} . M_{kin} is referred to as the *kinetic mass* of the meson in question.

Using \bar{E}_0 results from the fit, we find $aM_{\text{kin}}^{\theta=3} = 2.8394(60)$, from the $\theta = 3$ point, and $aM_{\text{kin}}^{\theta=5} = 2.858(11)$ from the $\theta = 5$ point. Taking the mean of these we find

$$aM_{\text{kin}} = 2.8488(125). \quad (5.11)$$

5.2.3 Decay Amplitude Ratios

At leading order in $1/m_b$ and α_s , the temporal axial current is recreated using simply $a_0^{0,0} = \langle \Omega | A_0 | B_c \rangle / \sqrt{2M_{B_c}}$. Recalling the definition of the decay constant for a pseudoscalar meson: $\langle \Omega | A_\mu | M \rangle = p_\mu f_M$, we see that

$$a_0^{0,0} = f_{B_c} \sqrt{\frac{E_{B_c}}{2}}. \quad (5.12)$$

Assuming a relativistic dispersion relation $E^2 = \mathbf{p}^2 + M^2$, taking the ratio of $a_0^{0,0}$ at non-zero and zero momenta results in

$$\frac{a_0^{0,0}(\mathbf{p})}{a_0^{0,0}(0)} = \sqrt{\frac{E_{B_c}(\mathbf{p})}{M_{B_c}}} = 1 + \frac{\mathbf{p}^2}{4M_{B_c}^2} + \mathcal{O}\left(\frac{\mathbf{p}^4}{M_{B_c}^4}\right). \quad (5.13)$$

This is our probe of the dispersion relation of the B_c meson. We took the ratio of $a_0^{0,0}$ fit parameters on the left-hand side, and compare this to the expected dependence of \mathbf{p}^2 on the right-hand side. This comparison is shown between the blue line and the grey dotted line in Fig. 5.2. We have used the kinetic mass (5.11) for the M_{B_c} mass here.

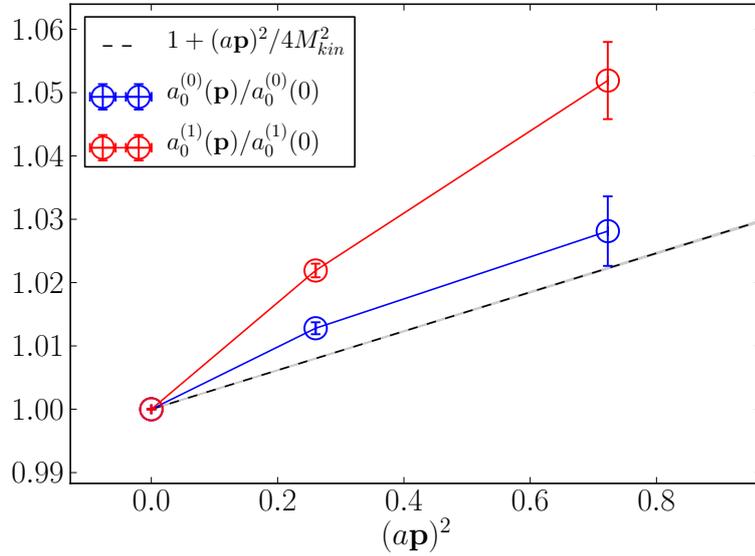


Figure 5.2: Decay amplitude ratios (colourful points) against the expected relativistic behaviour (grey dotted line and band). Adding the $A_{0,\text{lat}}^{(1)}$ piece of the current does not improve the relativistic behaviour of the ratio.

We gradually add corrections to this ratio by replacing $a_0^{0,0}$ with

$$\begin{aligned} a_0^{(0)}(\mathbf{p})\sqrt{2E_{B_c}(\mathbf{p})} &= \langle \Omega | A_{0,\text{lat}}^{(0)} | B_c(\mathbf{p}) \rangle, \\ a_0^{(1)}(\mathbf{p})\sqrt{2E_{B_c}(\mathbf{p})} &= \langle \Omega | (1 + z_0^{A_0} \alpha_s) [A_{0,\text{lat}}^{(0)} + A_{0,\text{lat}}^{(1)}] | B_c(\mathbf{p}) \rangle, \\ a_0^{(2)}(\mathbf{p})\sqrt{2E_{B_c}(\mathbf{p})} &= \langle \Omega | [(1 + z_0^{A_0} \alpha_s) A_{0,\text{lat}}^{(0)} + (1 + z_1^{A_0} \alpha_s) A_{0,\text{lat}}^{(1)} \\ &\quad + z_2^{A_0} \alpha_s A_{0,\text{lat}}^{(2)}] | B_c(\mathbf{p}) \rangle. \end{aligned} \quad (5.14)$$

We have here set the $\alpha = 0, n$ superscripts implicit to make room for the new superscripts. The lattice currents $A_{0,\text{lat}}^{(n)}$ are those defined in Eq. (5.2) for the temporal axial vector case. $a_0^{(0)}$ recreates the A_0 current to leading order in α_s and $1/m_b$, $a_0^{(1)}$ recreates A_0 up to order $\mathcal{O}(\alpha_s^2, (\Lambda_{\text{QCD}}/m_b)^2, (\mathbf{p}/m_b)^2, \alpha_s \Lambda_{\text{QCD}}/m_b, \alpha_s \mathbf{p}/m_b)$, and $a_0^{(2)}$ is up to order $\mathcal{O}(\alpha_s^2, (\Lambda_{\text{QCD}}/m_b)^2, (\mathbf{p}/m_b)^2)$.

Since the $z_0^{A_0}$ value is immediately available from [108], we can show the result of taking the ratio $a_0^{(1)}(\mathbf{p})/a_0^{(1)}(0)$ as the red line in Fig. 5.2. As can be seen here, going from $a_0^{(0)}$ to $a_0^{(1)}$ pushes the ratio in the wrong direction, away from the relativistic dispersion relation line (the grey dotted line).

We can determine values for $(z_1^{A_0} - z_0^{A_0})$ and $(z_2^{A_0} - z_0^{A_0})$ by demanding that $a_0^{(2)}(\mathbf{p})/a_0^{(2)}(0) = 1 + \mathbf{p}^2/4M_{\text{kin}}^2$. Then, using the known $z_0^{A_0}$ values from perturbative matching we find

$$z_1^{A_0} = -3.746267(44), \quad z_2^{A_0} = -0.000910(36). \quad (5.15)$$

$z_1^{A_0}$ here is required to be unnaturally large to overcome the suppression of α_s and drag the ratio downwards.

The above analysis shows that the truncation of NRQCD-HISQ temporal-axial current given in Eq. (5.3) is not sufficient to create a meson obeying a relativistic dispersion relation. This is perhaps indicative that further orders in the expansion are in fact important and should be included in lattice calculations.

5.3 $B_{(s)} \rightarrow D_{(s)}\ell\nu$ Form Factors

I attempted a calculation of the $B \rightarrow D\ell\nu$ and $B_s \rightarrow D_s\ell\nu$ form factors, $f_{0,+}(q^2)$ and $f_{0,+}^s(q^2)$, using the 2+1+1 MILC ensembles, HISQ l,s and c valence quarks, and an NRQCD valence b quark. This study was similar to previous studies of $B \rightarrow D\ell\nu$ form factors [36] and $B_s \rightarrow D_s\ell\nu$ form factors [152]. The main difference between this and the previous studies was that they used older MILC ensembles that do not take the charm into account in the sea.

Set	name	am_{s0}^{val}	am_{c0}^{val}	am_{b0}^{val}	u_0	$c_{1,6}$	c_5	c_4	T/a	a_{sm}/a
0	very coarse	0.0705	0.826	3.297	0.8195	1.36	1.21	1.22	8, 11, 14	0, 2.0, 4.0
1	coarse	0.0541	0.645	2.66	0.8340	1.31	1.16	1.20	9, 12, 15	0, 2.0, 4.0
2	fine	0.0376	0.450	1.91	0.8525	1.21	1.12	1.16	14, 19, 24	0, 3.425, 6.85

Table 5.1: Parameters used in our calculation. am_{s0}^{val} and am_{c0}^{val} are the bare masses of the strange and charm valence quarks, tuned in [153]. am_{b0}^{val} is the bare mass of the valence bottom quark, tuned in [119]. $u_0 = (\text{Re Tr}\langle \square \rangle / 3)^{1/4}$ is the ‘tadpole improvement parameter’ as used in [119] via a perturbative calculation of the plaquette \square . $\{c_i\}$ are the coefficients for the kinetic and chromomagnetic terms in the NRQCD action (Eq. (3.60)) [154]. $\{T\}$ is the set of temporal separations between source ($B_{(s)}$ creation operator) and sink ($D_{(s)}$ annihilation operator). a_{sm} are the radii of the exponential smearing function applied to the $B_{(s)}$ and $D_{(s)}$ creation operators.

One motivation for this study was to test how far down the q^2 range we could reach with lattice data before the noise in the correlators made the data useless. Another was to investigate the size and \mathbf{p} -dependence on the subleading currents $V_{\mu, \text{lat}}^{(2,3,4)}$ as one moves away from zero recoil.

This study was not completed on account of two major problems:

- The $\mathcal{O}(\alpha_s \mathbf{p}/m_b)$ terms in the vector NRQCD-HISQ current that we must ignore due to the lack of perturbative normalizations, $V_k^{(2)}$ and $V_k^{(4)}$, turned out to be significant in magnitude.
- On one ensemble in the $B_s \rightarrow D_s \ell \nu$ case, there was an anomalous result for the vector current matrix element extracted from the correlator fits.

I will give an outline of the calculation here for completeness, but the crucial findings of this section are these two issues.

5.3.1 Lattice Setup

Correlation functions were generated on three MILC ensembles, sets 0, 1 and 2 in Table 4.1. When using the NRQCD action, one is limited to the coarser end of the spectrum of ensembles. This is because in the $a \rightarrow 0$ limit subleading terms in δH (eq. (3.60)) and $J_\mu^{(n>0)}$ (eq. (5.2)) diverge, since the $1/m_b$ factors are in fact

proportional to $1/am_b$, resulting negative powers of the lattice spacing. However, NRQCD discretisation effects are small relative to other discretizations due to the lack of the b rest mass, so we can afford to use coarser lattices. Also, using coarse lattices means the project is computationally inexpensive since, for example, Dirac matrices to be inverted are smaller. The bare parameters used to generate the correlation functions are shown in table 5.1.

We generated 2-point correlation functions for $B_{(s)}$ and $D_{(s)}$ mesons, and 3-point correlators between $B_{(s)}$ and $D_{(s)}$ interpolating operators with $V_\mu^{(n)}$ currents inserted for all μ and $n = 0, 1, 2 (\mu = 0)$ and $n = 0, 1, 2, 3, 4 (\mu = 1, 2, 3)$. For the $B_{(s)}$ operator we use exponential smearing functions (like those introduced in Eq. (5.5)), smearing radii a_{sm} are given in Table 5.1.

The $B_{(s)}$ 2-point correlators, $C_{B_{(s)}}^{\alpha\beta}(t)$ were generated using Eq. (5.6), with the charm propagator replaced with a strange or light propagator, and $\mathcal{O}_m = \gamma_0 \gamma_5$. We also computed $D_{(s)}$ 2-point correlators at a number of spatial momenta $\{\mathbf{p}\}$ (given in Table 5.2), generated by

$$C_{D_{(s)}}^{\alpha\beta}(\mathbf{p}, t) = \sum_{\mathbf{x}, \mathbf{x}'} \sum_{\mathbf{y}, \mathbf{y}'} \phi^\alpha(\mathbf{x} - \mathbf{x}') \phi^\beta(\mathbf{y} - \mathbf{y}') \left\langle \text{Tr}_c [g_c^{\theta_{\mathbf{p}}}(\mathbf{x}, t, \mathbf{y}; t_0) g_{l(s)}^\dagger(\mathbf{x}', t; \mathbf{y}', t_0)] \right\rangle, \quad (5.16)$$

where $\phi^\alpha(\mathbf{x})$ are the smearing functions (Eq. (5.5)), $g_{l(s)}$ are light or strange staggered propagators, and $g_c^{\theta_{\mathbf{p}}}$ is a charm staggered propagator with momentum twist $\theta_{\mathbf{p}}$. Tr_c is over color.

We generated 3-point correlators for each individual piece of the NRQCD-HISQ current, and each \mathbf{p} , using

$$C_{V_\mu^{(n)}}^{\alpha\beta}(\mathbf{p}, t, T) = \sum_{\mathbf{x}, \mathbf{y}, \mathbf{z}} (-1)^{\sum_{k=1}^3 x_k/a} \phi^\alpha(\mathbf{x} - \mathbf{x}') \phi^\beta(\mathbf{z} - \mathbf{z}') \times \left\langle \text{Tr}_c \left(g_c^{\theta_{\mathbf{p}}}(\mathbf{x}, t_0; \mathbf{y}, t) g_{l(s)}^\dagger(\mathbf{x}, t_0; \mathbf{z}, T) \text{Tr}_s \left[\gamma_0 \Omega^\dagger(\mathbf{y}, t) \mathcal{O}_{n,\mu} G_b(\mathbf{y}, t; \mathbf{z}, T) \Omega(\mathbf{z}, T) \gamma_0 \right] \right) \right\rangle. \quad (5.17)$$

$\mathcal{O}_{n,\mu}$ are defined by $J_\mu^{(n)} = \bar{\psi}_c \mathcal{O}_{n,\mu} \Psi_b$, where $J_\mu^{(n)}$ are pieces of the NRQCD-HISQ vector current (Eq. (5.2)).

The list of twists we used on each ensemble is given in table 5.2. Due to the signal/noise degradation of the $D_{(s)}$ correlators as one adds more spatial momentum, our lattice data was limited to the high q^2 region.

	Set	θ	$ a\mathbf{p} $	$q^2[\text{GeV}^2]$
$B \rightarrow D$	0	0, 0.74, 1.47, 2.20, 2.94	0, 0.25, 0.5, 0.75, 1.00	11.8, 11.6, 10.8, 9.4, 7.3
	1	0, 1.58, 2.24, 4.53	0, 0.36, 0.51, 1.02	11.8, 10.8, 9.9, 5.0
	2	0, 1.76, 2.64	0, 0.30, 0.49	11.8, 10.7, 9.3
$B_s \rightarrow D_s$	0	0, 0.74, 1.47, 2.20, 2.94	0, 0.25, 0.5, 0.75, 1.03	11.8, 11.6, 10.9, 9.5, 7.6
	1	0, 1.10, 2.20, 3.31, 4.41	0, 0.25, 0.50, 0.75, 1.00	11.8, 11.3, 10.1, 8.1, 5.7
	2	0	0	11.8

Table 5.2: Momentum twists (and corresponding momenta and q^2 values) given to the charm propagator on each ensemble.

5.3.2 Correlator Fits

We extracted current matrix elements from the generated correlation functions, via simultaneous Bayesian fits as described in Sec. 4.2.1. For the set of 2-point correlators we use Eq. (5.7) (with $n = m = 0$), and 3-point correlators are fit to

$$\begin{aligned}
C_3^{\alpha\beta}(t, T)|_{\text{fit}} = & \sum_{j,k=0}^{N_{\text{exp}}, N_{\text{exp}}} \left(a_j^{B(s)} J_{jk}^{nn} a_k^{D(s)} f(E^{B(s)}, t) f(E_n^{D(s)}, T-t) \right. \\
& + a_{\alpha,j}^{B(s),o} J_{jk}^{on} a_{\beta,k}^{D(s)} (-1)^t f(E_n^{B(s),o}, t) f(E^{D(s)}, T-t) \\
& + a_{\alpha,j}^{B(s)} J_{jk}^{no} a_{\beta,k}^{D(s),o} (-1)^{T-t} f(E^{B(s)}, t) f(E_n^{M^*,o}, T-t) \\
& \left. + a_{\alpha,j}^{B(s),o} J_{jk}^{oo} a_{\beta,k}^{D(s),o} (-1)^T f(E_n^{B(s),o}, t) f(E^{D(s),o}, T-t) \right). \quad (5.18)
\end{aligned}$$

We set $N_{\text{exp}} = 5$ in each fit. We performed a single simultaneous fit containing each correlator computed for each ensemble, taking into account correlations between all time slices of all correlation functions involved in the fit.

We set Gaussian priors for the parameters J_{jk} , and log-normal priors for most other parameters. Using log-normal distributions ensures energies E_n^M and amplitudes a_n^M are positive and forbids them from moving arbitrarily close to zero, improving the stability of the fit. The exception is the smeared amplitudes $a_n^{\alpha>0}$, which can be zero or negative, so we simply set Gaussian priors for these.

Priors for the ground state energies and amplitudes (oscillating and non-oscillating) are set via an empirical Bayes approach. Plots of effective masses and amplitudes are inspected to find reasonable central values of priors (see Sec. 6.2.2

for definitions of effective masses and amplitudes). A generic variance of 10% of the central value is given in all cases. The typical precision of fit results for these parameters is of order 0.1%, much more precise than their priors. The log of all excited state energies are given generic priors $\log(\Lambda_{\text{QCD}} \pm \Lambda_{\text{QCD}}/2)$, where we set $\Lambda_{\text{QCD}} = 0.5\text{GeV}$. The log of the excited state (non-smear) amplitudes are set to $\log(0.3 \pm 0.2)$. The smeared excited state amplitudes are given priors of 0.3 ± 0.2 . All 3-point transition parameters J_{jk} are given priors of 0 ± 1 .

We typically set $t_{\text{cut}} = 3$ for 2- and 3-point correlators in the fit on the fine ensemble (set 2) and $t_{\text{cut}} = 2$ on very coarse and coarse (sets 0 and 1). For some specific correlators, a larger t_{cut} is necessary to achieve a good fit ($\chi^2/N_{\text{dof}} < 1$), these higher t_{cut} values are always in the range $t_{\text{cut}} \in [2, 8]$. An svd-cut is applied in each fit, with a specific choice of cut determined according to what achieves a good fit. The svd-cut is typically of the order 10^{-3} .

The current matrix element we require can be found via

$$\langle D_s | V_\mu^{(n)} | B_s \rangle |_{\text{lat}} = 2\sqrt{M_{B_s} E_{D_s}} J_{00}^{nn}. \quad (5.19)$$

5.3.3 Form Factors

We constructed 'continuum' vector currents $\langle D_{(s)} | V_\mu | B_{(s)} \rangle \equiv \langle V_\mu \rangle$ from the lattice expectation values $\langle D_{(s)} | V_\mu^{(n)} | B_{(s)} \rangle |_{\text{lat}}$ according to Eq. (5.3), i.e. only including the first two current terms $V_\mu^{(0)}$ and $V_\mu^{(1)}$. We also computed $V_0^{(2)}$ and $V_k^{(2,3,4)}$ to assess their size and the validity of ignoring them, this is addressed in Sec. 5.3.4.

We took the average over the spatial currents, resulting in two distinct current matrix elements $\langle V_0 \rangle$ and $\langle V_k \rangle$. Then, from the definition of pseudoscalar-pseudoscalar form factors (Eq. (2.24)), we find (defining $M_{B_{(s)}} \equiv M$, $M_{D_{(s)}} \equiv m$, $E_{D_{(s)}} \equiv E$, and $\mathbf{p}_{D_{(s)}} \equiv \mathbf{p}$)

$$\langle V_0 \rangle = f_+^{(s)}(q^2) \left[M + E - \frac{M^2 - m^2}{q^2} (M - E) \right] + f_0^{(s)}(q^2) \frac{M^2 - m^2}{q^2} (M - E), \quad (5.20)$$

$$\langle V_k \rangle = \frac{|\mathbf{p}|}{\sqrt{3}} \left[f_+^{(s)}(q^2) \left(1 + \frac{M^2 - m^2}{q^2} \right) - f_0^{(s)}(q^2) \frac{M^2 - m^2}{q^2} \right]. \quad (5.21)$$

By inverting these relations we deduce $f_{0,+}^{(s)}(q^2)$ from $\langle V_0 \rangle, \langle V_k \rangle$ at the lattice spacings of each ensemble.

We aim then to extrapolate these form factors to $a = 0$ and to the full physical q^2 range. We did not obtain any lattice data at light quark masses smaller than

$m_l/m_s = 0.2$ so cannot extrapolate to the physical light mass. In the $B_s \rightarrow D_s \ell \nu$ case, the size of such an effect is small (see chapters 6 and 7). In the $B \rightarrow D \ell \nu$ case, however, this could result in considerable systematic errors.

We parameterised the functional form of $f_{0,+}^{(s)}(q^2)$ using the BCL parameterization [155]. This involves first defining the map

$$z(q^2) = \frac{\sqrt{t_+ - q^2} - \sqrt{t_+ - t_0}}{\sqrt{t_+ - q^2} + \sqrt{t_+ - t_0}}, \quad (5.22)$$

where $t_{\pm} = (M_{B_{(s)}} \pm M_{D_{(s)}})^2$ and we choose $t_0 = t_+(1 - \sqrt{1 - t_-/t_+})$, and $z(q^2)$ has a very small magnitude throughout the entire q^2 range, in our case $|z| < 0.032 \forall$ physical q^2 . $f_{+,0}^{(s)}(q^2)$ can be expressed as a series expansion in z :

$$f_{0,+}(q^2) = \frac{1}{P_{0,+}(q^2)} \sum_{k=0}^K a_k^{0,+} z(q^2)^k. \quad (5.23)$$

We truncate this at $K = 2$, adding further terms appear to have no effect on the fit. The factors $P(q^2)$ are defined by

$$P_{0,+}(q^2) = \left(1 - \frac{q^2}{M_{0,+}^2}\right). \quad (5.24)$$

These are required due to subthreshold poles in the crossed channel of $\langle D_{(s)} | V_{\mu} | B_{(s)} \rangle$, which in our case is a W decay into a B_c^* meson. The pole is located where the W has the correct momentum q^2 to create the B_c , hence at $q^2 = M_{B_c^*}^2$. This is not within the q^2 range, but can create curvature in $f_{0,+}$ that can confound the expansion in z . $P_{0,+}$ effectively removes this pole from the z expansion.

The discretisation effects in our form factors are controlled for by modifying (5.23):

$$a_n^{0,+} \rightarrow a_n^{0,+} \times (1 + b_n^{0,+} (am_{c0}^{\text{val}})^2), \quad (5.25)$$

where $b_n^{0,+}$ are new fit parameters. $am_c \rightarrow 0$ in the continuum limit, and, since the charm mass is the largest scale involved in our calculation, it serves as a good order parameter for discretization effects. Hence, here we are extrapolating in both z and to continuum simultaneously. $\{a_n^{0,+}, b_n^{0,+}\}$ are all given Gaussian prior distributions of 0 ± 1 .

5.3.4 Results

The extrapolation of our lattice data to all q^2 and $a = 0$ is illustrated in Figures 5.3 and 5.4. As can be seen here, statistical errors in the lattice data increase

exponentially with momentum twist (as q^2 decreases). Besides this, the very coarse data suffers from large discretization effects as the twist is increased, pushing the results upwards.

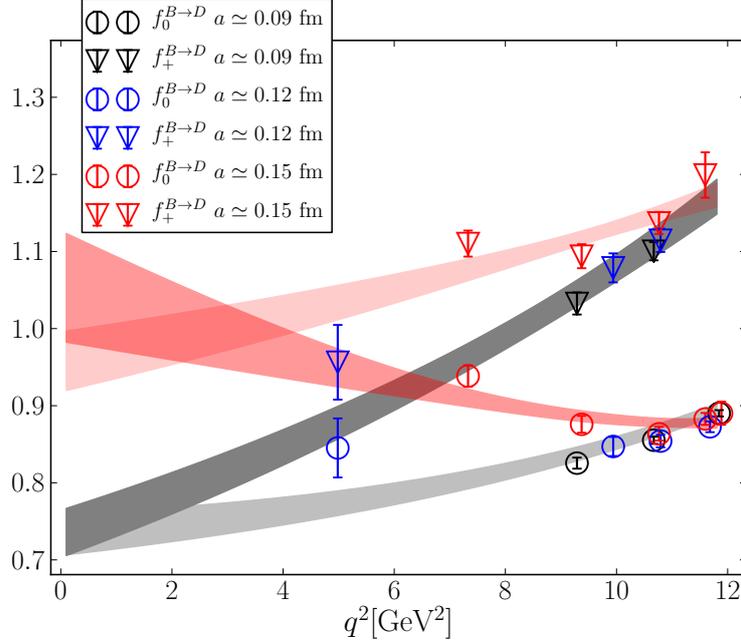


Figure 5.3: $B \rightarrow D\ell\nu$ form factors. The coloured points show lattice data, each color represents an ensemble. The grey band represents the continuum and kinematically extrapolated result. The red band shows the kinematically extrapolated result at $a \simeq 0.15\text{fm}$.

I now go on to discuss the two issues which led me to abandon this project.

Anomalous Results

In the $B_s \rightarrow D_s\ell\nu$ case I have not include lattice results from the fine ensemble in the q^2 and $a \rightarrow 0$ extrapolation. This is because the lattice results for $f_0^s(q^2)$ on this ensemble are clearly wrong. We have *a priori* knowledge of what, for example, the ballpark of $f_0^s(q_{\text{max}}^2)$ should be from a couple of sources:

- The result should not vary much more than $\mathcal{O}(a^2)$ (where a is the lattice spacing) from the same result on other ensembles.
- The result should not vary much more than $\mathcal{O}(am_{s_0}^{\text{val}} - am_{l_0}^{\text{val}})$ from the same number D on the same ensemble for the $B \rightarrow D$ calculation (Chiral symmetry).

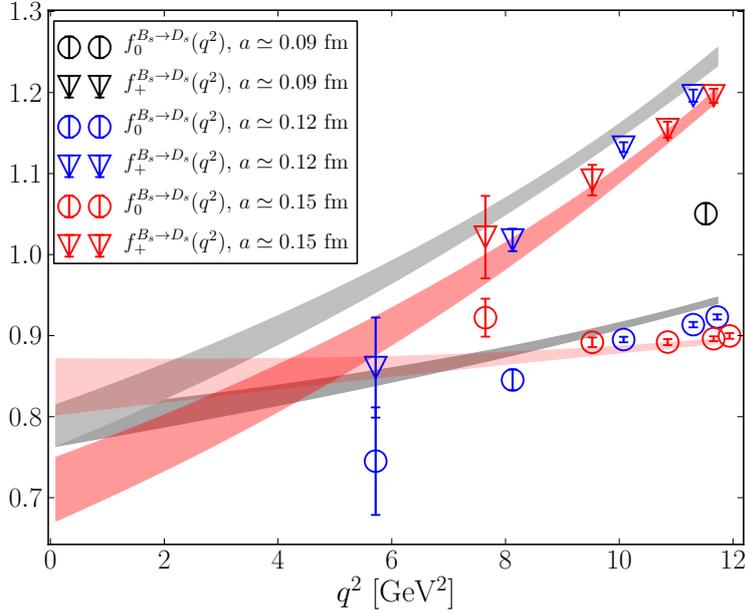


Figure 5.4: $B_s \rightarrow D_s \ell \nu$ form factors. The coloured points show lattice data, each color represents an ensemble. The grey band represents the continuum and kinematically extrapolated result. The red band shows the kinematically extrapolated result at $a \simeq 0.15\text{fm}$. Note that the $a \simeq 0.09\text{fm}$ data is not included in the extrapolation, since this prevents a good fit from being possible, see Sec. 5.3.4.

However, we find the fits to q_{max}^2 data on the fine ensemble produce a result for J_{00}^{mn} that is much larger than what is expected from these considerations. This is accompanied by the fits being very unstable, varying by a number of sigmas when different combinations of data are included, and different hyperparameters (svd-cut, t_{cut} , etc) are included. A number of tests have been carried out to find out exactly what is causing this issue, but no compelling evidence has emerged for any explanation. Fig. 5.5 illustrates the situation.

It is worth keeping in mind that NRQCD results have no continuum limit since the action and the currents are truncated sums of inverse masses in lattice units, therefore the truncation error grows like a^{-n} as $a \rightarrow 0$. What we are seeing here may be the result of large $1/(am_b)$ and $1/(am_b)^2$ corrections to the NRQCD-HISQ current being ignored.

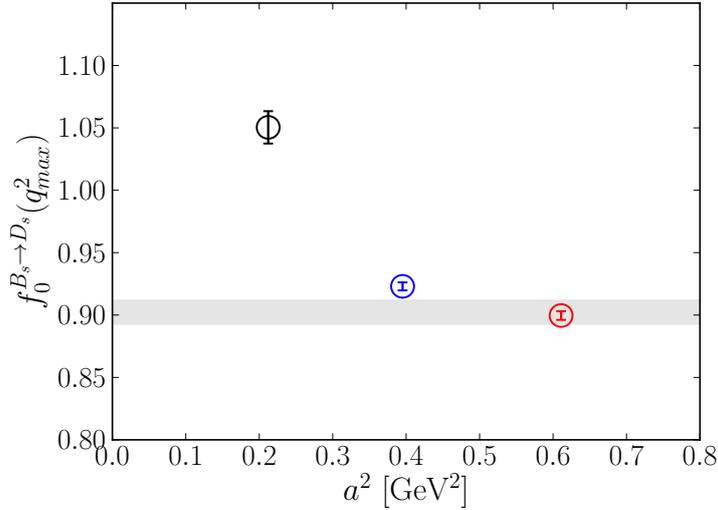


Figure 5.5: Lattice results for $f_0^s(q_{\max}^2)$ against a^2 . The grey band shows the result for $f_0^s(q_{\max}^2)$ computed in chapter 7 using the Heavy-HISQ approach for comparison. Clearly the result on the fine ensemble (the black point), and possibly on the coarse ensemble (the blue point), contain large unknown systematic errors.

Large Subleading Currents

Another problem that has uncovered itself in the NRQCD calculation is large subleading currents. Namely, the pieces $V_k^{(2,4)}$ of the spatial vector current. We determined these currents as part of the calculation in order to assess if they are suitably small such that they can be ignored. These turned out to have a magnitude $\sim 35\%$ of the leading order.

In Fig. 5.6, we show the ratios of (matrix elements of) NRQCD-HISQ currents in the $B_s \rightarrow D_s$ case. All ratios are between the subleading currents $V_\mu^{(n>0)}$ and the leading order current $V_\mu^{(0)}$, in order to show the size of the subleading currents relative to the leading order. $V_\mu^{(1)}$ is included in our result so we do not need to worry about its size. $V_0^{(2)}$ and $V_k^{(3)}$ are $\lesssim 10\%$ of the leading order, given that these also receive $\mathcal{O}(\alpha_s)$ suppression, their negligence is relatively harmless.

$V_k^{(2,4)}$, however, have considerable magnitude. Neglecting them implies a naive systematic error of $\mathcal{O}(35\% \times \alpha_s) \sim 8\%$. This would prevent any results from our calculation from being anywhere near competitive. These two currents are of order $\alpha_s \mathbf{p}/m_b$, so their magnitude would likely only increase as we move towards $q^2 = 0$.

It should be noted here that it is possible that the contribution from these cur-

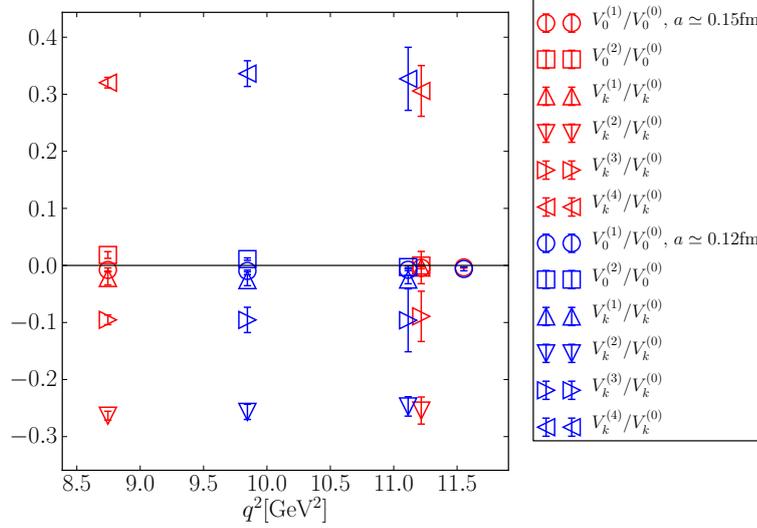


Figure 5.6: Ratios of (matrix elements of) the subleading NRQCD-HISQ currents $V_\mu^{(n>0)}$ to the leading order current $V_\mu^{(0)}$, in the $B_s \rightarrow D_s$ case.

rents turn out to be smaller than $\sim 8\%$. This is because individual terms of the NRQCD-HISQ current mix under renormalization, and can cause large cancellations (see for example [108]). To calculate the effect of this mixing would require a large perturbative calculation however. Since no calculation has been performed, we are stuck with the large systematic error implied by the large $V_k^{(2,4)}$ expectation values.

Since the problematic current pieces are exclusively part of the spatial vector current, we could remove this problem if we did not rely on the spatial vector current for extracting the form factors. The next section shows our attempt at such an alternative approach.

5.3.5 Form Factors from V_0 and S

Instead of using $\langle V_0 \rangle$ and $\langle V_k \rangle$ to extract $f_{0,+}^{(s)}(q^2)$, one could in principle instead use the combination $\langle V_0 \rangle$ and $\langle S \rangle$, where S is the scalar $b \rightarrow c$ density. The individual terms in the scalar NRQCD-HISQ current can be related to the temporal-vector current by using the property $\gamma_0 \Psi_b = \Psi_b$. Hence one can write the scalar NRQCD-

HISQ current as [151]:

$$S = (1 + z_0^S \alpha_s) V_0^{(0)} - (1 + z_1^S \alpha_s) V_0^{(1)} + z_2^S \alpha_s V_0^{(2)} \quad (5.26)$$

$$\begin{aligned} &+ \mathcal{O}(\alpha_s^2, (\Lambda_{\text{QCD}}/m_b)^2, (\mathbf{p}/m_b)^2), \\ &= (1 + z_0^S \alpha_s)(V_0^{(0)} - V_0^{(1)}) \quad (5.27) \\ &+ \mathcal{O}(\alpha_s^2, (\Lambda_{\text{QCD}}/m_b)^2, (\mathbf{p}/m_b)^2, \alpha_s \Lambda_{\text{QCD}}/m_b, \alpha_s \mathbf{p}/m_b). \end{aligned}$$

z_0^S can be derived from $z_0^{V_\mu}$. Hence we already have numerical results for matrix elements of the scalar current, via the vector current pieces $V_0^{(0,1)}$. The scalar and temporal vector currents are related at zero recoil via the PCVC relation

$$(M_{B_{(s)}} - E_{D_{(s)}}) \langle V_0 \rangle = \delta m \langle S \rangle, \quad (5.28)$$

where $\delta m \equiv (m_b - m_c)$. Using this one can relate the scalar current to the form factors, resulting in a new way to extract form factors via $\langle V_0 \rangle$ and $\langle S \rangle$:

$$f_0^{(s)}(q^2) = \frac{\delta m}{M_{B_{(s)}}^2 - M_{D_{(s)}}^2} \langle S \rangle, \quad (5.29)$$

$$f_+^{(s)}(q^2) = \frac{1}{2M_{B_{(s)}}} \frac{(M_{B_{(s)}} - E_{D_{(s)}}) \delta m \langle S \rangle - q^2 \langle V_0 \rangle}{\mathbf{p}_{D_{(s)}}^2}. \quad (5.30)$$

Care must be taken in choosing what masses to use in δm . One may want to use the bare valence charm and bottom masses, however these belong to different regularization schemes (HISQ and NRQCD), so taking their difference is not well defined. The solution is to use instead $\delta m = m_{b0}^{\text{val}} \times (1 - m_c/m_b)$, where m_c/m_b is a regularization-independent quantity computed to be $m_c/m_b = 1/4.51(4)$ in [115].

Results from adopting this alternative approach (in the $B \rightarrow D$ case) is shown in Fig. 5.7. One immediately notices that $f_+(q^2)$ results are diverging in the $|\mathbf{p}_D| \rightarrow 0$ limit. Why this occurs can be seen by inspecting Eq. (5.30). For f_+ to remain finite, the difference between currents on the numerator must tend to zero at the same rate as \mathbf{p}_D^2 . Our results for the currents $\langle S \rangle$ and $\langle V_0 \rangle$ are not precise enough to produce the delicate cancellation required to accurately determine f_+ in the high q^2 region.

A quick summary of the situation. We have 3 currents, $\langle S \rangle$, $\langle V_0 \rangle$, and $\langle V_k \rangle$. We require input from two of these currents in order to determine the form factors. Large contributions to $\langle V_k \rangle$ may be being ignored, so at the moment we do not consider $\langle V_k \rangle$ a 'trustworthy' estimation of the continuum spatial vector current. Using only $\langle S \rangle$ and $\langle V_0 \rangle$ leads to a divergence of $f_+(q^2)$ as $q^2 \rightarrow q_{\text{max}}^2$, so is also

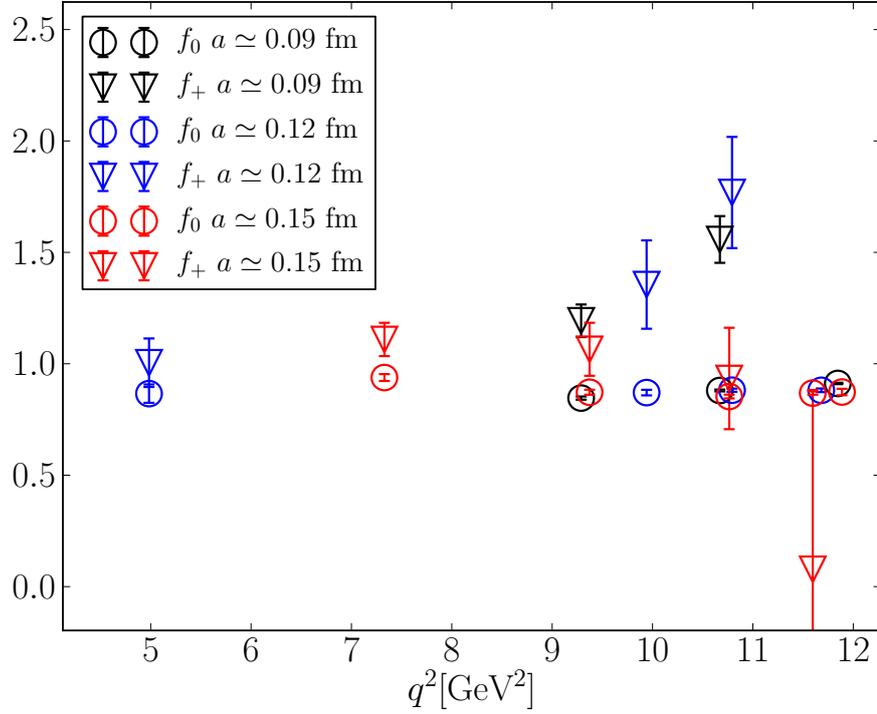


Figure 5.7: $B \rightarrow D$ form factors extracted from $\langle V_0 \rangle$ and $\langle S \rangle$. The statistical errors here are not well under control, a continuum and kinematic extrapolation is not possible using these form factors.

untenable. In the next section, we show an approach we attempted to finding new normalizations of these three currents such that all three can be 'trusted' as good approximations to the continuum current. One could then in principle use these trustworthy $\langle V_0 \rangle$ and $\langle V_k \rangle$ currents to extract the form factors.

5.4 Non-Perturbative Renormalization Using $B_c \rightarrow \eta_c$ Data

Here we define non-perturbative normalization constants Z_J for the NRQCD-HISQ currents via

$$\begin{aligned} J &= (1 + z_0^J \alpha_s)(J^{(0)} + J^{(1)}) + \mathcal{O}(\alpha_s^2, (\Lambda_{\text{QCD}}/m_b)^2, (\mathbf{p}/m_b)^2, \alpha_s \Lambda_{\text{QCD}}/m_b, \alpha_s \mathbf{p}/m_b) \\ &\equiv Z_J(1 + z_0^J \alpha_s)(J^{(0)} + J^{(1)}), \\ Z_J &= 1 + \mathcal{O}(\alpha_s^2, (\Lambda_{\text{QCD}}/m_b)^2, (\mathbf{p}/m_b)^2, \alpha_s \Lambda_{\text{QCD}}/m_b, \alpha_s \mathbf{p}/m_b). \end{aligned} \quad (5.31)$$

The Z_J factor compensates for the truncation of the series. One can imagine fixing Z_J by demanding some property of J . One may be concerned that Z_J is dependent on the spatial momenta in the current $Z_J = Z_J(\mathbf{p})$, however, we will see below that this variation is a negligible effect.

In this section, we outline an approach used to determine Z_S , Z_{V_0} and Z_{V_k} . In this process we use NRQCD data from another HPQCD project of determining form factors for $B_c \rightarrow \eta_c \ell \nu$ decays [6]. A schematic of how this is achieved is given in fig. 5.8. I thank Brian Colquhoun for supplying correlation functions from their calculation. The current in this calculation is the same as in the $B_{(s)} \rightarrow D_{(s)} \ell \nu$ calculation, so normalizations determined using this data can in principle be applied to the $B_{(s)} \rightarrow D_{(s)} \ell \nu$ calculation.

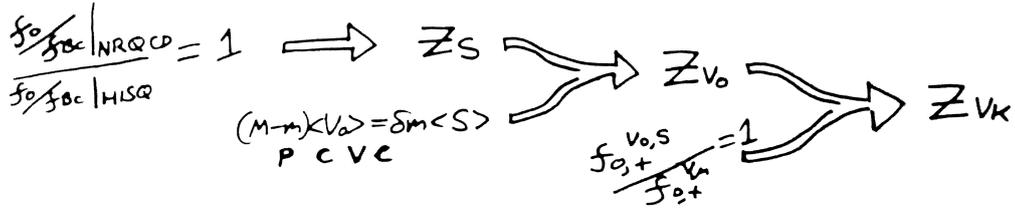


Figure 5.8: A schematic of the chain of steps towards normalizing the scalar, temporal vector and spatial vector NRQCD-HISQ currents. Details given in the next three sections; 5.4.1, 5.4.2 and 5.4.3.

We used the $B_c \rightarrow \eta_c \ell \nu$ data here since

- $B_c \rightarrow \eta_c$ correlators have much smaller statistical errors. This is because of the lack of the s spectator quark, degradation of the signal/noise ratio is less severe (see Sec. 4.2.2).

- A heavy-HISQ determination of $B_c \rightarrow \eta_c$ form factors was also available from their project [6]. This comes in useful in our approach to the normalization.

All analysis below is performed using data on the fine ensemble (set 2). In principle, it could be repeated for any other ensemble on which correlators for any $b \rightarrow c$ transition is available.

5.4.1 Z_S

We have available to us lattice results for $f_0^{B_c \rightarrow \eta_c}(q^2) \equiv f_0(q^2)$ for a number of q^2 values spanning the entire q^2 range (including q_{\max}^2 and $q^2 = 0$) from the NRQCD-HISQ S current on the fine ensemble. We also have a determination of f_{B_c} also from NRQCD-HISQ lattice currents on the same ensemble. We denote these finite- a lattice results as $\hat{f}_0(q^2)$, \hat{f}_{B_c} . We also have a continuum-extrapolated heavy-HISQ result for $f_0(q^2)/f_{B_c}$, for q_{\max}^2 and $q^2 = 0$. This is given in the form of this ratio since discretization effects largely cancel in this ratio improving the continuum extrapolation (see Chapters 6 and 7). The results are

$$\frac{f_0(q_{\max}^2)}{f_{B_c}} = 2.104(36), \quad \frac{f_0(0)}{f_{B_c}} = 1.288(42). \quad (5.32)$$

Since $f_0 \propto \langle S \rangle$, we can assert that $f_0 = Z_S \hat{f}_0$, i.e., the continuum f_0 contains the normalization that \hat{f}_0 is missing. Similarly for \hat{f}_{B_c} and Z_{A_0} . Hence by demanding that the NRQCD-HISQ finite- a results match the continuum heavy-HISQ results, we can find, for example

$$\left. \frac{Z_S}{Z_{A_0}} \right|_{q_{\max}^2} = \frac{f_0(q_{\max}^2)/f_{B_c}}{\hat{f}_0(q_{\max}^2)/\hat{f}_{B_c}} = 0.995(15). \quad (5.33)$$

As a test to see if Z_S varies with \mathbf{p} , we can compare this result to an analogous approach at $q^2 = 0$;

$$\left. \frac{Z_S}{Z_{A_0}} \right|_{q^2=0} = \frac{f_0(0)/f_{B_c}}{\hat{f}_0(0)/\hat{f}_{B_c}} = 0.962(33). \quad (5.34)$$

A similar approach cannot be applied for Z_{V_0} or Z_{V_k} , since f_0 has a complicated relationship to both $\langle V_0 \rangle$ and $\langle V_k \rangle$ that varies with q^2 , so it is not clear how one attribute discrepancies between \hat{f}_0/\hat{f}_{B_c} and f_0/f_{B_c} to Z_{V_0} and Z_{V_k} .

The comparison of these two ratios at q_{\max}^2 and $q^2 = 0$ show that any variation is small in comparison to statistical errors. We can absorb the variation in \mathbf{p} into a sub-leading term in the scalar current by demanding that $Z_S/Z_{A_0}|_{q_{\max}^2} = Z_S/Z_{A_0}|_{q^2=0}$.

Redefine the scalar current according to:

$$S = Z_S \left[(1 + z_0^S \alpha_s) (V_0^{(0)} - V_0^{(1)}) + \alpha_s z_2^S V_0^{(2)} \right]. \quad (5.35)$$

We can determine z_2^S by demanding that Z_S does not vary between q_{\max}^2 and $q^2 = 0$. This is equivalent to the $V_0^{(2)}$ term absorbing all of the variation in the normalization on \mathbf{p} , which one would expect since this current is proportional to the spatial momentum in the c -quark.

By defining \hat{f}_2 to be \hat{f}_0 but with $(1 + \alpha_s z_0^S)(V_0^{(0)} - V_0^{(1)})$ replaced with $\alpha_s V_0^{(2)}$, we can write:

$$\frac{Z_{A_0}}{Z_S} = \frac{(\hat{f}_0 + z_2^S \hat{f}_2)/\hat{f}_{B_c}}{f_0/f_{B_c}} \equiv \left(\frac{Z_{A_0}}{Z_S} \right)^{(0,1)} + z_2^S \left(\frac{Z_{A_0}}{Z_S} \right)^{(2)}. \quad (5.36)$$

With this further definition, and demanding that $Z_{A_0}/Z_S|_{q_{\max}^2} = Z_{A_0}/Z_S|_{q^2=0}$, we end up with

$$z_2^S = \frac{\left(\frac{Z_{A_0}}{Z_S} \right)^{(0,1)}|_{q^2=0} - \left(\frac{Z_{A_0}}{Z_S} \right)^{(0,1)}|_{q_{\max}^2}}{\left(\frac{Z_{A_0}}{Z_S} \right)^{(2)}|_{q_{\max}^2} - \left(\frac{Z_{A_0}}{Z_S} \right)^{(2)}|_{q^2=0}} = -1.1(1.5). \quad (5.37)$$

Now that we are able to include $V_0^{(2)}$ in the scalar current, we can consider the scalar current normalized up to $\mathcal{O}(\alpha_s^2, (\Lambda_{\text{QCD}}/m_b)^2, (\mathbf{p}/m_b)^2, \alpha_s \Lambda_{\text{QCD}}/m_b)$. We can use

$$S = (1 + z_0^S \alpha_s) (V_0^{(0)} - V_0^{(1)}) + z_2^S \alpha_s V_0^{(2)} + \mathcal{O}(\alpha_s^2, (\Lambda_{\text{QCD}}/m_b)^2, (\mathbf{p}/m_b)^2, \alpha_s \Lambda_{\text{QCD}}/m_b) \quad (5.38)$$

(Yes, we have not strictly determined Z_S in this case, but we will for the vector currents). The next steps are essentially to match results using the vector currents to this newly normalized scalar current, meaning the vector currents will be normalized up to $\mathcal{O}(\alpha_s^2, (\Lambda_{\text{QCD}}/m_b)^2, (\mathbf{p}/m_b)^2, \alpha_s \Lambda_{\text{QCD}}/m_b)$, hence we then will have theoretically accounted for the large subleading pieces in the spatial vector current, since we have accounted for $\alpha_s \mathbf{p}/m_b$ order terms.

5.4.2 Z_{V_0}

We normalized the temporal vector current via its PCVC relation with the now 'correctly' normalized scalar current.

This also has the benefit of partially removing the $f_+(q^2)$ divergence as $\mathbf{p}_{\eta_c}^2 \rightarrow 0$ when extracted from $\langle V_0 \rangle$ and $\langle S \rangle$. One can see this by inspecting the expression for $f_+(q^2)$ in terms of $\langle S \rangle$ and $\langle V_0 \rangle$ (Eq. (5.30)). As $\mathbf{p}_{\eta_c}^2 \rightarrow 0$ the numerator becomes

proportional to $\delta m \langle S \rangle - (M_{B_c} - M_{\eta_c}) \langle V_0 \rangle$, which vanishes when the PCVC relation is satisfied. So one would expect if we renormalize one of the currents so the Ward identity is satisfied (at q_{\max}^2), this divergence should be removed or at least reduced.

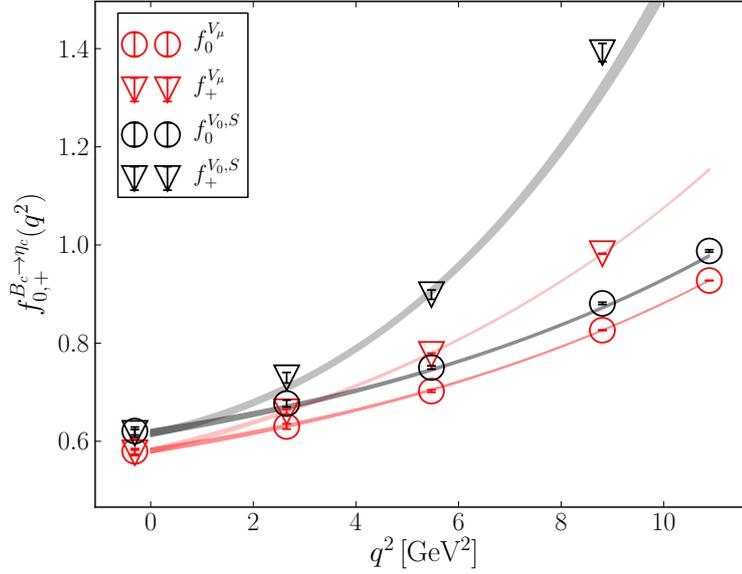


Figure 5.9: Comparison of form factors from $\langle V_0 \rangle$, $\langle S \rangle$ and $\langle V_0 \rangle$, $\langle V_k \rangle$, no additional normalizations. The bands show the results of fitting the data to the BCL parameterization (5.23), where coefficients $a_n^{0,+}$ are fit parameters. These bands are intended simply to guide the eye.

Hence we should normalize the temporal vector current using the already normalized scalar density. So to satisfy the PCVC, Multiply $\langle V_0 \rangle$ by

$$Z_{V_0} = \frac{m_b - m_c}{M_{B_c} - M_{\eta_c}} \frac{\langle S \rangle}{\langle V_0 \rangle} \Big|_{q_{\max}^2} = 1.0661(36). \quad (5.39)$$

This seems to deal with the f_+ divergence. Fig. 5.9 compares form factors determined using $\langle V_0 \rangle$, $\langle S \rangle$, and $\langle V_0 \rangle$, $\langle V_k \rangle$, *before* imposing Z_{V_0} , and Fig. 5.10 shows the same *after* $\langle V_0 \rangle$ has been multiplied by Z_{V_0} .

Unfortunately, the same technique does not solve the diverging f_+ issue in the $B_{(s)} \rightarrow D_{(s)}$ case, the statistics are not as good so the divergence is too severe.

5.4.3 Z_{V_k}

We can determine a Z_{V_k} by demanding that $f_{0,+}^{V_0,S}/f_{0,+}^{V_\mu} = 1$, with the knowledge that $\langle V_0 \rangle$, $\langle S \rangle$ require no further normalization. This can be done with both f_+ and

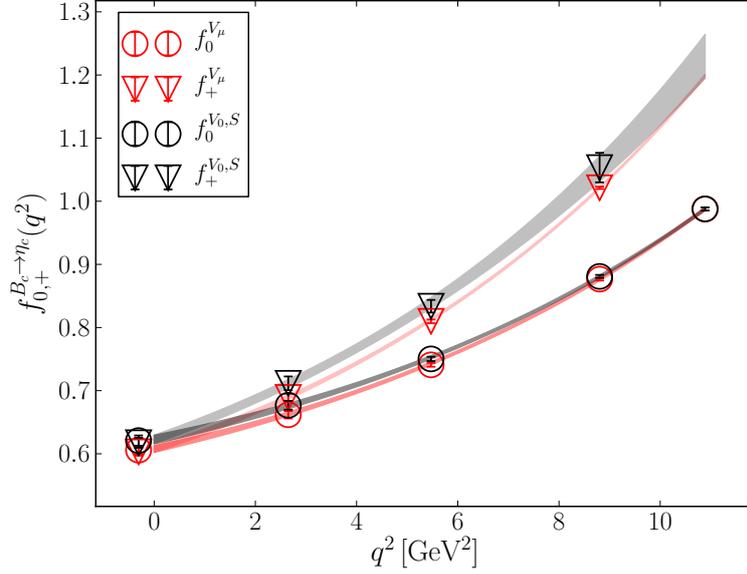


Figure 5.10: Comparison of form factors from $\langle V_0 \rangle, \langle S \rangle$ and $\langle V_0 \rangle, \langle V_k \rangle, \langle V_0 \rangle$ is multiplied by Z_{V_0} given in (5.39). The bands show the results of fitting the data to the BCL parameterization (5.23), where coefficients $a_n^{0,+}$ are fit parameters.

f_0 , at any q^2 ;

$$Z_{V_k} = 1 + \frac{f_+^{V_0,S} - f_+^{V_\mu}}{R_{+k} \langle V_k \rangle} \quad (5.40)$$

$$= 1 + \frac{f_0^{V_0,S} - f_0^{V_\mu}}{R_{0k} \langle V_k \rangle}, \quad (5.41)$$

where the R 's are some kinematic gunk: $R_{+k} = (M_{B_c} - E_{\eta_c})/2M_{B_c} \mathbf{p}_{\eta_c}/\sqrt{3}$, $R_{0k} = R_{+k} - (M_{B_c}^2 - M_{\eta_c}^2)(M_{B_c} - E_{\eta_c})/2M_{B_c} \mathbf{p}_{\eta_c}/q^2 \sqrt{3}$. The results are shown in Fig. 5.11. The fact that these are not varying by a statistically significant extent in q^2 implies that the spatial vector normalization does not vary strongly in \mathbf{p}_{η_c} . Hence, since these are all estimates of the same value, we can average over them to get

$$Z_{V_k} = 1.070(36). \quad (5.42)$$

When this normalization is given to $\langle V_k \rangle$, the $B_c \rightarrow \eta_c$ form factors from the two methods become consistent, see Fig. 5.12.

One could imagine using these normalizations Z_{V_0} and Z_{V_k} in the $B_{(s)} \rightarrow D_{(s)}$ calculation. The errors of this normalization are around 4%, which would push the

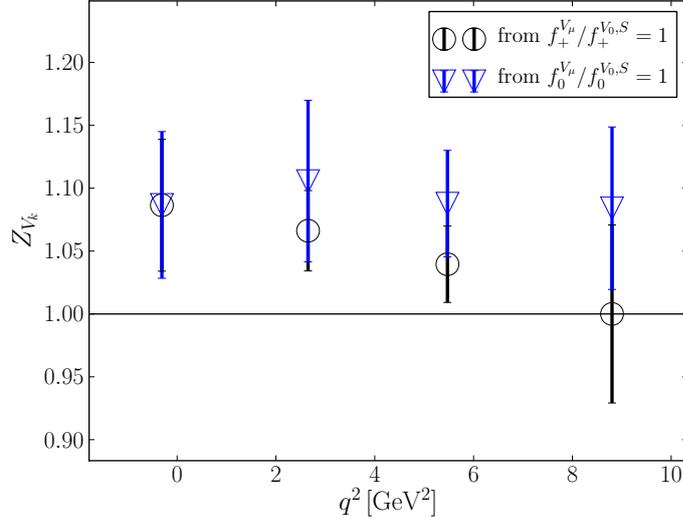


Figure 5.11: Z_{V_k} from constraining form factors to be the same from $\langle V_0 \rangle, \langle S \rangle$ and $\langle V_0 \rangle, \langle V_k \rangle$

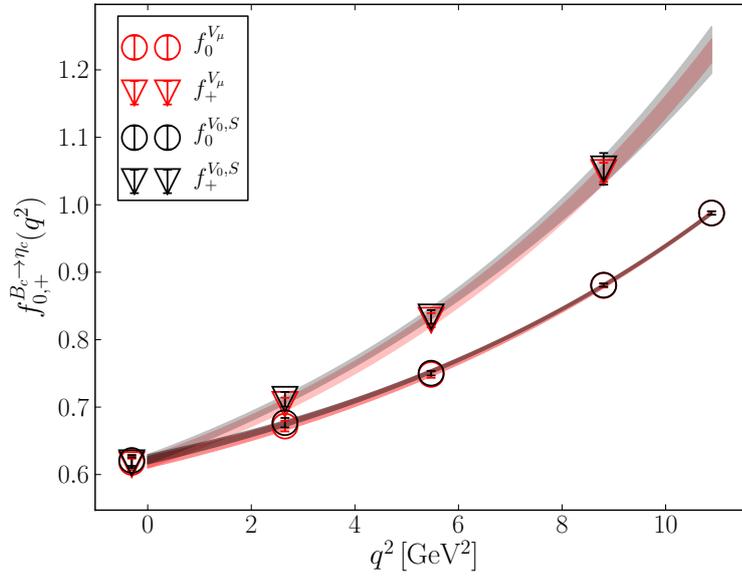


Figure 5.12: Comparison of form factors from $\langle V_0 \rangle, \langle S \rangle$ and $\langle V_0 \rangle, \langle V_k \rangle$, with $\langle V_0 \rangle$ normalized with Z_{V_0} and $\langle V_k \rangle$ normalized with Z_{V_k} . The bands show the results of fitting the data to the BCL parameterization (5.23), where coefficients $a_n^{0,+}$ are fit parameters.

final results of the $B_{(s)} \rightarrow D_{(s)}$ study up to the 5 – 10% range. At present this level of precision is not competitive in comparison to other calculations of these quantities.

If one chooses to trust the normalization of the scalar current, this analysis shows that (on the fine ensemble) the extra normalization required for the vector currents beyond what is usually implemented is a large effect ($Z_{V_0} - 1 \sim Z_{V_k} - 1 \sim 7\%$). This would suggest that the currently used truncation of the NRQCD-HISQ vector current is missing $\sim 7\%$ from the neglected $\mathcal{O}(1/m_b)$ terms.

5.5 Conclusion of NRQCD Work

As mentioned in the introduction - the main message of this thesis is that using NRQCD (namely NRQCD-HISQ currents) to compute form factors for $b \rightarrow c$ transitions, is far from optimal.

We rely here on truncations in many different interlocking series ($1/m_b, \alpha_s, \alpha_s/m_b, \dots$), which may leave out important information. The terms in the series we *do* have access to rely on perturbation theory via the matching factors. Some progress was made to normalize currents non-perturbatively in the above work, but our results fall short of what would be required to obtain precise continuum results.

I hope that in reading this chapter you experienced a similar feeling of confusion and anxiety to what I felt as I carried out this work. It was only after many months of grappling with the problems of these calculations that we decided to put them aside and attempt instead the heavy-HISQ approach from scratch. Reading the following two chapters, both dedicated to successful heavy-HISQ studies, will feel like a warm bath in comparison to the NRQCD experience. The heavy-HISQ approach is in contrast very elegant, contains far fewer assumptions, and results in cleaner signals.

$B_s \rightarrow D_s^* \ell \nu$ Axial Form Factor at Zero Recoil from Heavy-HISQ

This chapter concerns the simpler of our two heavy-HISQ studies, the calculation of the $B_s \rightarrow D_s^* \ell \nu$ axial form factor at zero recoil, $h_{A_1}^s(1)$. We give this quantity the superscript s to differentiate it from the quantity more commonly referred to as $h_{A_1}(1)$, the zero recoil axial form factor for $B \rightarrow D^* \ell \nu$ decays.

I will briefly review the definition of this form factor (at zero recoil) for ease of reading. The differential decay rate for the $\bar{B}_s^0 \rightarrow D_s^{*+} \ell^- \bar{\nu}_\ell$ decay is given in the SM by

$$\begin{aligned} \frac{d\Gamma}{dw}(\bar{B}_s^0 \rightarrow D_s^{*+} \ell^- \bar{\nu}_\ell) &= \frac{G_F^2 M_{D_s^*}^3 |\bar{\eta}_{\text{EW}} V_{cb}|^2}{4\pi^3} \\ &\times (M_{B_s}^2 - M_{D_s^*}^2) \sqrt{w^2 - 1} \chi(w) |\mathcal{F}^{B_s \rightarrow D_s^*}(w)|^2. \end{aligned} \quad (6.1)$$

where $w = v_{B_s} \cdot v_{D_s^*}$, $v_M = p_M/M_M$ is the 4-velocity of an M -meson, and $\chi(w)$ is a known function of w (see for example appendix G of [33]). $\bar{\eta}_{\text{EW}}$ accounts for electroweak corrections due to diagrams where photons or Z s are exchanged in addition to a W^- , as well as the Coulomb attraction of the final-state charged particles [29–31]. The differential decay rate for the $B_s^0 \rightarrow D_s^{*-} \ell^+ \bar{\nu}_\ell$ decay is identical.

The form factor $\mathcal{F}^{B_s \rightarrow D_s^*}(w)$ is a linear combination of hadronic form factors that parameterize the vector and axial-vector matrix elements between initial and final state hadrons. At zero recoil ($w = 1$), the vector matrix element vanishes, the axial-vector element simplifies to

$$\langle D_s^*(\epsilon) | A_\mu | B_s \rangle = 2 \sqrt{M_{B_s} M_{D_s^*}} h_{A_1}^s(1) \epsilon_\mu^*, \quad (6.2)$$

and $\mathcal{F}^{B_s \rightarrow D_s^*}(w)$ reduces to

$$\mathcal{F}^{B_s \rightarrow D_s^*}(1) = h_{A_1}^s(1). \quad (6.3)$$

Our goal is to compute $h_{A_1}^s(1)$.

All we need to do this is the matrix element $\langle D_s^*(\epsilon) | A_\mu | B_s \rangle$ with both the B_s and D_s^* at rest, with the D_s^* polarization ϵ in the same direction as the axial current.

6.1 Motivation

$B \rightarrow D^* \ell \nu$ decays supply one of the three methods used for precisely determining the CKM element $|V_{cb}|$ [38, 42, 156–171]. Measurements of branching fractions are extrapolated through q^2 to the zero recoil point to deduce $|h_{A_1}(1)V_{cb}|$, since $h_{A_1}(1)$ is the only form factor contributing at zero recoil. Then an SM determination of $h_{A_1}(1)$ (via Lattice QCD [33, 38]) can be divided out to infer $|V_{cb}|$.

A similar process that could also be used to determine $|V_{cb}|$, and test the SM, is $\bar{B}_s \rightarrow D_s^* \ell \bar{\nu}_\ell$. There is at time of writing no published measurements of this decay, but it is feasible to measure such a decay at a detector like LHCb. This decay is also attractive from the Lattice QCD side. The absence of valence light quarks means lattice results have smaller statistical errors, are less computationally expensive, a simpler chiral extrapolation to the physical light mass, and negligible finite volume effects. This makes the $\bar{B}_s \rightarrow D_s^* \ell \bar{\nu}_\ell$ both a useful test bed for lattice techniques (that may be later used to study $\bar{B} \rightarrow D^* \ell \bar{\nu}_\ell$ decays), and a key decay for future $|V_{cb}|$ determinations and tests of the SM.

Chiral symmetry implies that form factors for decays such as $B_s \rightarrow D_s^*$ and $B \rightarrow D^*$ are insensitive to the mass of the spectator quark, implying that form factors for these two decays are approximately equal [172]. This was seen in the recent lattice calculation [33] that found $h_{A_1}(1)/h_{A_1}^s(1) = 1.013(14)_{\text{stat}}(17)_{\text{sys}}$. We can then expect to learn about $B \rightarrow D^*$ by studying $B_s \rightarrow D_s^*$. We perform a further test of this claim that $B \rightarrow D^* \sim B_s \rightarrow D_s^*$, in the context of our formalism, in this study.

Lattice calculations of the $B_{(s)} \rightarrow D_{(s)}^*$ form factors at zero recoil have so far been performed by two collaborations. The Fermilab Lattice collaboration produced $h_{A_1}(1)$ in [38]. HPQCD computed both $h_{A_1}(1)$ and $h_{A_1}^s(1)$ in [33]. The RBC/UKQCD [173] and LANL-SWME [174] collaborations are also working towards lattice determinations of these form factors.

The presence of heavy quarks is a large consideration in designing a lattice calculation (as discussed in Sec. 3.3). A b quark introduces discretization effects of size $(am_b)^n$ where n is a positive integer dependent on the choice of action. To avoid such potentially large discretization effects, most lattice studies (including all of those mentioned in the previous paragraph), use some EFT approach for simulating heavy quarks. The Fermilab Lattice, RBC/UKQCD, and LANL-SWME calculations all used some variation of the Fermilab action [175–177] to simulate c and b quarks. The HPQCD calculation used the NRQCD action [95] for b quarks.

To relate the results from these approaches to full continuum QCD, each of the above studies requires perturbative matching of lattice currents to continuum QCD. The matching has only been performed to 1-loop, leading to each having matching errors as a key uncertainty. The use of NRQCD-HISQ currents in the HPQCD calculation brings in matching errors of $\mathcal{O}(\alpha_s^2, \alpha_s \Lambda_{\text{QCD}}/m_b, (\Lambda_{\text{QCD}}/m_b)^2)$. It is difficult to estimate the size of matching errors in lattice NRQCD, so to be conservative a large matching error was assigned to the result. This error contributes $\sim 80\%$ of the full error budget. The use of the Fermilab action in the Fermilab Lattice calculation leads to $\mathcal{O}(\alpha_s^2)$ errors. They avoid this issue to a large extent by analysing only ratios of correlation functions, however, the matching still contributes $\sim 30\%$ to the final error.

In this chapter, we report details and results of the first calculation of the $B_s \rightarrow D_s^*$ form factor at zero recoil using an approach free of perturbative matching.

Since the $B_s \rightarrow D_s$ form factor is approximately equal to the $B \rightarrow D$ form factor, and our results are non-perturbatively renormalised, this calculation can be seen as a check of the normalisation of the Fermilab Lattice and HPQCD determinations of $h_{A_1}(1)$ that contributed to $|V_{cb}|_{\text{excl}}$ (see Sec. 2.2.3).

Using the heavy-HISQ approach has the added benefit of elucidating the dependence of form factors on heavy quark masses, meaning we can test expectations from Heavy Quark Effective Theory (HQET). In this study, we produce an estimate of the HQET low energy constants $l_{V,A,P}$ associated with the $B_s \rightarrow D_s^*$ form factor at zero recoil.

6.2 Calculation Details

6.2.1 Lattice Setup

We used the MILC gluon field configurations detailed in Sec. 4.1.1 [7,8]. We used sets 2-5 in Table 4.1, i.e., the fine, fine-physical, superfine and ultrafine ensembles. Table 6.1 gives the valence quark masses we used in the generation of quark propagators. In three of the four ensembles (fine, superfine and ultrafine), the bare light mass is set to $m_{l0}/m_{s0} = 0.2$. The fact that the m_{l0} value is unphysically high is expected to have a small effect on $h_{A_1}^s(1)$, due to the lack of valence light quarks, and previous experience of the dependence of $h_{A_1}^s(1)$ on m_{l0} [33]. The small effect due to the unphysical m_{l0} is quantified by including the fine-physical ensemble with physical m_{l0} , and corrected for.

set	name	am_{s0}^{val}	am_{c0}^{val}	am_{h0}^{val}	$n_{\text{cfg}} \times n_{\text{src}}$	T/a
2	fine	0.0376	0.45	0.5, 0.65, 0.8	938×8	14,17,20
3	fine-physical	0.036	0.433	0.5, 0.8	284×4	14,17,20
4	superfine	0.0234	0.274	0.427, 0.525, 0.65, 0.8	250×8	22,25,28
5	ultrafine	0.0165	0.194	0.5, 0.65, 0.8	249×4	31,36,41

Table 6.1: Parameters relevant to our calculation. Columns 3 and 4 give the s and c valence quark masses, these values were tuned in [139] to reproduce the correct η_s and η_c masses. We used a number of heavy quark masses to assist the extrapolation to the physical b mass, given in column 5. Column 6 gives the number of gauge configurations (n_{cfg}) and the number of t_0 choices (n_{src}) used. Column 7 gives the temporal separations between B_s source and D_s^* sink, T/a , of the 3-point correlation functions computed on each ensemble.

We used a number of different masses for the valence heavy quark. This is in order to resolve the dependence of $h_{A_1}^s(1)$ on the heavy mass so that extrapolation to $m_h = m_b$ can be performed. By varying the heavy mass both within ensembles and between ensembles, we can resolve both the discretization effects that grow with large ($am_{h0}^{\text{val}} \lesssim 1$) masses and the physical dependence of the continuum form factor on m_h .

A considerable benefit to using unphysically light heavy quarks is that it reduces the signal/noise degradation in the correlation functions in comparison to using the physical b mass (as in e.g. the NRQCD approach). When using NRQCD, the large noise due to the heavy b -quark necessitated the computation of many correlation functions with different smeared operators in order to boost statistics. This is not necessary in the heavy-HISQ setting, we used only local creation/annihilation operators.

As detailed in Sec. 4.1.3, staggered correlation functions are built by a combination of staggered propagators $g(x, y)$ and staggered phases. In this calculation we only need local (non-point-split) operators, this is an advantage since point-split operators lead to correlation functions noisier than local operators.

We computed a number of correlation functions on each ensemble. To generate these correlators we used random wall sources, and used extended sources for the 3-point correlators, as described in Sec. 4.1.3. First, we computed 2-point correlation

functions between zero-momentum eigenstates, objects of the form

$$C_M(t) = \langle \Phi_M(t) \Phi_M^\dagger(0) \rangle, \quad (6.4)$$

$$\Phi_M(t) = \sum_{\mathbf{x}} \bar{q}(\mathbf{x}, t) \Gamma q'(\mathbf{x}, t),$$

where $\langle \rangle$ represents a functional integral, q, q' are valence quark fields of the flavours the M meson is charged under, and Γ is the spin-taste structure of M . I set $t_0 = 0$ here for notational simplicity. We computed these for all t values, i.e. $0 \leq t \leq T_{\text{lat}}$.

We computed correlation functions for a heavy-strange pseudoscalar, H_s , with spin-taste structure $(\gamma_5 \otimes \gamma_5)$. In terms of staggered propagators, this takes the form

$$C_{H_s}(t) = \sum_{\mathbf{x}, \mathbf{y}} \left\langle \text{Tr}_c \left[g_h(x, y) g_s^\dagger(x, y) \right] \right\rangle, \quad (6.5)$$

where $g_q(x, y)$ is a staggered propagator for flavour q , and the trace is over color. Here $x_0 = 0$ and $y_0 = t$. We also computed correlators for a charm-strange vector meson D_s^* , with structure $(\gamma_\mu \otimes \gamma_\mu)$, using

$$C_{D_s^*}(t) = \sum_{\mathbf{x}, \mathbf{y}} (-1)^{x_\mu + y_\mu} \left\langle \text{Tr}_c \left[g_c(x, y) g_s^\dagger(x, y) \right] \right\rangle. \quad (6.6)$$

In order to non-perturbatively renormalise the axial vector current, we computed correlation functions for two heavy-charm mesons, denoted H_c and \hat{H}_c respectively. H_c has spin-taste structure $(\gamma_5 \otimes \gamma_5)$ and \hat{H}_c has structure $(\gamma_5 \gamma_0 \otimes \gamma_5 \gamma_0)$. H_c correlators are computed using (6.5) (with g_s replaced with g_c), while \hat{H}_c correlators are given by

$$C_{\hat{H}_c}(t) = \sum_{\mathbf{x}, \mathbf{y}} (-1)^{\bar{x}_0 + \bar{y}_0} \left\langle \text{Tr}_c \left[g_h(x, y) g_c^\dagger(x, y) \right] \right\rangle, \quad (6.7)$$

where we use the notation $\bar{z}_\mu = \sum_{\nu \neq \mu} z_\nu$. H_c and \hat{H}_c are referred to as goldstone and non-goldstone pseudoscalars respectively.

The heavy-mass extrapolation requires masses of η_h mesons, heavy-heavy pseudoscalars artificially forbidden to annihilate. To quantify mistuning of the charm and strange quark masses, we also require masses for η_c and η_s mesons, identical to η_h with h replaced c and s quarks respectively. We computed correlators for each of these, using a spin-taste $(\gamma_5 \otimes \gamma_5)$, taking the form of (6.5).

We then generate the 3-point correlation functions

$$C_3(t, T) = \sum_{\mathbf{y}} \langle \Phi_{D_s^*(c)}(T) A_\mu(\mathbf{y}, t) \Phi_{H_s}(0) \rangle, \quad (6.8)$$

$$A_\mu(\mathbf{y}, t) = \bar{c}(\mathbf{y}, t) \gamma_5 \gamma_\mu h(\mathbf{y}, t).$$

In terms of the staggered formalism, the H_s source is given structure $(\gamma_5 \otimes \gamma_5)$, the D_s^* sink is given $(\gamma_\mu \otimes \gamma_\mu)$, and the current insertion $(\gamma_5 \gamma_\mu \otimes \gamma_5 \gamma_\mu)$. In terms of staggered propagators this is given by

$$C_3(t, T) = \sum_{\mathbf{x}, \mathbf{y}, \mathbf{z}} (-1)^{\bar{y}_\mu + \bar{z}_\mu} \left\langle \text{Tr}_c \left[g_h(x, y) g_c(y, z) g_s^\dagger(x, z) \right] \right\rangle, \quad (6.9)$$

where we fix $x_0 = 0$, $y_0 = t$ and $z_0 = T$. We computed these for all t values within $0 \leq t \leq T$, and 3 T values that vary between ensembles, given in Table 6.1.

In the $C_{D_s^*}$ and C_3 cases, dependant on a polarization μ , we computed the cases with $\mu = x, y, z$, and took the average over these.

6.2.2 Correlator Fits

We extracted current matrix elements from the generated correlation functions via simultaneous Bayesian fits as described in Sec. 4.2.1. We used fit forms given by Eq. (4.41) for 2-point and Eq. (5.18) for 3-point correlators. We set $N_{\text{exp}} = 5$ in each fit. We performed a single simultaneous fit containing each correlator computed ($H_s, D_s^*, \eta_h, \eta_c, \eta_s, H_c, \hat{H}_c$, and 3-point) for each ensemble. We also marginalized out the highest energy excited states (the $N_{\text{exp}} = 5$ states) in the interest of speeding up the fits.

The marginalization is implemented in the following way. The expected contribution to correlation functions from the $N_{\text{exp}} = 5$ state is estimated using the prior distributions of the associated fit parameters. The contribution is then numerically added to the lattice correlation functions being fit, and the fit function can be truncated at $N_{\text{exp}} = 4$.

We set Gaussian priors for the parameters J_{jk} and log-normal priors for all other parameters. The prior values we chose are summarized below.

$$\begin{aligned} \log(\tilde{E}_0^M) &= \log((am_{q0} + am_{q'0} + a\Lambda_{\text{QCD}}) \pm 2a\Lambda_{\text{QCD}}), \\ \log(\tilde{E}_0^{M,o}) &= \log((am_{q0} + am_{q'0} + 2a\Lambda_{\text{QCD}}) \pm 2a\Lambda_{\text{QCD}}), \\ \log(\tilde{E}_i^{M(o)} - \tilde{E}_{i-1}^{M(o)}) &= \log(2a\Lambda_{\text{QCD}} \pm a\Lambda_{\text{QCD}}), \quad i > 0, \\ \log(\tilde{a}_0^{M(o)}) &= \text{Empirical Bayes}, \\ \log(\tilde{a}_i^M) &= -1.20(67), \quad i > 0, \\ \log(\tilde{a}_i^{M,o}) &= -3.0(2.0), \quad i > 0, \\ J_{jk} &= 0 \pm 1 \text{ except for } J_{00}^{nn} = 1 \pm 0.6 \end{aligned} \quad (6.10)$$

Ground state energies E_0^M were given priors of $(am_{q0} + am_{q'0} + a\Lambda_{\text{QCD}}) \pm 2a\Lambda_{\text{QCD}}$, where $m_{q0, q'0}$ are the masses of the flavours the meson M is charged under, and Λ_{QCD} is the confinement scale, which we set to 0.5 GeV. For $q = h$ or c , this corresponds to the leading order HQET expression for a heavy meson mass. In the η_s case, the prior becomes approximately $2am_{s0} + a\Lambda_{\text{QCD}} \simeq a\Lambda_{\text{QCD}}$, which one would expect. Ground-state energies of oscillating states, $E_0^{M,o}$, are given priors of $(am_{q0} + am_{q'0} + 2a\Lambda_{\text{QCD}}) \pm 2\Lambda_{\text{QCD}}$. Excited state energy differences, $E_i^{M(o)} - E_{i-i}^{M(o)}$, $i > 0$ are given prior values $2a\Lambda_{\text{QCD}} \pm a\Lambda_{\text{QCD}}$. Priors for ground state amplitudes $a_0^{M(o)}$, are set according to an empirical-Bayes approach, plots of the effective amplitudes of the correlation functions (defined in Eq. 6.15 below) are inspected to deduce reasonable priors. The resulting priors always have a variance at least 10 times that of the final result. The excited state log-amplitudes, $\log(a_i^{M(o)})$, $i > 0$ are given priors of $-1.20(67)$ for non-oscillating states, and $-3.0(2.0)$ for oscillating states. The ground-state non-oscillating to non-oscillating 3-point parameter, J_{00}^{nn} is given a prior of 1 ± 0.6 , and the rest of the 3-point parameters J_{jk}^{nn} are given 0 ± 1 .

The current matrix element we require to find $h_{A_1}^s(1)$ is given by

$$\langle D_s^*(\hat{k}) | A_k | H_s \rangle |_{\text{lat}} = 2\sqrt{M_{H_s} M_{D_s^*}} J_{00}^{nn}. \quad (6.11)$$

We performed a number of tests on the fits to demonstrate the robustness of the fits to various hyperparameter choices. Results are given in Fig. 4.2.1. I will refer to these tests throughout the remainder of this section. In test #2 we loosened priors to test stability. We tested the effects of changing N_{exp} , to $N_{\text{exp}} = 6$ in test #3 and $N_{\text{exp}} = 4$ in test #4.

To ensure that truncating the sum at N_{exp} is a good approximation to the infinite sum containing all excited states, we only include data with $t \geq t_{\text{cut}}$ and $t \leq T_{\text{lat}} - t_{\text{cut}}$ in the 2-point case and $t \leq T - t_{\text{cut}}$ in the 3-point case. We can in principle use a different t_{cut} for every correlation function included in our fit, so must choose a set $\{t_{\text{cut}}^c\}$ (where c labels the correlator).

To ensure the optimal choice for the $\{t_{\text{cut}}^c\}$ set, we employ the `scikit-optimize` python package [178]. The process consists of defining a function f with an input of $\{t_{\text{cut}}^c\}$ and an output of some loss function f . Then, the minimum of f with respect to $\{t_{\text{cut}}^c\}$ is found via a Gaussian process. We used the loss function

$$f(\{t_{\text{cut}}^c\}) = -\log \text{GBF} + \theta (\chi^2 - N_{\text{dof}}) \rho \frac{\chi^2}{N_{\text{dof}}}. \quad (6.12)$$

GBF is the Gaussian Bayes factor corresponding to the comparison between the resulting model of the fit (the fit function with parameters set by the fit), and a

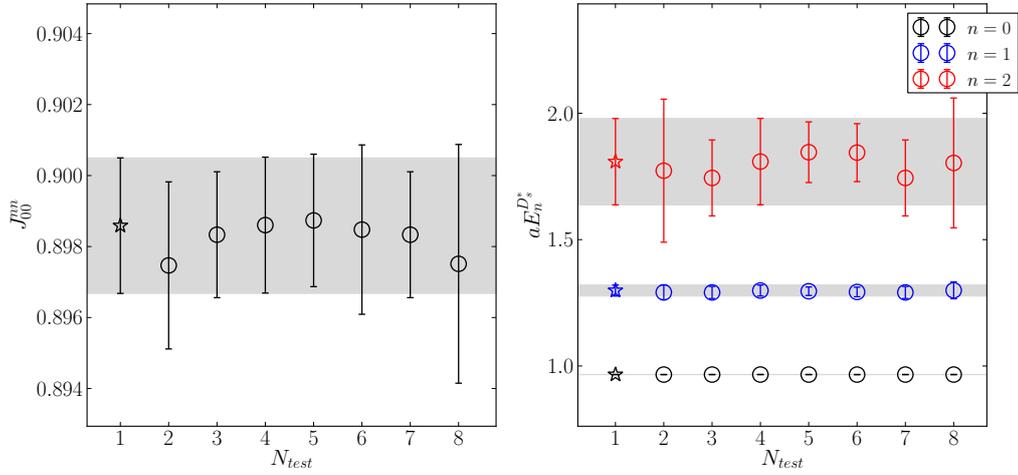


Figure 6.1: Tests of the correlator fits on the fine ensemble. The left panel shows J_{00}^{nn} at the heavy mass $am_{h0}^{\text{val}} = 0.5$. The errors shown here as statistical, estimated by taking the second derivative of the best-fit χ^2 with respect to the fit parameter in question. At $N_{\text{test}} = 1$ we give the final accepted result. $N_{\text{test}} = 2$ gives the results when all priors are broadened by 50%. $N_{\text{test}} = 3$ and 4 gives the results of setting $N_{\text{exp}} = 4$ and 6 respectively. $N_{\text{test}} = 5, 6$ gives the results of setting $t_{\text{cut}} = 2, 4$ respectively for all correlators. $N_{\text{test}} = 7$ gives the result without marginalising out the $n = 5$ excited state. $N_{\text{test}} = 8$ gives the result of moving the SVD cut from 10^{-3} to 10^{-2} .

random model (the fit function with randomly sampled parameters). The second term gives a strong punishment to fits with $\chi^2/N_{\text{dof}} > 1$. We set $\rho = 10^5$, in order to make the second term of comparable size of the first, which for typical fits we attempted had a magnitude of order 10^4 . The output of this process is shown in Table 6.2. A couple of more naive choices for $\{t_{\text{cut}}^c\}$ are given in tests #5 & #6.

An appropriate value for the svd cut is found by comparing estimates of eigenvalues of the data’s covariance matrix between different bootstrap samples of the data (see Sec. 2.7 of the **CorrFitter** documentation [146]). Typically the smallest eigenvalues are sensitive to taking new bootstrap copies, suggesting they are poorly estimated. A cut is placed such that any poorly estimated eigenvalues are replaced with more conservative (larger) values. The resulting svd cut varies between ensembles since it depends on the quality of the dataset, but are always of order 10^{-3} . For example, in fits to the fine ensemble correlators, we set the svd cut to exactly

set	name	H_s	D_s	H_c	\hat{H}_c	η_q	A_0
2	fine	2	2	5	5	3	1
3	fine-physical	4	4	8	8	3	2
4	superfine	4	4	6	6	3	1
5	ultrafine	2	8	4	4	2	1

Table 6.2: t_{cut} values used for each correlator on each ensemble. A_0 denotes the 3-point correlators. The rest are for 2-point correlators.

10^{-3} . We tested to see if this had any effect by also running the fit with svd cut 10^{-2} in test #8.

We can perform further sanity checks on the fits by plotting certain functions of the 2-point correlators. To obtain useful forms, first one can approximately flush out the oscillating states from correlators by performing a so-called *superaverage*, $C(t) \rightarrow [C(t) + C(t+a)]/2$. We perform a doubled and symmetric version of this operator on correlators to obtain

$$C(t) \rightarrow \tilde{C}(t) = \frac{1}{4}(C(t-a) + 2C(t) + C(t+a)). \quad (6.13)$$

We can check the non-oscillating ground-state energy of the correlator by looking at the large- t behaviour of

$$E_{\text{eff}}(t) = \log \left(\frac{\tilde{C}(t)}{\tilde{C}(t-a)} \right). \quad (6.14)$$

It is straightforward to show from plugging in the fit form for 2-point correlators (Eq. (4.36)) that in the large t (but $t < T_{\text{lat}}/2$) limit, this should tend towards the ground-state energy for the correlator. One can also construct a similar function for the amplitude:

$$a_{\text{eff}}(t) = \sqrt{\frac{2\tilde{C}(t)e^{E_{\text{eff}}(t)t}}{\cosh E_{\text{eff}}(t) - 1}}. \quad (6.15)$$

The $\tilde{C}(t)e^{E_{\text{eff}}(t)t}$ factor would produce the correct amplitude (in the large- t limit) in the absence of superaveraging, and the other factor corrects for the effect of the superaveraging. These functions, for various relevant correlators on the fine ensemble, are plotted in comparison with the full fit results in Fig. 6.2.

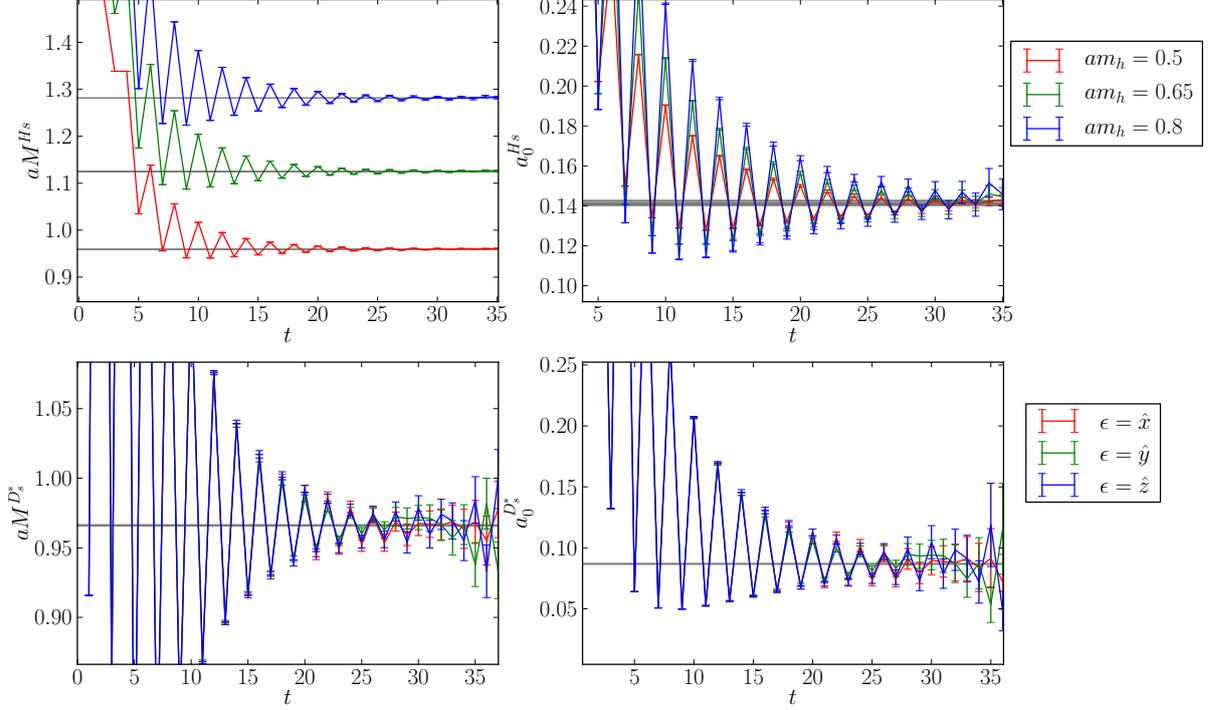


Figure 6.2: Effective energies and amplitudes for H_s and D_s^* correlators on the fine ensemble. The energies are obtained from Eq. (6.14), and amplitudes from Eq. (6.15). The grey bands give the results of the full multiexponential fit.

A similar approach can be applied to the 3-point correlators. The ratio $\tilde{C}_3(t, T) / \tilde{C}_{H_s}(t) \tilde{C}_{D_s^*}(T - t)$ approaches $J_{00}^{nn} / a_0^{H_s} a_0^{D_s^*}$ for $t \gg 0$ and $t \ll T$. This is illustrated in Fig. 6.3. From inspecting these figures for 2- and 3-point sanity tests, one can reassure themselves that the fits to the correlators are well behaved.

6.2.3 Normalization of the Axial Current

Conserved and partially conserved currents require no renormalization (see Sec. 2.3.2). However, the staggered conserved axial-vector current is not simply $(\gamma_5 \gamma_\mu \otimes \gamma_5 \gamma_\mu)$, it is a complicated linear combination of many local and point-split lattice currents. In this study we used only local axial vector currents, this simplifies the lattice calculation but creates the need for our resulting current matrix element to be multiplied by a matching factor Z_A to produce the appropriate continuum current. We found Z_A via a fully non-perturbative method [113, 179].

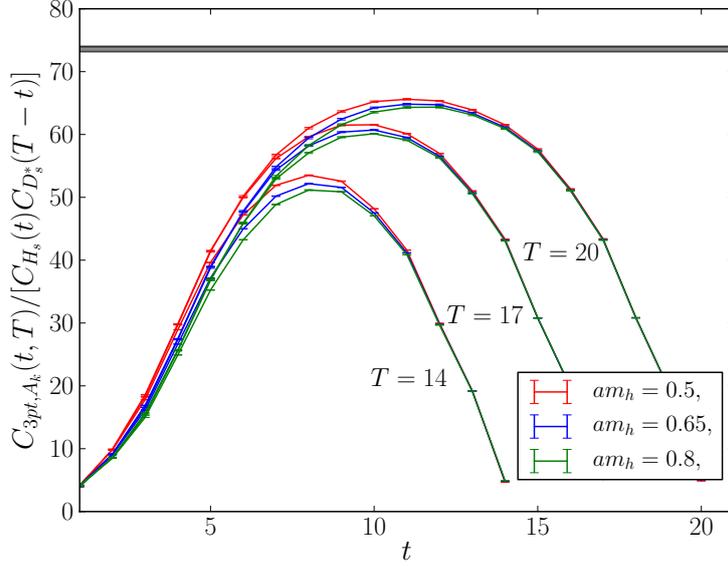


Figure 6.3: Sanity check for fits to the 3-point correlation functions. This ratio should approach J_{00}^{nn} for $t \gg 0$ and $t \ll T$. The grey bands show the result for J_{00}^{nn} from the full Bayesian simultaneous fit for each am_h (here they are all bunched up so appear as a single band). If the T values were larger, one can envision the data reaching a plateau at the same height as the grey band.

We leveraged the fact that the staggered local pseudoscalar current ($\gamma_5 \otimes \gamma_5$), multiplied by the sum of masses of quark flavours the current is charged under, is absolutely normalized. We extract from the 2-point H_c and \hat{H}_c correlators the decay amplitudes $\langle \Omega | \bar{c}(\gamma_5 \otimes \gamma_5) h | H_c \rangle \equiv \langle \Omega | P | H_c \rangle$ and $\langle \Omega | \bar{c}(\gamma_0 \gamma_5 \otimes \gamma_0 \gamma_5) h | \hat{H}_c \rangle = \langle \Omega | A_0 | \hat{H}_c \rangle$ from $a_0^{H_c}$ and $a_0^{\hat{H}_c}$. Then, the normalization for A_0 (common to that of spacial axial currents A_k) Z_A , is fixed by demanding that the partially conserved axial current relation holds:

$$(m_{h0}^{\text{val}} + m_{c0}^{\text{val}}) \langle \Omega | P | H_c \rangle |_{\text{lat}} = M_{\hat{H}_c} Z_A \langle \Omega | A_0 | \hat{H}_c \rangle |_{\text{lat}}. \quad (6.16)$$

The Z_A values found on each ensemble and am_{h0}^{val} are given in Table 7.3.

There is an ambiguity in which mass to use on the right-hand side of Eq. (6.16), we here use the non-goldstone mass $M_{\hat{H}_c}$, but one could just as well replace this with M_{H_c} . Using M_{H_c} here changes Z_A only by discretization effects, the effect on Z_A this causes never exceeds 0.0015% throughout the ensembles and heavy masses. The choice between these two definitions of Z_A has a negligible effect on our final result for $h_{A_1}^s(1)$.

Set	am_h^{val}	Z_A	Z_{disc}
2	0.5	1.03178(57)	0.99819
	0.65	1.03740(58)	0.99635
	0.8	1.04368(56)	0.99305
3	0.5	1.03184(47)	0.99829
	0.8	1.04390(39)	0.99315
4	0.427	1.0141(12)	0.99931
	0.525	1.0172(12)	0.99859
	0.65	1.0214(12)	0.99697
	0.8	1.0275(12)	0.99367
5	0.5	1.00896(44)	0.99889
	0.65	1.01363(49)	0.99704
	0.8	1.01968(55)	0.99375

Table 6.3: Normalization constants applied to the lattice axial vector current in Eq. (6.18). Z_A is found from Eq. (6.16) and Z_{disc} from Eq. (6.17).

We also remove tree-level mass-dependent discretization effects using a normalization constant derived in [108, 114]:

$$Z_{\text{disc}} = \sqrt{\tilde{C}_h \tilde{C}_c}, \quad (6.17)$$

$$\tilde{C}_q = \cosh am_{q,\text{tree}} \left(1 - \frac{1 + \epsilon_{\text{Naik}}}{2} \sinh^2 am_{q,\text{tree}} \right),$$

where ϵ_{Naik} is the Naik parameter in the HISQ action and $am_{q,\text{tree}}$ is the tree-level pole mass in HISQ defined in Eq. (3.55). The effect of Z_{disc} is very small, never exceeding 0.2%. Z_{disc} values on each ensemble for each am_{h0}^{val} are given in Table 7.3.

Combining these normalizations with the lattice current from the 3-point fits, we find a value for the form factor at a given heavy mass and lattice spacing:

$$h_{A_1}^s(1) = \frac{1}{3} \sum_{k=0}^3 \frac{Z_A Z_{\text{disc}} \langle D_s^*(\hat{k}) | A_k | H_s \rangle |_{\text{lat}}}{2\sqrt{M_{H_s} M_{D_s^*}}}. \quad (6.18)$$

6.2.4 Extrapolation to the Physical Point

We now address the extrapolation of the lattice $h_{A_1}^s(1)$ values to continuum and physical b and l masses. In the process of the extrapolation, we also aim to determine the HQET low energy constants $l_{V,A,P}^s$. This process requires a number of considerations.

1: Heavy Mass Dependence

Our extrapolation in the m_h direction can be guided by HQET. The HQET expression for $h_{A_1}^s(1)$ (where here we consider both h and c to be heavy quarks in the HQET context) is given by Eq. (2.72) from Sec. 2.4.1:

$$h_{A_1}^s(1) = \eta_A \left(1 - \frac{l_V}{(2m_c)^2} + \frac{l_A}{2m_c m_h} - \frac{l_P}{(2m_h)^2} \right) + \mathcal{O} \left(\frac{1}{m_c^n m_h^m}, n + m \geq 3 \right). \quad (6.19)$$

Luke's theorem dictates that this form factor has no $\mathcal{O}(1/m_{h,c})$ corrections. η_A is an ultraviolet matching factor between HQET and QCD, and contains (weak) dependence on m_h .

2: Quark Mass Proxies

Attention must be paid to what to input for the masses $m_{h,c}$ in the above expression (6.19). Finding continuum quark masses corresponding to lattice bare masses would be a considerable task. Even if we took this on, what renormalization scheme should the masses belong to? In HQET, the masses that define the power counting should be pole masses [180]. Due to renormalons, the definition of a pole mass m also has an ambiguity of order Λ_{QCD}/m [181].

Because of this, we cannot exactly reproduce the HQET expression for $h_{A_1}^s(1)$ (6.19) in our fit. We instead test a number of proxies for the quark masses. Since we are not exactly reproducing the HQET expression, our results for $l_{V,A,P}$ are not exact but rather should be interpreted as ballpark estimations.

One possible approach is the following. The quark masses in Eq. (6.19) could be related to the meson masses (that we have access to via the correlator fits) using HQET. To see this, first consider the HQET expansion of a heavy-light meson [182]:

$$M_{H_s} = m_{h,S} + \bar{\Lambda}_S + \frac{\mu_{\pi,S}^2 - d_{H^{(*)}} \mu_{G,S}^2}{m_{h,S}} + \mathcal{O} \left(\frac{1}{m_h^2} \right). \quad (6.20)$$

where $d_{H^{(*)}} = 1$ for pseudoscalar mesons and $-1/3$ for vectors. $\bar{\Lambda}_S, \mu_{\pi,S}, \mu_{G,S}$ are HQET parameters. q labels the light quark in the meson. As already mentioned, m_h is defined in some renormalization scheme S , and since the meson mass M_{H_q} is scheme-independent, the HQET parameters must also take on scheme dependence to cancel the dependence in m_h .

A simple rearrangement of the above (6.20) gives us

$$\begin{aligned} m_{h,S} &= M_{H_s} - \bar{\Lambda}_S - \frac{\mu_{\pi,S}^2 - d_{H^{(*)}} \mu_{G,S}^2}{M_{H_s} - \bar{\Lambda}_S} + \mathcal{O}\left(\frac{1}{m_h^2}\right), \\ &\equiv \frac{1}{\varepsilon_{h,S}} + \mathcal{O}\left(\frac{1}{m_h^2}\right). \end{aligned} \quad (6.21)$$

For two heavy-light mesons, for example $M_{H_s}, M_{D_s^*}$, one can show (recognising that $\varepsilon_h \sim \mathcal{O}(1/m_h)$)

$$\frac{1}{m_{h,S} m_{c,S}} = \varepsilon_{h,S} \varepsilon_{c,S} + \mathcal{O}\left(\frac{1}{m_c^n m_h^m}, n+m \geq 3\right). \quad (6.22)$$

Since we are aiming to find the low energy constants in the context of HQET at order below $\mathcal{O}(1/m_c^n m_h^m, n+m \geq 3)$, we can safely replace the quark masses m_c, m_h in (6.19) with $\varepsilon_{h/c}^{-1}$.

For our calculation, we used HQET parameters calculated in [182] in the minimal renormalon-subtraction scheme: $\bar{\Lambda}_{\text{MRS}} = 0.552(30)\text{GeV}$, $\mu_{\pi,\text{MRS}}^2 = 0.06(22)\text{GeV}^2$, $\mu_{G,\text{MRS}}^2 = 0.38(1)\text{GeV}^2$. We are free to arbitrarily choose this choice of scheme, since the resulting ambiguity in the masses, $\Lambda_{\text{QCD}}/m_{h,c}$, are absorbed into higher orders in the HQET expansion.

Unfortunately, the quark mass dependence in η_A prevents this approach from resulting in exactly the correct $l_{V,A,P}$ values. η_A contains ratios m_c/m_h and logs of m_c/m_h , that cannot simply be redefined in this way such that ambiguities are pushed into higher orders of $1/m_{h,c}$.

We also implement the fit with more simple proxies for $m_{h,c}$. We tried replacing $m_{h,c}$ with M_{H_s, D_s^*} or $M_{\eta_{h,c}}/2$. We find the results of the extrapolation are very insensitive to the choice of proxy (see Fig. 6.9). Therefore in the end, we take our final fit function using the simplest choice of replacing $m_{h,c}$ with $M_{\eta_{h,c}}/2$. This means we have not inserted any ambiguity due to renormalization scheme choice.

3: Implementation of η_A

η_A accounts for matching between HQET and QCD, and has been computed to 2-loop: $\eta_A = 0.960(7)$ [183]. It is dependent on $m_{h,c}$, so one may worry that, if we

are going to use this expression for the extrapolation in m_h , we must account for the m_h dependence in η_A . However this dependence is weak in the region of m_h we are interested in ($m_c \leq m_h \leq m_b$). This can be seen by examining how the 1-loop expression for η_A varies with m_h [184]:

$$\eta_A(m_h) = 1 - \frac{\alpha_s}{\pi} \left(\frac{m_h + m_c}{m_h - m_c} \ln \left(\frac{m_c}{m_h} \right) + \frac{8}{3} \right). \quad (6.23)$$

Fig. 6.4 shows the variation of η_A throughout this range, the value changes by around 1.5%. The two-loop correction is an order of magnitude smaller than this [183].

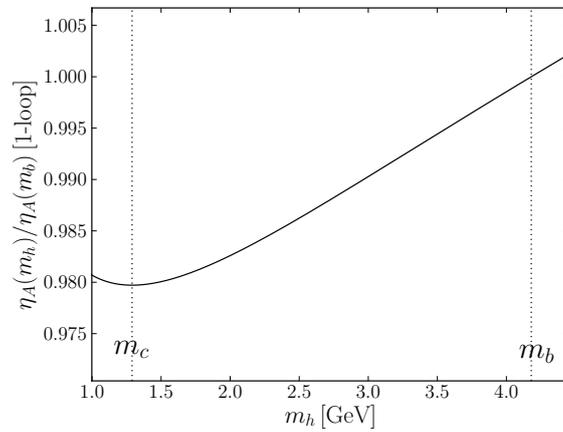


Figure 6.4: The variation of the 1-loop expression for η_A (Eq. (6.23)) throughout the $m_c \leq m_h \leq m_b$ range. For m_b and m_c values we used $m_b^{\overline{\text{MS}}}(m_b^{\overline{\text{MS}}})$ and $m_c^{\overline{\text{MS}}}(m_c^{\overline{\text{MS}}})$. For the coupling constant we used $\alpha_s(\sqrt{m_b m_c})$.

We cannot consistently include η_A in our fit function for the continuum and heavy mass extrapolation since we do not have access to the pole $m_{h,c}$ masses. We ran the extrapolation using a number of reasonable approaches to estimating the η_A behaviour and found that the final result was very insensitive to our choice of approach. We implemented the extrapolation with

- $\eta_A = 1$,
- $\eta_A = 1$ -loop expression with m_c/m_h replaced with M_{η_c}/M_{η_h} ,
- $\eta_A = 1 + \rho \log(M_{\eta_c}/M_{\eta_h})$, where ρ is a fit parameter with prior distribution 0 ± 1 .

The final result was stable upon varying these choices (see Fig. 6.9). The last bullet point, for testing to see if logarithms of m_h can be resolved in the data, resulted in a ρ consistent with zero and a decrease of the Bayes factor for the fit by a factor of 25. Clearly, the lattice data cannot resolve logarithms in m_h .

We choose $\eta_A = 1$ for simplicity. The $l_{V,A,P}$ results however are sensitive to the η_A implementation, since not properly accounting for the m_h dependence in η_A can lead to that variance in m_h being absorbed into $l_{V,A,P}$. This is another reason to take our $l_{V,A,P}$ results as estimates rather than determinations.

The fit form we used for the full continuum and heavy mass extrapolation of $h_{A_1}^s(1)$ is

$$h_{A_1}^s(1)|_{\text{fit}} = 1 - \left(\frac{1}{M_{\eta_c}}\right)^2 l_V + \frac{2}{M_{\eta_h} M_{\eta_c}} l_A - \left(\frac{1}{M_{\eta_h}}\right)^2 l_P + \mathcal{N}_{\text{disc}} + \mathcal{N}_{\text{mistuning}}. \quad (6.24)$$

$\mathcal{N}_{\text{disc}}$ and $\mathcal{N}_{\text{mistuning}}$ are nuisance parameters to account for discretization and mass mistuning effects, defined in the following subsections. $l_{V,A,P}$ are taken here as fit parameters with prior distributions $0 \pm 1 \text{GeV}^2$.

4: Discretization Effects

Discretization effects in the data are accounted for by including (following the methodology of [5]):

$$\mathcal{N}_{\text{disc}} = \sum_{i,j,k=0}^{2,2,2} d_{ijk} \left(\frac{2\Lambda_{\text{QCD}}}{M_{\eta_h}}\right)^i \left(\frac{am_{h0}^{\text{val}}}{\pi}\right)^{2j} \left(\frac{am_{c0}^{\text{val}}}{\pi}\right)^{2k}. \quad (6.25)$$

d_{ijk} are fit parameters with prior distributions 0 ± 1 . We account here for discretization effects from the two largest scales in the system; the heavy and charm masses. All discretization effects are of even order by construction of the HISQ action.

We tried including extra terms of size $(a\Lambda_{\text{QCD}})^2, (am_{s0}^{\text{val}})^2, (am_{l0}^{\text{val}})^2$, but the data could not resolve effects of that size, so it made no difference to the fit. We also tested the effects of increasing the number of terms in each sum (see Fig. 6.9), but the final result remained unchanged.

5: Mass Mistunings

Any possible mistuning of the charm mass is automatically accounted for in HQET part of the fit function (6.24). To obtain the final result we set M_{η_c} to the physical value given in the PDG [3], hence any charm mistuning is removed.

The strange and light mistunings are accounted for using a formalism introduced in [139]. To deal with possible (valence and sea) strange mistuning, we define the terms $\delta_s^{(\text{val})} = m_{s0}^{(\text{val})} - m_s^{\text{tuned}}$, where m_s^{tuned} is defined by

$$m_s^{\text{tuned}} = m_{s0} \left(\frac{M_{\eta_s}^{\text{phys}}}{M_{\eta_s}} \right)^2. \quad (6.26)$$

$M_{\eta_s}^{\text{phys}} = 0.6885(40)\text{GeV}$ is determined in lattice simulations from the masses of the pion and kaon [141].

We similarly account for (sea) light quark mistuning by defining $\delta_l = m_{l0} - m_l^{\text{tuned}}$. We can find m_l^{tuned} from m_s^{tuned} , by leveraging the fact that the ratio of quark masses is regularization independent and was calculated in [114] to be

$$\left. \frac{m_s}{m_l} \right|_{\text{phys}} = 27.18(10). \quad (6.27)$$

We set m_l^{tuned} to m_s^{tuned} divided by this ratio.

Chiral perturbation theory dictates that perturbations in quark masses cause linear contributions to the form factor. Hence the full term we include to account for mistuning is given by

$$\mathcal{N}_{\text{mistuning}} = \frac{c_s^{\text{val}} \delta_s^{\text{val}} + c_s \delta_s + 2c_l \delta_l}{10m_s^{\text{tuned}}}, \quad (6.28)$$

where c_s^{val} , c_s and c_l are fit parameters with prior distributions 0 ± 1 . We divide all terms by m_s^{tuned} to absorb any running of the quark masses in $\delta_{l,s}^{(\text{val})}$ with the cutoff, that varies between ensembles. The factor of 10 in the denominator is to bring this term close in magnitude to the chiral perturbation theory contributions that it represents. We neglect $\delta_{s,l}^{(\text{val})2}$ contributions since these are an order of magnitude smaller and are not resolved by the data.

6: Negligible Effects

The finite volume effects in our lattice results are negligible. Since the lightest valence quarks in our simulation are s quarks, the lightest particles that can arise from loop diagrams in the decay are Kaons. In appendix F of [33], the HMS χ PT finite volume effect on the fine-physical ensemble, as a function of the lightest meson appearing in loops, was found (from the formulas derived in [172]). At the Kaon mass, the finite volume effect is many orders of magnitude smaller than any of our other sources of error.

In our simulation we set $m_u = m_d \equiv m_l$, our results do not account for the difference $m_d - m_u$. We tested for any possible influence this has on our fit, by moving the m_l^{tuned} value up and down by the PDG value for $m_d - m_u$ [3]. The effect was negligible in comparison to the other sources of error.

Since we take the physical result at $M_{\eta_h} = M_{\eta_b}$, the uncertainty in M_{η_b} will contribute an uncertainty in the final result. We use the PDG result for M_{η_b} [3]. However, $\bar{b}-b$ annihilation and electroweak corrections make this somewhat different to the appropriate value on the lattice. We estimate the corresponding uncertainty in M_{η_b} to be no greater than $\pm 10 \text{ MeV}$. To see how this changes the result of the extrapolation, we varied M_{η_b} up and down by 10 MeV and studied how our final result for $h_{A_1}^s(1)$ changes. The change is less than 10^{-5} , which is negligible in comparison to our other errors.

6.3 Results

6.3.1 $h_{A_1}^s(1)$

The values extracted from 3-point correlation function fits for $h_{A_1}^s(1)$, along with quantities required for its extrapolation to the physical point, are given in Tables 6.4 and 6.5.

The results of the extrapolation through heavy mass of $h_{A_1}^s(1)$ is depicted in Fig. 6.5. By evaluating our fit form (6.24) at $a = 0$, $M_{\eta_{h,c}} = M_{\eta_{h,c}}^{\text{phys}}$ and $\delta_{s,l}^{(\text{sea})} = 0$, we reach our final, fully non-perturbative result for the $B_s \rightarrow D_s^*$ form factor at zero recoil:

$$\mathcal{F}^{B_s \rightarrow D_s^*}(1) = h_{A_1}^s(1) = 0.9020(96)_{\text{stat}}(90)_{\text{sys}}. \quad (6.29)$$

Adding the statistical and systematic errors in quadrature, we find a total fractional error of 1.45%. The error budget for this result is given in Table 6.6. The continuum/quark mass extrapolation had a goodness of fit of $\chi^2/N_{\text{dof}} = 0.16$ (for $N_{\text{dof}} = 12$).

We include in Fig. 6.5 a determination from the only other unquenched lattice calculation of this quantity [33]. They report a value of $h_{A_1}^s(1) = 0.883(12)_{\text{stat}}(28)_{\text{sys}}$. Our two studies, containing independent systematic uncertainties, are in agreement. Their study used the same gluon ensembles, with HISQ s and c valence quarks, and an NRQCD b quark. Using NRQCD meant they could perform their simulation directly at the physical b mass. However, the matching

Set	am_h^{val}	$h_{A_1}^s(1)$	aM_{H_s}	$aM_{D_s^*}$
2	0.5	0.9255(20)	0.95972(12)	0.96616(44)
	0.65	0.9321(22)	1.12511(16)	
	0.8	0.9434(24)	1.28128(21)	
3	0.5	0.9231(21)	0.95462(12)	0.93976(42)
	0.8	0.9402(27)	1.27577(22)	
4	0.427	0.9107(46)	0.77453(24)	0.63589(49)
	0.525	0.9165(49)	0.88487(31)	
	0.65	0.9246(65)	1.02008(39)	
	0.8	0.9394(66)	1.17487(54)	
5	0.5	0.9143(51)	0.80245(24)	0.47164(39)
	0.65	0.9273(62)	0.96386(33)	
	0.8	0.9422(72)	1.11787(43)	

Table 6.4: Values extracted from correlation function fits for $h_{A_1}^s(1)$, along with masses of the two mesons on either end of the transition. $h_{A_1}^s(1)$ values are found from Eq. (6.18). Errors are statistical.

Set	am_h^{val}	aM_{H_c}	af_{H_c}	aM_{η_h}	aM_{η_c}	aM_{η_s}
2	0.5	1.419515(41)	0.186299(70)	1.471675(38)	1.367014(40)	0.313886(75)
	0.65	1.573302(40)	0.197220(77)	1.775155(34)		
	0.8	1.721226(39)	0.207068(78)	2.064153(30)		
3	0.5	1.400034(28)	0.183472(62)	1.470095(25)	1.329291(27)	0.304826(52)
	0.8	1.702456(23)	0.203407(45)	2.062957(19)		
4	0.427	1.067224(46)	0.126564(70)	1.233585(41)	0.896806(48)	0.207073(96)
	0.525	1.172556(46)	0.130182(72)	1.439515(37)		
	0.65	1.303144(46)	0.133684(75)	1.693895(33)		
	0.8	1.454205(46)	0.137277(79)	1.987540(30)		
5	0.5	1.011660(32)	0.098970(52)	1.342639(65)	0.666586(89)	0.15412(17)
	0.65	1.169761(34)	0.100531(60)	1.650180(56)		
	0.8	1.321647(37)	0.101714(70)	1.945698(48)		

Table 6.5: Values extracted from correlation function fits. f_{H_c} is the H_c decay constant derived from Eq. (4.37).

Source	% Fractional Error
Statistics & Z_A	1.06
$a \rightarrow 0$	0.73
$m_h \rightarrow m_b, c$ -mistuning	0.69
l and s mistuning	0.20
Total	1.45

Table 6.6: Error budget for $h_{A_1}^s(1)$. The value for statistics & Z_A is given by the partial standard deviation of $h_{A_1}^s(1)$ with respect to the lattice data. The value for $a \rightarrow 0$ is the partial standard deviation of $h_{A_1}^s(1)$ with respect to priors of the fit parameters in $\mathcal{N}_{\text{disc}}$. Similarly for $m_h \rightarrow m_b, c$ -mistuning the value is the partial standard deviation with respect to priors of $l_{V,A,P}$, and the mistuning value comes from priors of parameters in $\mathcal{N}_{\text{mistuning}}$.

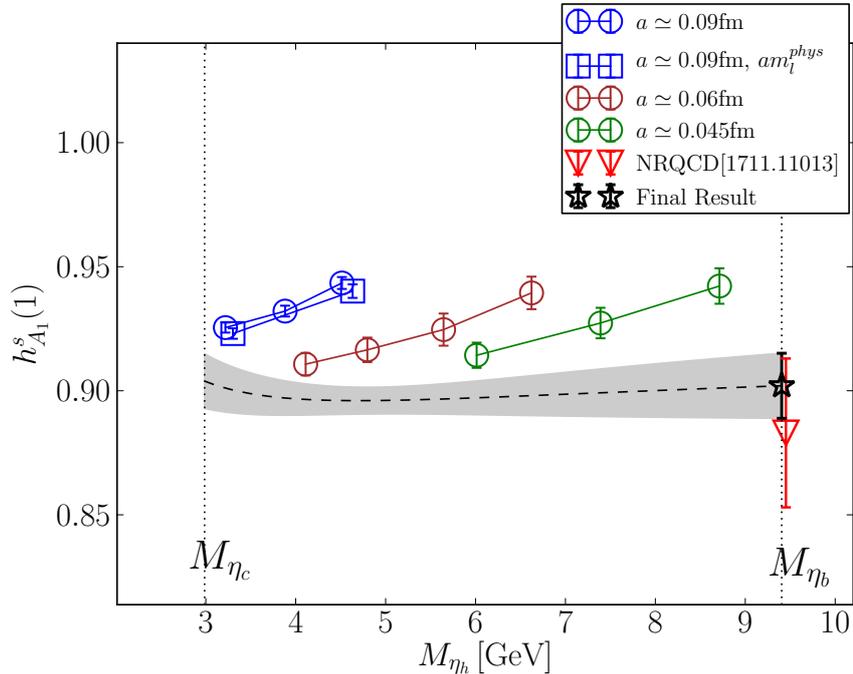


Figure 6.5: $h_{A_1}^s(1)$ against M_{η_h} (a proxy for the heavy quark mass). The grey band shows the result of the extrapolation at $a = 0$ and physical l, s and c masses. Sets listed in the legend follow the order of sets in Table 6.1. The red point represents a determination of the same quantity from a previous study using the NRQCD action for the b [33].

of lattice NRQCD-HISQ currents to continuum QCD causes their dominant error. Their result contains errors associated with the truncation of the NRQCD-HISQ current, of sizes $\mathcal{O}(\alpha_s^2)$, $\mathcal{O}(\alpha_s \Lambda_{\text{QCD}}/m_b)$ and $\mathcal{O}((\Lambda_{\text{QCD}}/m_b)^2)$. Adding these corrections in quadrature we find a 2.8% error, while their total error is reported as 2.9%. Our result is much more precise since it does not suffer from these large matching errors.

6.3.2 Implications for $B \rightarrow D^*$

Chiral symmetry implies that the $B_s \rightarrow D_s^*$ form factor should be very close to the equivalent $B \rightarrow D^*$ form factor [172]. This was found to be the case in previous studies (e.g. [33]).

As an additional test of this claim, we obtained lattice data for $h_{A_1}(1)$ on the fine ensemble, for comparison with the $h_{A_1}^s(1)$ data within our formalism. This involved

an identical process to that of obtaining $h_{A_1}^s(1)$, except with the strange valence quark replaced with a valence quark of a mass equal to am_{l0} , the sea light quark mass.

The $h_{A_1}(1)$ data is shown in comparison to the $h_{A_1}^s(1)$ data in Fig. 6.6. Errors are statistical. The error on $h_{A_1}(1)$ is much larger due to the presence of the valence light quark. There is no statistically significant difference between $h_{A_1}(1)$ and $h_{A_1}^s(1)$ here.

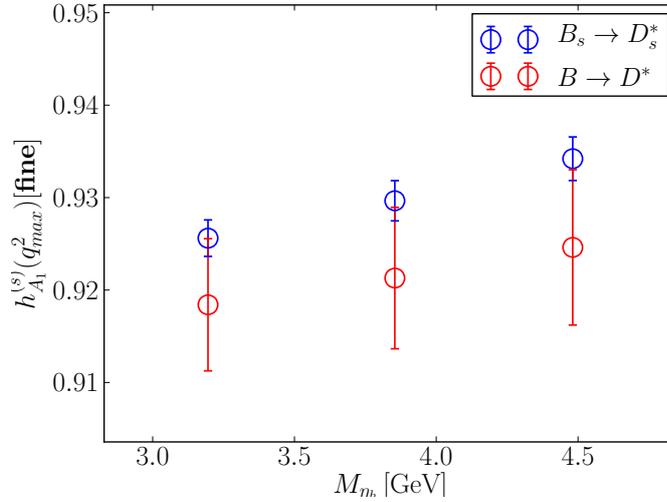


Figure 6.6: $h_{A_1}(1)$ and $h_{A_1}^s(1)$ data on the fine ensemble. Note that on this ensemble the light quark is not physical, $m_l/m_s = 0.2$.

In [33], the ratio between these two quantities was computed - $h_{A_1}(1)/h_{A_1}^s(1) = 1.013(14)_{\text{stat}}(17)_{\text{sys}}$. Multiplying this by our result for $h_{A_1}^s(1)$, one finds a result consistent with the two previous $h_{A_1}(1)$ determinations:

$$\mathcal{F}^{B \rightarrow D^*}(1) = h_{A_1}(1) = 0.914(24). \quad (6.30)$$

While this result does rely on NRQCD, it in principle suffers from much smaller perturbative matching errors. This is because the overall normalization of the axial vector NRQCD-HISQ current cancels in the ratio $h_{A_1}(1)/h_{A_1}^s(1)$. Errors due to the truncation of the NRQCD-HISQ currents in the $1/m_b$ series will remain however.

In Fig. 6.7, we show all current lattice results for $h_{A_1}(1)$ and $h_{A_1}^s(1)$. In Fig. 6.8, we show lattice data from previous FNAL/MILC and HPQCD studies, along with their final results, and the final result of this study, against 'pion mass'. Here

pion mass refers to the mass of a pion containing quarks with the mass of the spectator quark. Here we can see that the FNAL/MILC lattice data is very flat in the spectator quark mass, so if they were to extrapolate their data to find $h_{A_1}^s(1)$, it would likely be in agreement with our result. Since our result requires no perturbative normalization, while the other two studies do, we can see this agreement as an important check of the normalization of the previous studies.

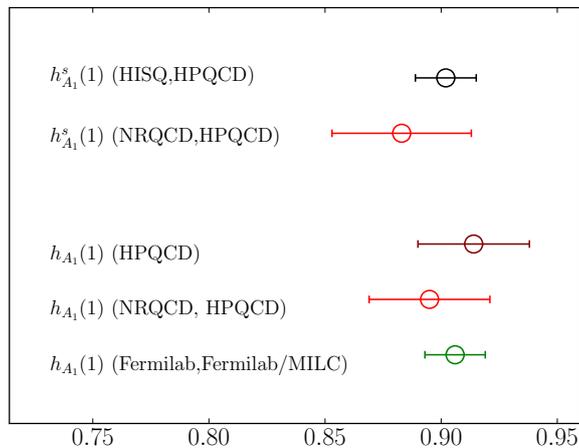


Figure 6.7: $h_{A_1}^{(s)}(1)$ from different calculations. Our result is marked (HISQ,HPQCD). Those marked (NRQCD,HPQCD) are from [33]. The quantity marked (HPQCD) is the result of multiplying our result for $h_{A_1}^s(1)$ with the ratio $h_{A_1}(1)/h_{A_1}^s(1)$ computed at [33]. The quantity marked (Fermilab,Fermilab/MILC) is from [38]. Note that our methodology is very different to that of Fermilab/MILC in a number of ways, so the comparison between our and their results is a very robust test.

6.3.3 HQET Low Energy Constants

Our fit of the lattice data to our fit function (Eq. (6.24)) produced the fit parameters $l_{V,A,P}$, which as discussed in Sec. 6.2.4 are numerically approximately equal to the

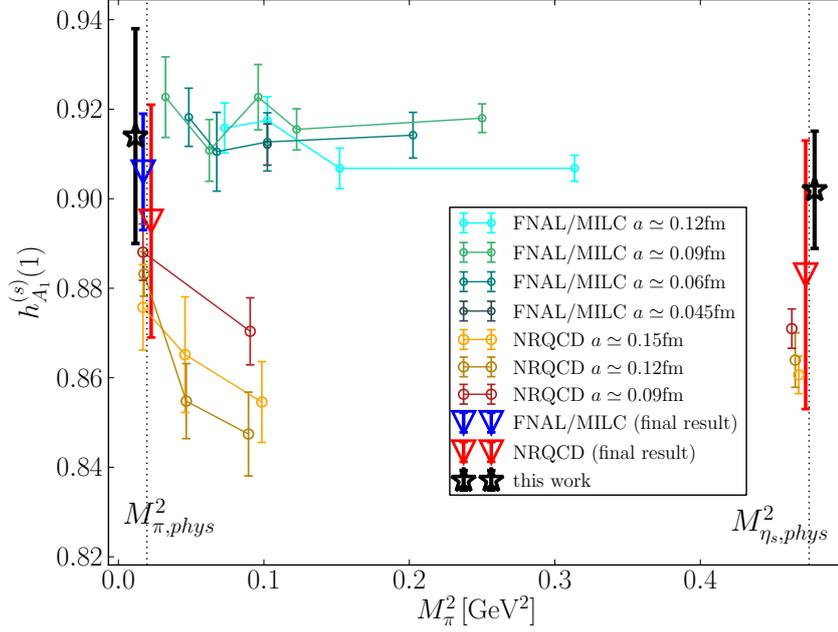


Figure 6.8: Lattice data and continuum extrapolated data for three studies of $h_{A_1}(1)$ and $h_{A_1}^s$, against the pion mass. Points labeled FNAL/MILC are from [38], and those labeled NRQCD are from [33]. The x-axis must be taken with a pinch of salt, the points at $M_\pi = M_{\eta_s}$ have pions in the sea of smaller masses than M_{η_s} , but we place them here to signify that the spectator quark has the mass of a strange quark.

low energy HQET constants of the same name. We find

$$\begin{aligned}
 l_V &= 0.71(28)\text{GeV}^2, \\
 l_A &= -0.34(32)\text{GeV}^2, \\
 l_P &= -0.53(34)\text{GeV}^2.
 \end{aligned} \tag{6.31}$$

An estimate from the ISGW model for $B \rightarrow D^*$ decays gives [185]

$$l_P \simeq l_V \simeq 0.39\text{GeV}^2. \tag{6.32}$$

These however do not come with any error, preventing a meaningful comparison between the ISGW model and our results.

6.3.4 Extrapolation Stability

We performed a number of tests of the continuum/heavy mass extrapolation. The results of each of these tests are given in Fig. 6.9.

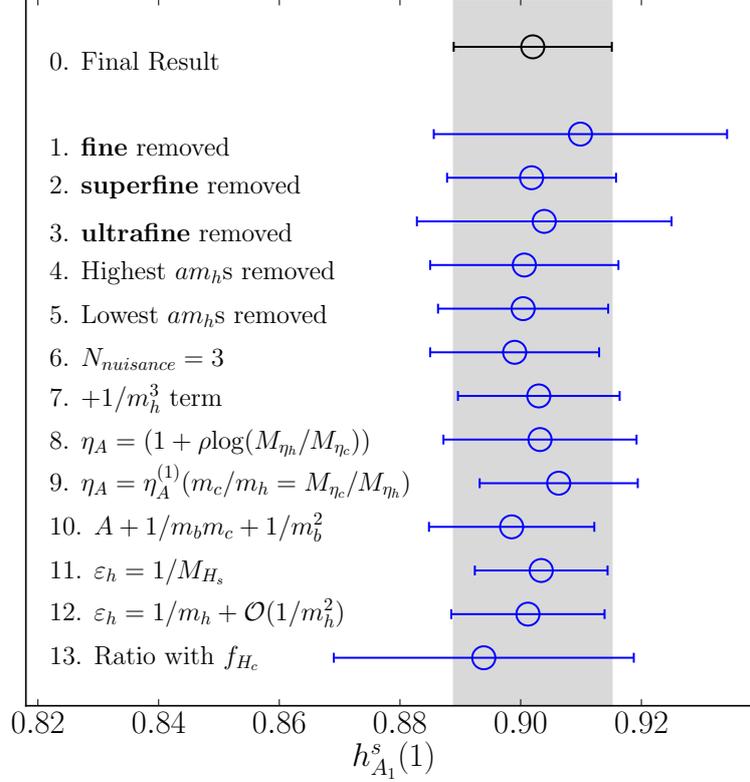


Figure 6.9: Results of $h_{A_1}^s(1)$ extrapolation tests. Points **1-3** show the final result if data from the fine, superfine or ultrafine ensembles are not used in the fit. Points **4 & 5** points show the result if data at the highest/lowest $am_{h_0}^{\text{val}}$ value on each ensemble are removed. Point **6**, ' $N_{\text{nuisance}} = 3$ ' shows the result of truncating each sum in $\mathcal{N}_{\text{disc}}$ (Eq. (6.25)) at 3 rather than 2. Point **7**, ' $+1/m_b^3$ ' results from adding an extra term to (6.24) of the form $p/M_{\eta_h}^3$ where p is a fit parameter with the same prior as $l_{V,A,P}^s$. In this case, the Bayes factor falls by a factor of 7, suggesting that the data does not contain a cubic dependence on the heavy mass. Points **8 & 9** show the results of the implementations of η_A described in Sec. 6.2.4. ρ is a fit parameter with prior distribution 0 ± 1 . Including this factor causes the Bayes factor to drop by a factor of 20, implying that the data cannot resolve logarithms in m_h . Point 9 shows the result of using the 1-loop expression for η_A (Eq. (6.23)), with m_c/m_h replaced with M_{η_c}/M_{η_h} . Point **10**, ' $A + 1/m_b m_c + 1/m_b^2$ ' is the result of replacing $1 + l_V/m_c^2$ in the fit with simply a fit parameter A with prior distribution 1 ± 1 . The fact that this does not affect the fit implies that charm mistuning does not strongly affect the extrapolation. Points **11 & 12** show the result of replacing the heavy mass proxy $M_{\eta_h}/2$ with M_{H_s} and Eq. (6.21) respectively. Point **13**, 'Ratio with f_{H_c} ' is the result of an alternative extrapolation described in Sec. 6.3.4.

One of the tests requires some explanation, the result of which is given in Fig. 6.9, labelled 'Ratio with f_{H_c} '. We performed a continuum/heavy mass extrapolation in the ratio $h_{A_1}^s(1)/(f_{H_c} \sqrt{M_{H_c}})$. f_{H_c} is found from fitting the H_c correlation functions to obtain $a_0^{H_c}$, and using Eq. (4.37). Since we create the H_c mesons with a local HISQ pseudoscalar current, which is absolutely normalized, no renormalization of f_{H_c} is required here. Details of the extrapolation are given below.

Discretization effects cancel to a large extent in this ratio. It however varies strongly with changing heavy mass. This makes the extrapolation very different from the extrapolation in $h_{A_1}^s(1)$, which has large discretization effects but has little variation in the heavy mass. The two extrapolations have quite different systematics, so testing their agreement is a stringent test of our formalism.

In order to compare the result of the two extrapolations, we must multiply $h_{A_1}^s(1)/(f_{B_c} \sqrt{M_{B_c}})$ by $f_{B_c} \sqrt{M_{B_c}}$. We can use the PDG value for M_{B_c} [3]. For an f_{B_c} value, we extrapolate our f_{H_c} data to the physical point.

We used a similarly structured fit form for both the $h_{A_1}^s(1)/(f_{B_c} \sqrt{M_{B_c}})$ and f_{H_c} extrapolations. We followed the methodology of [5]. Both extrapolations use a fit function of the form

$$\begin{aligned} \text{fit} = & A \left(\frac{\alpha_s(M_{\eta_h}/2)}{\alpha_s(M_{\eta_c}/2)} \right)^{6s/25} M_{\eta_h}^{n/2} \sum_{i,j,k=0}^{2,2,2} d_{ijk} \left(\frac{2\text{GeV}}{M_{\eta_h}} \right)^i \left(\frac{am_{h0}^{\text{val}}}{\pi} \right)^{2j} \left(\frac{am_{c0}^{\text{val}}}{\pi} \right)^{2k} \\ & \times (1 + \mathcal{N}_{\text{mistuning}} + \mathcal{N}_{\text{mistuning}}^c). \end{aligned} \quad (6.33)$$

$\alpha_s(M)$ is the QCD coupling evaluated at scale M (according to results from [139] with $N_f = 5$). $s = +1$ and $n = 0$ for $h_{A_1}^s(1)/(f_{H_c} \sqrt{M_{H_c}})$, $s = -1$ and $n = -1$ for f_{H_c} . The $M_{\eta_h}^{n/2}$ accounts for the leading order dependence of f_{H_c} in HQET, and the α_s ratio comes from renormalization group improved matching between QCD and HQET of f_{H_c} . $\mathcal{N}_{\text{mistuning}}$ is defined in Eq. (6.28). We have introduced a new mistuning term for the charm:

$$\mathcal{N}_{\text{mistuning}}^c = c_c \left(\frac{M_{\eta_c} - M_{\eta_c}^{\text{phys}}}{M_{\eta_c}^{\text{phys}}} \right), \quad (6.34)$$

where $M_{\eta_c}^{\text{phys}}$ is taken from the PDG [3], and c_c is a fit parameter with prior distribution 0 ± 1 .

A is given prior distribution $0 \pm 4\text{GeV}^{3/2}$ in the f_{H_c} case and $0 \pm 2\text{GeV}^{-3/2}$ in the ratio case. d_{ijk} are given priors of 0 ± 2 in all cases except d_{000} which is set to 1.

The result of the extrapolation of $h_{A_1}^s(1)/(f_{H_c} \sqrt{M_{H_c}})$ at the physical point was multiplied by our $f_{B_c} M_{B_c}$ result to obtain a second determination of $h_{A_1}^s(1)$. This

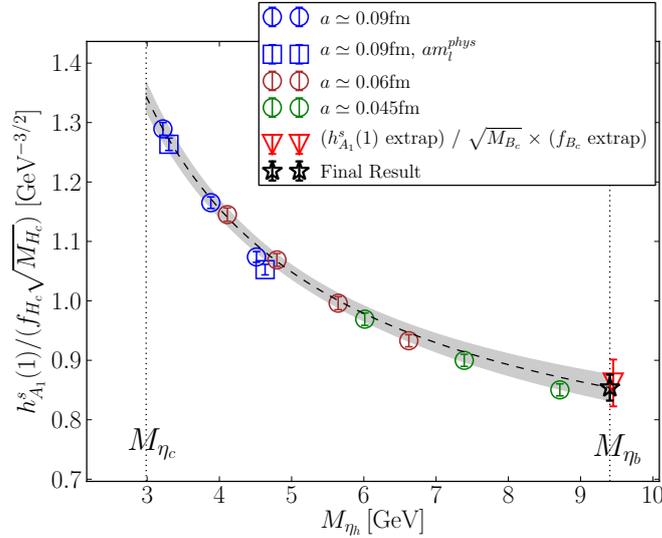


Figure 6.10: $h_{A_1}^s(1)/(f_{H_c}\sqrt{M_{H_c}})$ against M_{η_h} (a proxy for the heavy quark mass). The grey band shows the result of the extrapolation at $a = 0$ and physical l, s and c masses. Sets listed in the legend follow the order of sets in Table 6.1. The black point shows our final result for $h_{A_1}^s(1)$ divided by $\sqrt{M_{B_c}}$ from the PDG [3] and f_{B_c} from our extrapolation of f_{H_c} to continuum and physical b mass.

is the result given in Fig. 6.9 labelled 'Ratio with f_{H_c} '.

The extrapolation in f_{H_c} is shown in Fig. 6.11. We here include the result from a previous heavy-HISQ determination of f_{B_c} on $N_f = 2 + 1$ MILC ensembles [5]. Our final result for this quantity is

$$f_{B_c} = 0.4178(45) \text{ GeV}. \quad (6.35)$$

6.3.5 H_s and D_s^* Masses

As a further consistency check of our results, we can check if the masses for the H_s and D_s^* mesons, extracted from our correlator fits, reproduce what we expect physically.

Fig. 6.12 shows D_s^* mass extracted from correlators on each ensemble. Each are consistent with the experimentally measured D_s^* mass (the grey band).

We performed an extrapolation of $M_{H_s} - M_{\eta_h}/2$ masses to continuum $m_h = m_b$ and $m_h = m_c$, for comparison with the known value for $M_{B_s} - M_{\eta_b}/2$ and

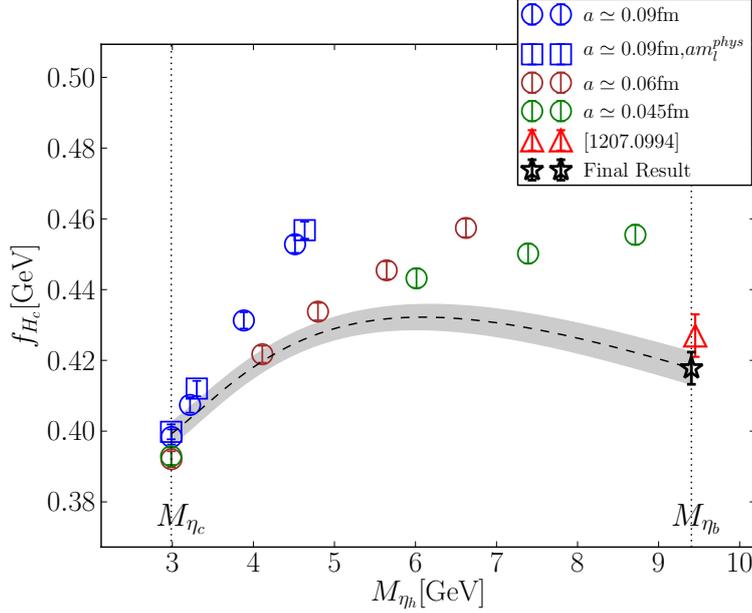


Figure 6.11: f_{H_c} against M_{η_h} (a proxy for the heavy quark mass). The grey band shows the result of the extrapolation at $a = 0$ and physical l, s and c masses. Sets listed in the legend follow the order of sets in table 6.1. The red point shows the result from a previous heavy-HISQ determination of f_{B_c} on 2+1 gauge ensembles [5].

$M_{D_s} - M_{\eta_c}/2$. To perform this extrapolation we use the fit form

$$\left(M_{H_s} - \frac{M_{\eta_h}}{2} \right) \Big|_{\text{fit}} = \left(\sum_{n=-1}^{+1} c_n \left(\frac{M_{\eta_h}}{2\text{GeV}} \right)^n \right) \times \quad (6.36)$$

$$\left(1 + \sum_{i,j=0}^{2,2} d_{ij} \left(\frac{am_{h0}^{\text{val}}}{\pi} \right)^{2i} \left(\frac{a\Lambda_{\text{QCD}}}{\pi} \right)^{2j} + \mathcal{N}_{\text{mistuning}} \right).$$

c_n are fit parameters. Since the lattice data for $M_{H_s} - M_{\eta_h}/2$ is close to linear, priors can be set for c_1 and c_0 by inspecting the approximate gradient and intercept of $(M_{H_s} - M_{\eta_h}/2)$ against M_{η_h} . Accordingly c_1 is given prior 0.05(5), and c_0 is given 0.5(5). c_{-1} is given 0 ± 1 . d_{ij} are given priors 0 ± 1 . $\mathcal{N}_{\text{mistuning}}$ is defined in Equation (6.28). We tested the effect of including $\mathcal{O}(1/M_{\eta_h}^2)$ and $\mathcal{O}(1/M_{\eta_h}^3)$ terms, this does not change the result in any statistically significant way.

Fig. 6.13 depicts this extrapolation. We find

$$M_{B_s} - \frac{M_{\eta_b}}{2} = 0.6588(61)\text{GeV}, \quad (6.37)$$

$$M_{D_s} - \frac{M_{\eta_c}}{2} = 0.4755(37)\text{GeV}. \quad (6.38)$$

As can be seen from Fig. 6.13, both are in agreement with the physical result.

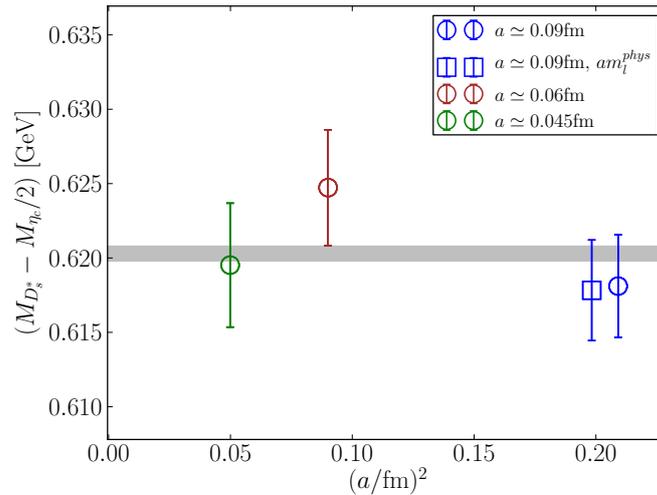


Figure 6.12: Lattice results for $M_{D_s^*} - M_{\eta_c}/2$ on each ensemble. The grey band shows the PDG result [3].

6.4 Conclusions

We have produced a fully non-perturbative determination of $h_{A_1}^s(1)$, sometimes called $\mathcal{F}^{B_s \rightarrow D_s}(1)$, using unquenched lattice data from a fully relativistic and highly improved lattice action, along with an estimation of the low energy constants $l_{V,A,P}$, given in (6.29) and (6.31) respectively. We used gauge ensembles with 3 lattice spacings, including an ensemble with approximately physical light sea quark masses, and obtained data corresponding to 12 different heavy quark masses.

This study supplies an independent check of the NRQCD formalism used in previous HPQCD studies. It is also much more precise, in the case of $h_{A_1}^s$, the total fractional error has been halved in comparison to the NRQCD determination. The comparative precision resulting from the heavy-HISQ method suggests that it is well suited to computing other form factors associated with b -decays.

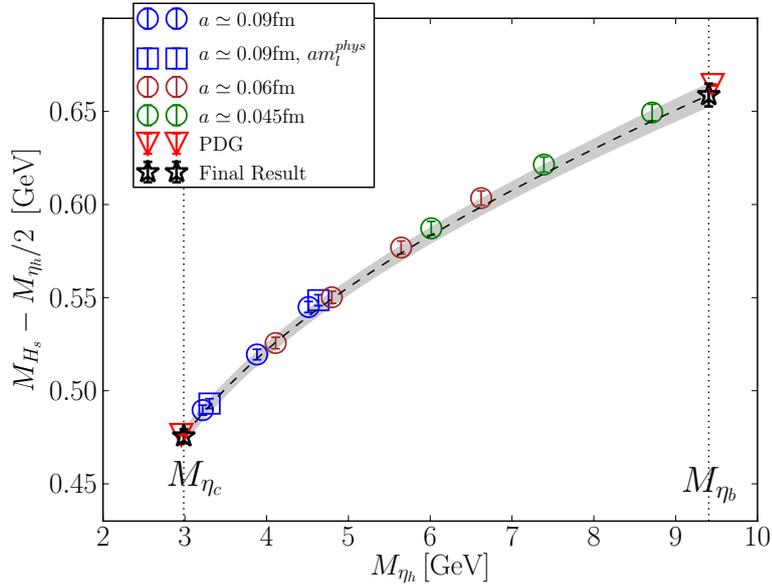


Figure 6.13: Extrapolation of $M_{H_s} - M_{\eta_h}/2$ to the physical point. The grey band shows the result at $a = 0$ and physical charm, strange and light masses.

This study also clearly demonstrates the power of the heavy-HISQ approach. It produces a result approximately twice as precise as the NRQCD result and contains fewer assumptions while being consistent with all other lattice studies of $h_{A_1}^s(1)$ and $h_{A_1}(1)$.

$B_s \rightarrow D_s \ell \nu$ Form Factors at All Physical q^2 from Heavy-HISQ

In this chapter, I present the second of our two heavy-HISQ studies, the calculation of the $B_s \rightarrow D_s \ell \nu$ form factors $f_0^s(q^2)$ and $f_+^s(q^2)$ throughout all physical q^2 , as defined in Sec. 2.2.2. Like for $h_{A_1}^s(1)$, I'm giving this quantity the superscript s to differentiate it from the more often referred to form factors for $B \rightarrow D \ell \nu$ decays.

I briefly review the definition of the form factors here for ease of reading. The differential decay rate for $B_s \rightarrow D_s \ell \nu$ decays are given in the SM by [3]:

$$\begin{aligned} \frac{d\Gamma}{dq^2} = \eta_{\text{EW}} \frac{G_F^2 |V_{cb}|^2}{24\pi^3 M_{B_s}^2} \left(1 - \frac{m_\ell^2}{q^2}\right)^2 |\mathbf{p}_{D_s}| \times \\ \left[\left(1 + \frac{m_\ell^2}{2q^2}\right) M_{B_s}^2 |\mathbf{p}_{D_s}|^2 f_+^{s2}(q^2) + \frac{3m_\ell^2}{8q^2} (M_{B_s}^2 - M_{D_s}^2)^2 f_0^{s2}(q^2) \right] \end{aligned} \quad (7.1)$$

where m_ℓ is the mass of the lepton, η_{EW} is the electroweak correction [29–31], $q^2 = (p_{B_s} - p_{D_s})^2$ is the momentum transfer, and $f_0^s(q^2)$, $f_+^s(q^2)$ are the scalar and vector form factors that parameterize the non-perturbative contribution to the decay. The allowed range of q^2 values if the final states are on-shell is

$$m_\ell^2 \leq q^2 \leq (M_{B_s} - M_{D_s})^2. \quad (7.2)$$

The form factors parameterize matrix elements of the electroweak current between B_s and D_s states, $\langle D_s | (V - A)_\mu | B_s \rangle$ where $V_\mu = \bar{b} \gamma_\mu c$ is the vector component and $A_\mu = \bar{b} \gamma_5 \gamma_\mu c$ is the axial vector component. In a pseudoscalar-to-pseudoscalar amplitude, only V_μ contributes, since $\langle D_s | A_\mu | B_s \rangle$ does not satisfy the parity invariance of QCD. The vector current in terms of form factors is given by

$$\begin{aligned} \langle D_s | V^\mu | B_s \rangle = f_+^s(q^2) \left[p_{B_s}^\mu + p_{D_s}^\mu - \frac{M_{B_s}^2 - M_{D_s}^2}{q^2} q^\mu \right] \\ + f_0^s(q^2) \frac{M_{B_s}^2 - M_{D_s}^2}{q^2} q^\mu. \end{aligned} \quad (7.3)$$

Analyticity of this matrix element demands that

$$f_+^s(0) = f_0^s(0). \quad (7.4)$$

Via the PCVC relation;

$$q_\mu \langle D_s | V^\mu | H_s \rangle = (m_h - m_c) \langle D_s | S | H_s \rangle, \quad (7.5)$$

the form factor $f_0^s(q^2)$ is also directly related to the matrix element of the scalar current $S = \bar{b}c$:

$$(m_b - m_c) \langle D_s | S | B_s \rangle = (M_{B_s}^2 - M_{D_s}^2) f_0^s(q^2). \quad (7.6)$$

In our calculation we access the form factors by computing matrix elements of the temporal vector current V_0 and the scalar current S . The form factors can be extracted from this combination using expressions derived from equations (7.3) and (7.6):

$$f_0^s(q^2) = \frac{m_b - m_c}{M_{B_s}^2 - M_{D_s}^2} \langle D_s | S | B_s \rangle, \quad (7.7)$$

$$f_+^s(q^2) = \frac{1}{2M_{B_s}} \frac{\delta^M \langle D_s | S | B_s \rangle - q^2 \langle D_s | V_0 | B_s \rangle}{\mathbf{p}_{D_s}^2}, \quad (7.8)$$

$$(\delta^M = (m_b - m_c)(M_{B_s} - E_{D_s})).$$

Our goal is to compute $f_0^s(q^2)$ and $f_+^s(q^2)$ throughout the range of q^2 values $0 \leq q^2 \leq (M_{B_s} - M_{D_s})^2 \equiv q_{\max}^2$. We extend the range to $q^2 = 0$ in order to take advantage of the constraint in Eq. (7.4). To achieve this, we compute $\langle D_s | S | H_s \rangle$ and $\langle D_s | V_0 | H_s \rangle$ on the lattice, where the H_s meson is at rest and D_s mesons are given an appropriate array of spatial momenta.

7.1 Motivation

$B_s \rightarrow D_s \ell \nu$ decays can supply a new method for precisely determining the CKM element $|V_{cb}|$. Determination of $|V_{cb}|$ in this way requires both a measurement of the branching fraction and a theoretical determination of the form factors, as explained in Sec. 2.2.2. To obtain the highest possible precision, data for both the form factors and branching fractions are required throughout the largest possible range of momentum transfer. Analogous approaches were already performed using $B \rightarrow D \ell \nu$ decays [36, 37, 42–44, 186, 187].

The $B_s \rightarrow D_s \ell \nu$ decay can also supply a new test of the SM, by comparing the theoretical and experimental determinations of the ratio R_{D_s} , defined in Eq. (2.31). This would be especially illuminating since tension has been found in the intimately related ratios $R_{D^{(*)}}$. The presence or absence of an anomaly in R_{D_s} would help to confirm or dismiss a new physics explanation for such a family of anomalies.

The $B_s \rightarrow D_s \ell \nu$ scalar form factor is useful in the experimental extraction of $B_s \rightarrow \mu^+ \mu^-$ branching fractions. Taking a ratio of the $B_s \rightarrow D_s$ and $B \rightarrow D$ scalar form factors gives the so-called fragmentation ratio, the ratio of probabilities of a b quark hadronizing into a B or B_s meson. In analyses such as [188], $B_s \rightarrow \mu^+ \mu^-$ branching fractions are measured using $B_u^+ \rightarrow J/\psi(\mu^+ \mu^-)K^+$ and $B_d^0 \rightarrow K^+ \pi^-$ as normalization channels, in this case one requires a value for the fragmentation ratio.

Similar to the $B_{(s)} \rightarrow D_{(s)}^*$ case, chiral perturbation theory implies that form factors for $B_s \rightarrow D_s$ and $B \rightarrow D$ decays are insensitive to the mass of the spectator quark, implying that form factors for these two decays are approximately equal [172]. This expectation has been validated by previous lattice calculations, for example in [189] the ratio of scalar form factors for the two decays at momentum transfer $q^2 = M_\pi^2$ was found to be $f_0^s(M_\pi^2)/f_0(M_\pi^2) = 1.054(50)$, while [152] found the value $f_0^s(M_\pi^2)/f_0(M_\pi^2) = 1.006(62)$. Hence we can expect to learn about $B \rightarrow D$ form factors by studying $B_s \rightarrow D_s$.

While $B \rightarrow D$ form factors are currently more phenomenologically useful, $B_s \rightarrow D_s$ form factors are more attractive on the lattice QCD side. The absence of valence light (u or d) quarks means lattice QCD results have smaller statistical errors, are less computationally expensive, have a more simple chiral extrapolation to the physical light mass, and negligible finite volume effects. This makes the $B_s \rightarrow D_s \ell \nu$ decay a useful test bed for lattice techniques that may be later used to study $B \rightarrow D \ell \nu$ decays.

A number of lattice calculations of $B_{(s)} \rightarrow D_{(s)}$ form factors have already been performed. The FNAL/MILC collaboration produced $B \rightarrow D$ form factors on the $N_f = 2 + 1$ MILC gluon ensembles using the Fermilab action for b and c valence quarks and ASQTAD light quarks [37]. They also, in an earlier work, computed the ratio of scalar form factors for $B_s \rightarrow D_s$ and $B \rightarrow D$ to obtain the fragmentation ratio [189]. The HPQCD collaboration computed both $B \rightarrow D$ and $B_s \rightarrow D_s$ form factors on the $N_f = 2 + 1$ MILC gluon ensembles using the NRQCD action for the valence b , and the HISQ action for all other quarks [36, 152]. Atoui et. al. also produced $B_s \rightarrow D_s$ form factors using maximally twisted Wilson quarks on

$N_f = 2$ gluon ensembles [190]. A calculation of the $B_s \rightarrow D_s$ form factors by the RBC/UKQCD collaboration is currently underway [191].

A considerable limitation in the FNAL/MILC and HPQCD studies is the requirement for perturbative matching between the lattice effective field theories and continuum QCD. FNAL/MILC required a matching that was only available to 1-loop, resulting in an $\mathcal{O}(\alpha_s^2)$ systematic error. In the HPQCD calculation, NRQCD-HISQ currents were truncated, prompting them to report large systematic errors. Besides the reported errors, as was discussed in Chapter 5, parts of the NRQCD-HISQ vector current that contribute away from zero recoil have a large magnitude ($\sim 30\%$ of the leading order). The currents used in this study did not take these large sub-leading currents into account, so the result from this may have large uncontrolled systematic errors.

Another limitation present in each of the aforementioned studies is that the lattice data is limited to a region of high q^2 . To generate lattice points at lower q^2 , a large spatial momentum must be given to one of the quarks on the lattice. Due to signal/noise degradation, adding momentum leads to an exponential increase of noise in correlation functions. Hence, in cases like $B_{(s)} \rightarrow D_{(s)}$, lattice data close to $q^2 = 0$ would be uselessly noisy. Lattice results must be limited to high q^2 . This fact necessitates an extrapolation from the data in the high q^2 region to the rest of the physical range. Since there has been some controversy in choices of form factor extrapolations through q^2 recently (for example see [39, 46]), it would be desirable to instead have lattice data covering the entire q^2 range. We can in fact achieve this with the heavy-HISQ approach. This is because in heavy-HISQ the b quarks are lighter than physical, this shrinks the q^2 range, meaning smaller spatial momenta are required to cover the range.

7.2 Calculation Details

7.2.1 Lattice Setup

This calculation closely followed the approach employed in our calculation of the $B_s \rightarrow D_s^*$ axial form factor at zero recoil, given in the last chapter. The main modifications required for this calculation were 1) the form factors are not protected by Luke's theorem, so a more general fit form for the extrapolation in m_h was required, and 2) to cover the q^2 range we gave the charm quark a number of spatial momentum values via a momentum twist (Sec. 4.1.3) and interpolated the results

set	am_{s0}^{val}	am_{c0}^{val}	am_{h0}^{val}	$ a\mathbf{p}_{D_s} $	T/a
2	0.0376	0.45	0.5	0, 0.056	14, 17, 20
			0.65	0, 0.142, 0.201	
			0.8	0, 0.227, 0.323	
3	0.036	0.433	0.5	0, 0.0279	14, 17, 20
			0.8	0, 0.162	
4	0.0234	0.274	0.427	0, 0.113, 0.161	22, 25, 28
			0.525	0, 0.161, 0.244	
			0.65	0, 0.244, 0.338	
			0.8	0, 0.338, 0.438	
5	0.0165	0.194	0.5	0, 0.202, 0.281	31, 36, 41
			0.65	0, 0.202, 0.281, 0.382	
			0.8	0, 0.281, 0.382, 0.473	

Table 7.1: Simulation details. Columns 2 and 3 give the s and c valence quark masses, which were tuned in [139]. Column 4 gives the bare heavy quark masses, we use a number of heavy quark masses to assist the extrapolation to the physical b mass. Column 5 gives the absolute value of the spatial momentum given to the D_s meson, using a momentum twist, in lattice units. These values were chosen with the following rationale: when only 2 are used, these correspond to the $q^2 = 0$ and q_{max}^2 points (except on the fine-physical ensemble, where we compute at the points q_{max}^2 and $q_{\text{max}}^2/2$). When 3 twists are used, the momenta correspond to $q^2 = 0$, $q^2 = q_{\text{max}}^2/2$, and q_{max}^2 points. When 4 are used, these are points for q_{max}^2 , $3q_{\text{max}}^2/4$, $q_{\text{max}}^2/2$, $q_{\text{max}}^2/4$, $q^2 = 0$. To give the D_s meson these spatial momenta we gave the charm an appropriate momentum twist in the $(1, 1, 1)$ direction. Column 6 gives the temporal separations between source and sink, T , of the 3-point correlation functions computed on each ensemble.

to all q^2 .

We used the same set of ensembles as in the $B_s \rightarrow D_s^*$ study. In three of the four ensembles (sets 2, 4 and 5), the bare light mass is set to $m_{l0}/m_{s0} = 0.2$. The

fact that the m_{l0} value is unphysically high is expected to have a small effect on the form factors, due to the lack of valence light quarks, and previous experience of the form factor dependence on m_{l0} [152]. The small effect due to the unphysical m_{l0} is quantified by including a fourth ensemble (set 3) with roughly physical m_{l0} , and corrected for.

We used a number of different masses for the valence heavy quark am_{h0}^{val} . Unphysically light h -quarks reduce the q^2 range, meaning we can obtain lattice data at both ends of the range while the statistical noise remains under control, unlike previous studies of these form factors.

As detailed in Sec. 4.1.3, staggered correlation functions are built by a combination of staggered propagators $g(x, y)$ and staggered phases. In this calculation we only need local (non-point-split) operators, this is an advantage since point-split operators lead to correlation functions noisier than those using local operators.

We computed a number of correlation functions on each ensemble. Valence masses, momenta and other inputs to the calculation are given in Table 7.1. First, we computed 2-point correlation functions between eigenstates of momentum \mathbf{p} , objects of the form

$$C_M(\mathbf{p}, t) = \langle \tilde{\Phi}_M(\mathbf{p}, t) \tilde{\Phi}_M^\dagger(\mathbf{p}, 0) \rangle, \quad (7.9)$$

$$\tilde{\Phi}_M(\mathbf{p}, t) = \sum_{\mathbf{x}} e^{-i\mathbf{p}\cdot\mathbf{x}} \bar{q}(\mathbf{x}, t) \Gamma q'(\mathbf{x}, t),$$

where $\langle \rangle$ represents a functional integral over all fields, q, q' are valence quark fields of the flavours the M meson is charged under, and Γ is the spin-taste structure of M . We computed these for all t values, i.e. $0 \leq t \leq N_t$.

We computed correlation functions for a heavy-strange pseudoscalar, H_s , with spin-taste structure $(\gamma_5 \otimes \gamma_5)$, at rest. In terms of staggered propagators, this takes the form

$$C_{H_s}(\mathbf{0}, t) = \sum_{\mathbf{x}, \mathbf{y}} \left\langle \text{Tr} \left[g_h(x, y) g_s^\dagger(x, y) \right] \right\rangle, \quad (7.10)$$

where $g_q(x, y)$ is a staggered propagator for flavour q , and the trace is over color. Here $x_0 = 0$ and $y_0 = t$. We also computed correlators for a charm-strange pseudoscalar meson D_s , with structure $(\gamma_\mu \otimes \gamma_\mu)$ and momentum \mathbf{p} , using

$$C_{D_s}(\mathbf{p}, t) = \sum_{\mathbf{x}, \mathbf{y}} \left\langle \text{Tr} \left[g_c^{\theta_{\mathbf{p}}}(x, y) g_s^\dagger(x, y) \right] \right\rangle, \quad (7.11)$$

where $g_q^{\theta_{\mathbf{p}}}(x, y)$ denotes a propagator with momentum twist $\theta_{\mathbf{p}}$ corresponding to momentum \mathbf{p} . We computed this using a number of twists to produce the range

of momenta given in Table 7.1. We designed the c propagators to have momentum $a\mathbf{p} = |a\mathbf{p}|(1, 1, 1)$, by imposing a twist $\theta = N_x|a\mathbf{p}|/\pi\sqrt{3}$ in each spatial direction.

We also computed non-goldstone pseudoscalar heavy-strange mesons at rest, denoted \hat{H}_s . These are necessary for extracting the vector current. This has spin-taste structure $(\gamma_0\gamma_5 \otimes \gamma_0\gamma_5)$. \hat{H}_s correlators were computed using

$$C_{\hat{H}_s}(t) = \sum_{\mathbf{x}, \mathbf{y}} (-1)^{\bar{x}_0 + \bar{y}_0} \left\langle \text{Tr} \left[g_h(x, y) g_s^\dagger(x, y) \right] \right\rangle, \quad (7.12)$$

where I use the notation $\bar{z}_\mu = \sum_{\nu \neq \mu} z_\nu$.

We also computed correlators for H_c mesons, heavy-charmed pseudoscalars, using the same form as the H_s correlator (Eq. (7.10)). This is used to find H_c decay constants, these are useful in our continuum and m_h extrapolation.

The heavy-mass extrapolation requires masses of η_h mesons, heavy-heavy pseudoscalars artificially forbidden to annihilate. To quantify mistuning of the charm and strange quark masses, we also required masses for η_c and η_s mesons, identical to η_h with h replaced c and s quarks respectively. We computed correlators for each of these at rest, using a spin-taste structure $(\gamma_5 \otimes \gamma_5)$, taking the same form as the H_s correlator (Eq. (7.10)).

We then computed 3-point correlation functions. We required two sets of such correlation functions, one with a scalar and one with a temporal vector current insertion. The first takes the form

$$C_S(\mathbf{p}, t, T) = \sum_{\mathbf{y}} \langle \tilde{\Phi}_{D_s}(\mathbf{p}, T) S(\mathbf{y}, t) \tilde{\Phi}_{H_s}(\mathbf{0}, 0) \rangle, \quad (7.13)$$

$$S(\mathbf{y}, t) = \bar{c}(\mathbf{y}, t) h(\mathbf{y}, t).$$

In terms of the staggered formalism, both the H_s source and D_s sink are given structure $(\gamma_5 \otimes \gamma_5)$, and the current insertion is given $(1 \otimes 1)$. We generated these with staggered propagators using

$$C_S(\mathbf{p}, t, T) = \sum_{\mathbf{x}, \mathbf{y}, \mathbf{z}} \left\langle \text{Tr} \left[g_h(x, y) g_c^{\theta_{\mathbf{p}}}(y, z) g_s^\dagger(x, z) \right] \right\rangle, \quad (7.14)$$

where we fix $x_0 = 0$, $y_0 = t$ and $z_0 = T$, and once again the charm propagator is given the appropriate twist $\theta_{\mathbf{p}}$. We computed these for all t values within $0 \leq t \leq T$, and 3 T values that vary between ensembles, given in Table 7.1.

To extract the temporal vector current, we required the function

$$C_{V_0}^{\mathbf{p}D_s}(t, T) = \sum_{\mathbf{y}} \langle \tilde{\Phi}_{D_s}(\mathbf{p}, T) V_0(\mathbf{y}, t) \tilde{\Phi}_{\hat{H}_s}(\mathbf{0}, 0) \rangle, \quad (7.15)$$

$$V_0(\mathbf{y}, t) = \bar{c}(\mathbf{y}, t) \gamma_0 h(\mathbf{y}, t).$$

This was generated using structures $(\gamma_0 \gamma_5 \otimes \gamma_0 \gamma_5)$ at the \hat{H}_s source, $(\gamma_5 \otimes \gamma_5)$ at the D_s sink, and $(\gamma_0 \otimes \gamma_0)$ at the current insertion. To achieve this we evaluated

$$C_{V_0}(\mathbf{p}, t, T) = \sum_{\mathbf{x}, \mathbf{y}, \mathbf{z}} (-1)^{\bar{x}_0 + \bar{y}_0} \left\langle \text{Tr} \left[g_h(x, y) g_c^{\theta \mathbf{p}}(y, z) g_s^\dagger(x, z) \right] \right\rangle. \quad (7.16)$$

The non-goldstone \hat{H}_s , as opposed to simply H_s , was required here to ensure all taste structure cancels in the fermion loop.

7.2.2 Analysis of Correlation Functions

We then extracted current matrix elements from the generated correlation functions, via simultaneous Bayesian fits, as described in Sec. 4.2.1. We performed a single simultaneous fit containing each correlator computed $(C_{H_s}, C_{\hat{H}_s}, C_{D_s}, C_{\eta_h}, C_{\eta_c}, C_{\eta_s}, C_{H_c}, C_S, C_{V_0})$ at every m_h and every $|\mathbf{p}_{D_s}|$, for each ensemble. This means that our extrapolation to the physical point can take into account correlations between data at different heavy masses and D_s momenta.

We chose not to perform tuning on $\{t_{\text{cut}}\}$ as was performed in the $B_s \rightarrow D_s^*$ study. This is because the fits are much larger than in the $B_s \rightarrow D_s^*$ case ($B_s \rightarrow D_s^*$ only required data at q_{max}^2), and tuning, which would involve many serial fits, would take a prohibitively long time. The t_{cut} 's we set are given in table 7.2.

set	name	H_s	\hat{H}_s	D_s	H_c	η_q	S	V_0
2	fine	2	2	2	2	2	2	2
3	fine-physical	4	4	4	5	4	2	2
4	superfine	5	5	5	10	5	4	4
5	ultrafine	2	2	2	8	2	4	4

Table 7.2: t_{cut} values used for each correlator on each ensemble. S and V_0 denote the 3-point correlators with the corresponding current S or V_0 . The rest are for 2-point correlators. There is one exception to the values here: on the 3-point correlators on the ultrafine ensemble with $am_{h0}^{\text{val}} = 0.8$ and $q^2 = 0$, the t_{cut} given here is replaced with 8, 10 and 12 in the $T/a = 31, 36, 41$ cases respectively. This is due to the signal/noise degradation from the large D_s momentum causing noise that makes data close to the H_s source useless to the fit.

These simultaneous fits are very large. For example, on set 4 (ultrafine) we

fit 109 correlation functions to 1080 fit parameters, taking all correlations in the data into account. Both the stability (e.g. invariance under changes of arbitrary hyperparameters such as N_{exp} and t_{cut}) and the speed of such simultaneous fits take a hit when there is such a large amount of data and a large number of parameters.

The stability issue is likely due to the size of the covariance matrix for the data, which must be inverted to estimate χ^2 . We took two steps toward mitigating this. The first is to impose an svd-cut. The second step we took was to employ a chained-fitting approach (this was required on the superfine and ultrafine data only). We first performed an array of smaller 'individual' fits, each fitting the correlators relevant only to one m_h and one $|a\mathbf{p}_{D_s}|$ value. In the case of set 4, for example, this results in 11 separate fits. Then, a full simultaneous fit of all of the correlators was carried out, with priors set to the results of the smaller fits. To be conservative we multiplied these priors by 1 ± 1.5 , i.e., the priors end up with over 150% variance. This both speeds up the full fit and improves the stability of the results. We tested the validity of the results by varying the additional fractional error between 100% and 200%, this caused negligible changes in the results. Since we did not need to take this measure on the fine ensemble, we performed both a standard simultaneous fit and chained fits of this type as a check for the chained fits. Tests 8, 9 and 10 on Fig. 7.1 show the results of these chained fits in comparison to the more traditional fit.

The priors for the 'traditional' fits to fine and fine-physical (sets 2 and 3) data, and individual chained superfine and ultrafine (sets 4 and 5) fits, were set up as follows. We set Gaussian priors for the parameters J_{jk} , and log-normal priors for all other parameters. Using log-normal distributions forbids ground state energies E_0^M , excited energy differences $E_n^M - E_{n-1}^M$, and amplitudes a_n^M from both becoming negative and moving arbitrarily close to zero, improving the stability of the fit. The priors we chose are summarized below.

$$\begin{aligned}
\log(\tilde{E}_0^M) &= \text{Empirical Bayes}, \\
\log(\tilde{E}_0^{M,o}) &= \mu_{\log(\tilde{E}_0^M)} \pm (\sigma_{\log(\tilde{E}_0^M)} \times 1.5), \\
\log(\tilde{E}_i^{M(o)} - \tilde{E}_{i-1}^{M(o)}) &= \log(2a\Lambda_{\text{QCD}} \pm a\Lambda_{\text{QCD}}), \quad i > 0, \\
\log(\tilde{a}_0^M) &= \text{Empirical Bayes}, \\
\log(\tilde{a}_i^M) &= -1.9(3.3), \quad i > 0, \\
\log(\tilde{a}_i^{M,o}) &= -3.0(2.0), \quad i > 0, \\
J_{jk} &= 0 \pm 1 \text{ except for } J_{00}^{nn} = 1 \pm 0.5
\end{aligned} \tag{7.17}$$

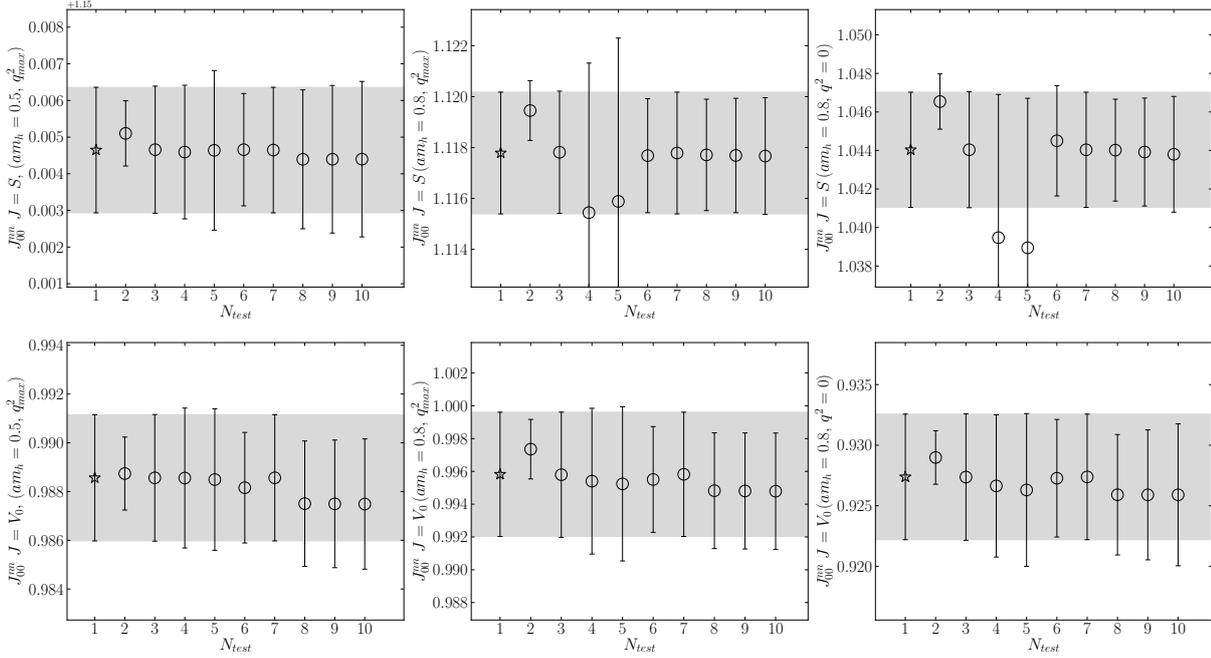


Figure 7.1: Tests on the correlator fits on the fine ensemble. $N_{\text{test}} = 1$ gives the final accepted result. $N_{\text{test}} = 2$ and 3 gives the results of setting $N_{\text{exp}} = 4$ and 6 respectively. $N_{\text{test}} = 4$ gives the results when all priors are given a standard deviation one and a half times the size of the final fit. $N_{\text{test}} = 5$ gives the result of setting $t_{\text{cut}} = 4$ rather than 2 for all correlators. $N_{\text{test}} = 6$ gives the result without marginalising out the $n = 5$ excited state. $N_{\text{test}} = 7$ gives the result of moving the svd cut from 10^{-3} to 10^{-2} . $N_{\text{test}} = 8, 9, 10$ gives the result of using a chained fit described above with priors for the full correlated fit given additional fractional errors of 100%, 150% and 200% respectively.

Priors for ground state energies $\log(E_0^M)$ and amplitudes $\log(a_0^M)$ are set according to an empirical-Bayes approach, plots of effective energies and amplitudes of the correlation functions are inspected to deduce reasonable priors. The ground-state oscillating parameters $\log(a_0^{M,o})$, $\log(E_0^{M,o})$, are given the same priors as the non-oscillating states, with errors inflated by 50%. The resulting priors always have a standard deviation of at least 10 times that of the final result. The log of oscillating and non-oscillating excited state energies, $\log(E_i^{M,(o)} - E_{i-1}^{M,(0)})$, $i > 0$ are given prior values of $\log(2\Lambda_{\text{QCD}} \pm \Lambda_{\text{QCD}})$. We set $\Lambda_{\text{QCD}} = 0.5\text{GeV}$. The excited state amplitudes $\log(a_i^M)$, $i > 0$ are given priors of -1.9 ± 3.3 for non-oscillating states, and -3.0 ± 2.0 for oscillating states. The ground-state non-oscillating to non-oscillating 3-point parameter, J_{00}^{nn} is given a prior of 1 ± 0.5 , and the rest of the 3-point parameters J_{jk}^{nn} are given 0 ± 1 .

The current matrix elements we require can be extracted from the fit parameters via

$$\langle D_s | J | H_s \rangle |_{\text{lat}} = 2\sqrt{M_{H_s} E_{D_s}} J_{00}^{nn}. \quad (7.18)$$

7.2.3 Vector Current Renormalization

In HISQ, the local scalar current ($1 \otimes 1$) (multiplied by the mass difference of flavours it is charged under) requires no renormalization due to its connection to the partially conserved vector current through the PCVC relation. This is not the case for the local temporal vector current ($\gamma_0 \otimes \gamma_0$). The partially conserved vector current is a complicated linear combination of many local and point-split lattice currents. In this calculation we use only the local part of the vector current, this improves the statistics of our results but creates the need for the resulting current matrix element to be multiplied by a matching factor Z_V to produce the appropriate continuum current. We found Z_V via a fully non-perturbative method [192].

When both meson states in the matrix elements are at rest (the zero recoil point), the scalar and local vector matrix elements are related via the PCVC relation:

$$(M_{H_s} - M_{D_s})Z_V \langle D_s | V_0 | \hat{H}_s \rangle |_{\text{lat}}(q_{\text{max}}^2) = (m_{h0}^{\text{val}} - m_{c0}^{\text{val}}) \langle D_s | S | H_s \rangle |_{\text{lat}}(q_{\text{max}}^2). \quad (7.19)$$

Z_V can be extracted from this relation since the matrix elements are already computed as part of the calculation. Our calculation is self-renormalizing, in the sense that the normalization can be found at no extra computational cost. The Z_V values found on each ensemble and am_{h0}^{val} are given in Table 7.3.

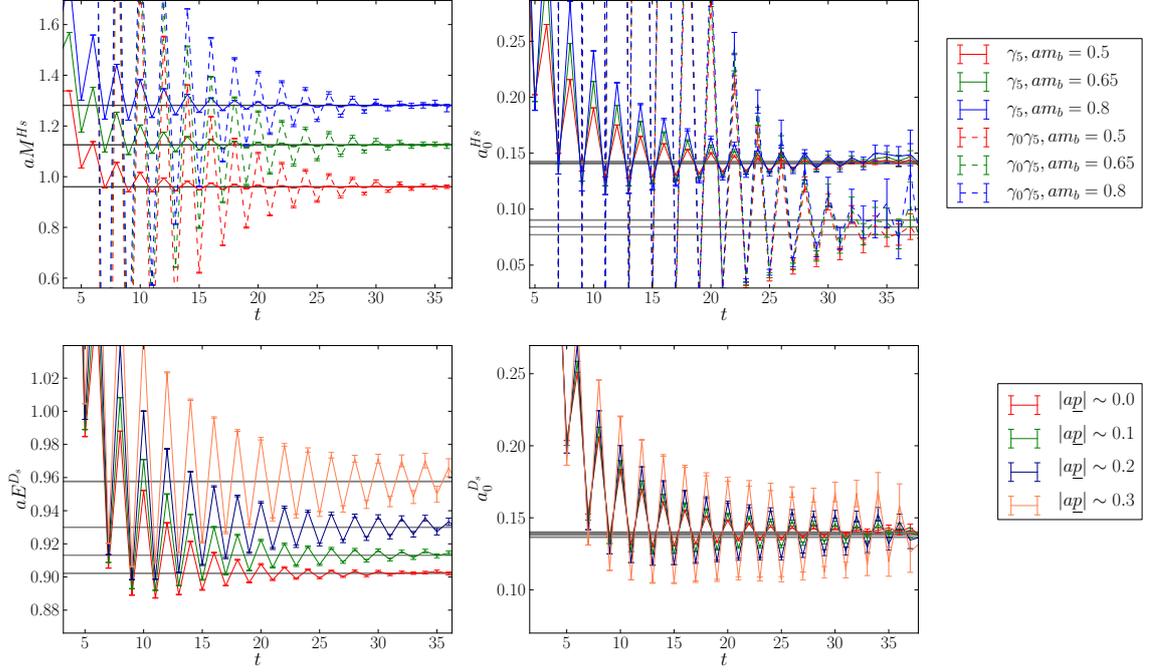


Figure 7.2: Superaveraged effective energies (given by Eq. (6.14)) and amplitudes (given by Eq. (6.15)) for a selection of 2-point correlators on the fine ensemble. The grey bands show the corresponding results for these quantities from the simultaneous Bayesian fits. These plots supply a further check of our correlator fits - results of the Bayesian fits are in good agreement with the effective energies and amplitudes.

We also removed tree-level mass-dependent discretization effects using a normalization constant Z_{disc} defined in Eq. (6.17).

Combining these normalizations with the lattice current from the correlation function fits, we found values for the form factors at a given heavy mass, lattice spacing, and q^2 :

$$\begin{aligned}
 f_0^s(q^2) &= \frac{m_{h0}^{\text{val}} - m_{c0}^{\text{val}}}{M_{H_s}^2 - M_{D_s}^2} Z_{\text{disc}} \langle D_s | S | H_s \rangle_{\text{lat}}(q^2) \\
 f_+^s(q^2) &= \frac{Z_{\text{disc}}}{2M_{H_s}} \times \\
 &\quad \frac{\delta^M \langle D_s | S | B_s \rangle_{\text{lat}}(q^2) - q^2 Z_V \langle D_s | V_0 | B_s \rangle_{\text{lat}}(q^2)}{\mathbf{p}_{D_s}^2}.
 \end{aligned} \tag{7.20}$$

I have here explicitly denoted the dependence of the matrix elements on q^2 as a

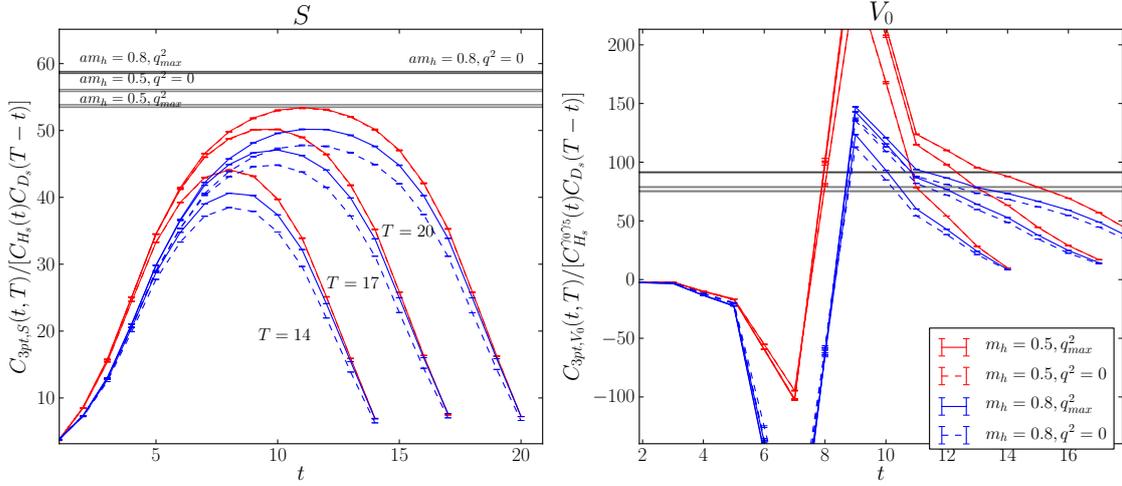


Figure 7.3: $\tilde{C}_3(t, T)/\tilde{C}^{H_s}(t)\tilde{C}^{D_s}(T-t)$, which should plateau at J_{00}^{nn} , in the fine ensemble for the $J = S$ and $J = V_0$ cases. The grey bands show the corresponding results for these quantities from the simultaneous Bayesian fits. Unfortunately, in the V_0 case the oscillating component dominates the correlation function, preventing any plateau from being visible.

reminder that the matrix elements have q^2 dependence via the D_s momentum.

7.2.4 Extrapolation to the Physical Point

I will now address the extrapolation of the $f_0^s(q^2)$ and $f_+^s(q^2)$ values to continuum, physical quark masses and arbitrary q^2 . We took two complementary approaches to the extrapolation.

One we refer to as the *ratio* approach, in which one extrapolates the quantity

$$R_{0,+}^s(q^2) \equiv \frac{f_{0,+}^s(q^2)}{f_{H_c}\sqrt{M_{H_c}}}. \quad (7.21)$$

to the physical point. Discretisation effects appear to cancel to a large extent in this ratio, resulting in a better controlled continuum extrapolation. The value at the physical point is then multiplied by $f_{B_c}\sqrt{M_{B_c}}$ to isolate the form factors, where we find f_{B_c} via a separate extrapolation (detailed in Sec. 6.3.4), and take the PDG value for M_{B_c} [3]. While this approach improves the continuum extrapolation, it has the downside of introducing errors from scale-setting on account of $R_{0,+}^s(q^2)$ being dimensionful quantities (as opposed to $f_{0,+}^s(q^2)$ which are dimensionless).

Set	am_h^{val}	Z_V	Z_{disc}
2	0.5	1.0151(32)	0.99819
	0.65	1.0240(37)	0.99635
	0.8	1.0368(49)	0.99305
3	0.5	1.0134(24)	0.99829
	0.8	1.0348(29)	0.99315
4	0.427	1.0027(25)	0.99931
	0.525	1.0059(29)	0.99859
	0.65	1.0108(37)	0.99697
	0.8	1.0197(49)	0.99367
5	0.5	1.0037(40)	0.99889
	0.65	1.0087(46)	0.99704
	0.8	1.0160(53)	0.99375

Table 7.3: Normalization constants applied to the lattice axial vector current in (7.20). Z_V is found from (7.19) and Z_{disc} from (6.17).

In the other method, that we refer to as the *direct* approach, one simply extrapolates the form factors to the physical point. The form factors by themselves have larger discretisation effects than $R_{0,+}^s(q^2)$, but since $f_{0,+}^s$ are dimensionless, the results are completely insensitive to scale-setting uncertainty.

We use identical fit functions for both approaches. In the below discussion, we use the notation $F_{0,+}^s(q^2)$ to denote either $f_{0,+}^s(q^2)$ or $R_{0,+}^s(q^2)$, depending on the approach being applied.

Kinematic Behaviour

Our fit form for the extrapolation is a modified version of the Bourely-Caprini-Lellouch (BCL) parameterization for pseudoscalar \rightarrow pseudoscalar form factors [193]:

$$F_0^s(q^2)|_{\text{fit}} = \frac{1}{1 - \frac{q^2}{M_{H_c^0}^2}} \sum_{n=0}^{N-1} a_n^0 z^n(q^2), \quad (7.22)$$

$$F_+^s(q^2)|_{\text{fit}} = \frac{1}{1 - \frac{q^2}{M_{H_c^*}^2}} \sum_{n=0}^{N-1} a_n^+ \left(z^n(q^2) - \frac{n}{N} (-1)^{n-N} z^N(q^2) \right).$$

The function $z(q^2)$ maps q^2 to a small region inside the unit circle on the complex plane, defined by

$$z(q^2) = \frac{\sqrt{t_+ - q^2} - \sqrt{t_+ - t_0}}{\sqrt{t_+ - q^2} + \sqrt{t_+ - t_0}}, \quad (7.23)$$

where $t_+ = (M_{H_s} + M_{D_s})^2$, and we chose t_0 to be $t_0 = 0$. This t_0 choice means that at $q^2 = 0$ the fit functions simplify to $F_{0,+}^s(0) = a_0^{0,+}$. Throughout the physical range of q^2 , z is restricted to the range $0 < z < 0.06$, resulting in a fast converging series in powers of z . We truncate at $N = 3$, adding further orders of z^n does not affect the results of the fit.

The factors in front of the sums in the BCL parameterization account for sub-threshold poles in the form factors due to the production of on-shell H_{c0} and H_c^* states in the crossed channel of the semileptonic decay.

To estimate $M_{H_{c0}}$, the scalar heavy-charm meson mass, at each of the heavy masses we used, we leveraged the fact that the splitting $\Delta_0 = M_{H_{c0}} - M_{H_c}$ is due to an orbital excitation and therefore independent of the heavy quark mass. This has been calculated in [194] to be $\Delta_0 = 0.429(13)\text{GeV}$. Combined with an H_c mass from our correlators, we can construct the H_{c0} mass: $M_{H_{c0}} = M_{H_c} + \Delta_0$. We did

not take the error on Δ_0 into account in the fit since the precise position of the pole has a small effect on the fit results.

To estimate $M_{H_c^*}$, the vector heavy-charm mass, we use the fact that the splitting $M_{H_c^*} - M_{H_c}$ vanishes in the infinite heavy mass limit. $M_{H_c^*}$ then takes the approximate form $M_{H_c^*} \simeq M_{H_c} + \mathcal{O}(1/m_h)$. To reproduce this behaviour we use the ansatz $M_{H_c^*} = M_{H_c} + x/M_{\eta_h}$, and fix x at the physical point to find $x = 0.508 \text{ GeV}^2$. To do this we used the result $M_{B_c^*} - M_{B_c} = 54(3) \text{ MeV}$ from [194].

Heavy Mass and Discretisation Effects

To account for variation in heavy mass and discretisation effects in a general way, we gave the following form to each of the $a_n^{0,+}$ coefficients:

$$a_n^{0,+} = \left(1 + \rho_n^{0,+} \log \left(\frac{M_{\eta_c}}{M_{\eta_h}} \right) \right) \times \sum_{i,j,k=0}^{2,2,2} d_{ijkn}^{0,+} \left(\frac{2\Lambda_{\text{QCD}}}{M_{\eta_h}} \right)^i \left(\frac{am_{h0}^{\text{val}}}{\pi} \right)^{2j} \left(\frac{aE_{D_s}}{\pi} \right)^{2k} \times \left(1 + \mathcal{N}_{\text{mistuning},n}^{0,+} \right). \quad (7.24)$$

To understand this form, focus first on the sum. Powers of $(2\Lambda_{\text{QCD}}/M_{\eta_h})$ give an HQET inspired way of quantifying the variation in the results due to the changing heavy mass. M_{η_h} varies strongly and monotonically with the heavy quark mass, so acts as a suitable proxy. Λ_{QCD} is the confinement scale, which we set to 0.5 GeV .

The two scales expected to be the largest sources of discretisation effects are the heavy mass am_{h0}^{val} , and the energy in the D_s meson, aE_{D_s} , especially when it is given a large spatial momentum. Adding further, smaller scales, like $a\Lambda_{\text{QCD}}$, had no effect on the results.

The coefficients $d_{ijkn}^{0,+}$ are fit parameters given prior distributions of 0 ± 2 . In the ratio case, these carry mass dimension $\text{GeV}^{-3/2}$, hence this prior corresponds to a prior of $0 \pm (2\Lambda_{\text{QCD}})^{-3/2}$.

To account for any required matching between HQET and QCD, we included a log term in front of the sum. $\rho_n^{0,+}$ are fit parameters with prior distribution 0 ± 1 .

The fact that $f_+^s(0) = f_0^s(0) (\Rightarrow a_0^+ = a_0^0)$ is a very powerful constraint within the heavy-HISQ approach. Since this must be true at all m_h , this translates to constraints in the fit parameters: $d_{i000}^+ = d_{i000}^0 \forall i$ and $\rho_0^+ = \rho_0^0$. We imposed these constraints in the fit, which serve to stabilize the extrapolation in the heavy mass direction.

Mass Mistunings

We dealt with possible mistuning in the c , s and l masses in the same way as in the $B_s \rightarrow D_s^*$ study. We included the terms $\mathcal{N}_{\text{mistuning},n}^{0,+}$ in each $a_n^{0,+}$ coefficient, given by

$$\mathcal{N}_{\text{mistuning},n} = \frac{c_{s,n}^{\text{val}}\delta_s^{\text{val}} + c_{s,n}\delta_s + 2c_{l,n}\delta_l}{10m_s^{\text{tuned}}} + c_{c,n}^{0,+} \left(\frac{M_{\eta_c} - M_{\eta_c}^{\text{phys}}}{M_{\eta_c}^{\text{phys}}} \right), \quad (7.25)$$

where $c_{l,n}^{0,+}$, $c_s^{0,+}$ and $c_{ci}^{0,+}$ are fit parameters with prior distributions 0 ± 1 . $\delta_{s,l}^{(\text{val})}$ are defined in Sec. 6.2.4 and m_s^{tuned} is defined in Eq. (6.26).

All higher order contributions, like $\delta_{s,l}^{(\text{val})2}$ or $(M_{\eta_c} - M_{\eta_c}^{\text{phys}})^2$ are too small to be resolved by our lattice data, so are not included in the fit.

Negligible Effects

Finite volume effects are negligible in our calculation, we do not include any associated error. Finite volume corrections to the $B \rightarrow D\ell\nu$ form factors in chiral perturbation theory were calculated in [172]. They found the $B \rightarrow D\ell\nu$ form factor at zero recoil, with a lattice size of $L = 2.5\text{fm}$, and pion mass equal to or greater than physical, never exceeded 10^{-4} . There is no reason to believe changing the spectator quark from light to strange, and moving away from zero recoil, will increase this effect.

In our simulation we set $m_u = m_d \equiv m_l$, this means our results do not account for isospin breaking. By moving the m_l^{tuned} value up and down by the PDG value for $m_d - m_u$, we tested for any signs of isospin breaking having an effect on the results. The resulting effect was negligible in comparison to all other sources of error.

Since we take the physical result at $M_{\eta_n} = M_{\eta_b}$, the uncertainty in M_{η_b} will contribute an uncertainty in the final result. We used the PDG result for M_{η_b} [3]. However, $\bar{b}-b$ annihilation and electroweak corrections make this somewhat different to the appropriate value on the lattice. We estimated the corresponding uncertainty in M_{η_b} to be no greater than $\pm 10\text{MeV}$. To see how this changes the result of the extrapolation, we varied M_{η_b} up and down by 10MeV and studied how our result at $f_0(q_{\text{max}}^2)$ changes. The change is never greater than 10^{-4} , which is negligible in comparison to the other errors.

7.3 Results

Tables 7.9 and 7.10 in Sec. 7.6 give numerical values for the form factors, the ratios $R_{0,+}^s(q^2)$, and parameters extracted from the correlation function fits required for the extrapolations to the physical point.

I will first show results from the ratio method, then the direct method. In both cases, we performed a simpler fit at zero recoil first, then a larger fit taking into account all data throughout q^2 . In each case, our results are statistics dominated. The results from the two methods are in good agreement.

7.3.1 Ratio Method

Zero Recoil

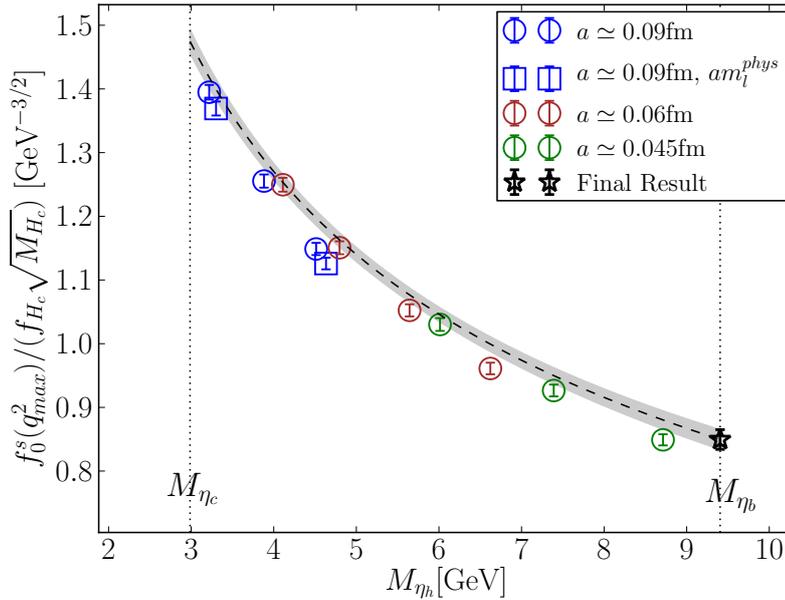


Figure 7.4: $R_0^s(q_{\max}^2) = f_0^s(q_{\max}^2)/(f_{H_c} \sqrt{M_{H_c}})$ against M_{η_h} (a proxy for the heavy quark mass). The grey band shows the result of the extrapolation at $a = 0$ and physical l, s and c masses. Sets listed in the legend follow the order of sets in table 4.1.

We performed an isolated extrapolation of $R_0^s(q_{\max}^2)$ to the physical point. To do this we used a simplified fit form for $R_0^s(q_{\max}^2)$ consisting of the right hand side

of (7.24), with the index n discarded. We find

$$R_0^s(q_{\max}^2) = \frac{f_0^s(q_{\max}^2)}{f_{B_c} \sqrt{M_{B_c}}} = 0.843(18) \text{GeV}^{-3/2}. \quad (7.26)$$

The extrapolation against M_{η_h} is illustrated in Fig. 7.4. As can be seen here, data for this ratio has a very weak dependence on the lattice spacing. The error budget for this result is given in Table 7.4.

Source	% Fractional Error
Scale Setting	1.35
Statistics	1.07
$a \rightarrow 0$	0.16
$m_h \rightarrow m_b$,	0.87
mistuning	0.72
Total	2.16

Table 7.4: Error budget for $R_0^s(q_{\max}^2)$. The value for scale setting is given by the partial standard deviation of $R_0^s(q_{\max}^2)$ with respect to w_0 and w_0/a . The value for statistics is given by the partial standard deviation of $R_0^s(q_{\max}^2)$ with respect to the lattice data (with scale setting error subtracted in quadrature). The value for $a \rightarrow 0$ is the partial standard deviation of $R_0^s(q_{\max}^2)$ with respect to priors of the fit parameters d_{0jk}^0 for $i + j \neq 0$. Similarly for $m_h \rightarrow m_b$, the value is the partial standard deviation with respect to priors of d_{ijk}^0 for $k \neq 0$. The mistuning value comes from priors of parameters in $\mathcal{N}_{\text{mistuning}}$.

We performed a number of tests on this zero recoil extrapolation to test the stability of our fit form, results are given in Fig. 7.5.

Non-zero Recoil

Fig. 7.6 shows the result of the full extrapolation of $R_{0,+}^s(q^2)$ throughout the q^2 range described in Sec. 7.2.4. As the heavy mass increases, the q^2 range, $0 < (M_{H_s} - M_{D_s})^2$ expands.

To isolate the form factors, the resulting functions $R_{0,+}^s(q^2)$ were multiplied by $\sqrt{M_{B_c}}$ (using the PDG value) and f_{B_c} from our determination detailed in Sec. 6.3.4.

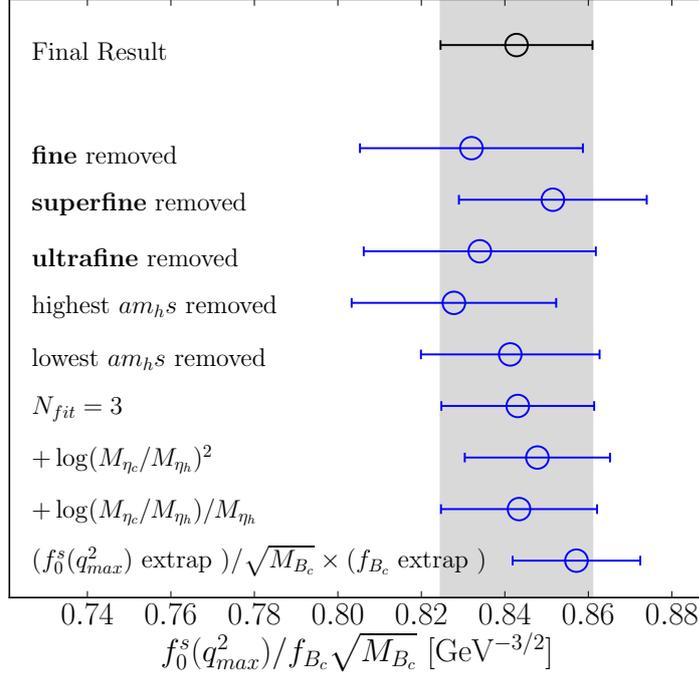


Figure 7.5: Results of tests of the $R_0^s(q_{\max}^2)$ extrapolation. The top three blue points show $R_0^s(q_{\max}^2)$ at continuum and physical quark mass, if data from the fine, superfine or ultrafine ensembles are not used in the fit. The fourth and fifth blue points show the result if data at the highest/lowest am_{h0}^{val} value on each ensemble are removed. The point labelled $N_{fit} = 3$ is the result of extending the sum in (7.24) such that it truncates at 3 rather than 2 in each of the i, j, k directions. The point labelled $+\log(M_{\eta_h}/M_{\eta_c})^2$ represents the result of adding a $\rho_{2,n} \log(M_{\eta_h}/M_{\eta_c})^2$ term in the first set of brackets in (7.24), where $\rho_{2,n}$ are new fit parameters with the same prior distributions as ρ_n . Similarly $+\log(M_{\eta_h}/M_{\eta_c})/M_{\eta_h}$ shows the result of adding this term multiplied by $\rho_{2,n}$. The lowest point shows the result of our direct extrapolation of $f_0^s(q_{\max}^2)$ to the physical point, divided by the PDG value for $\sqrt{M_{B_c}}$ [3] and the result of our extrapolation of f_{B_c} to the physical point detailed in Sec. 6.3.4.

The resulting form factors are shown in Fig. 7.12, against the form factors found via the direct method.

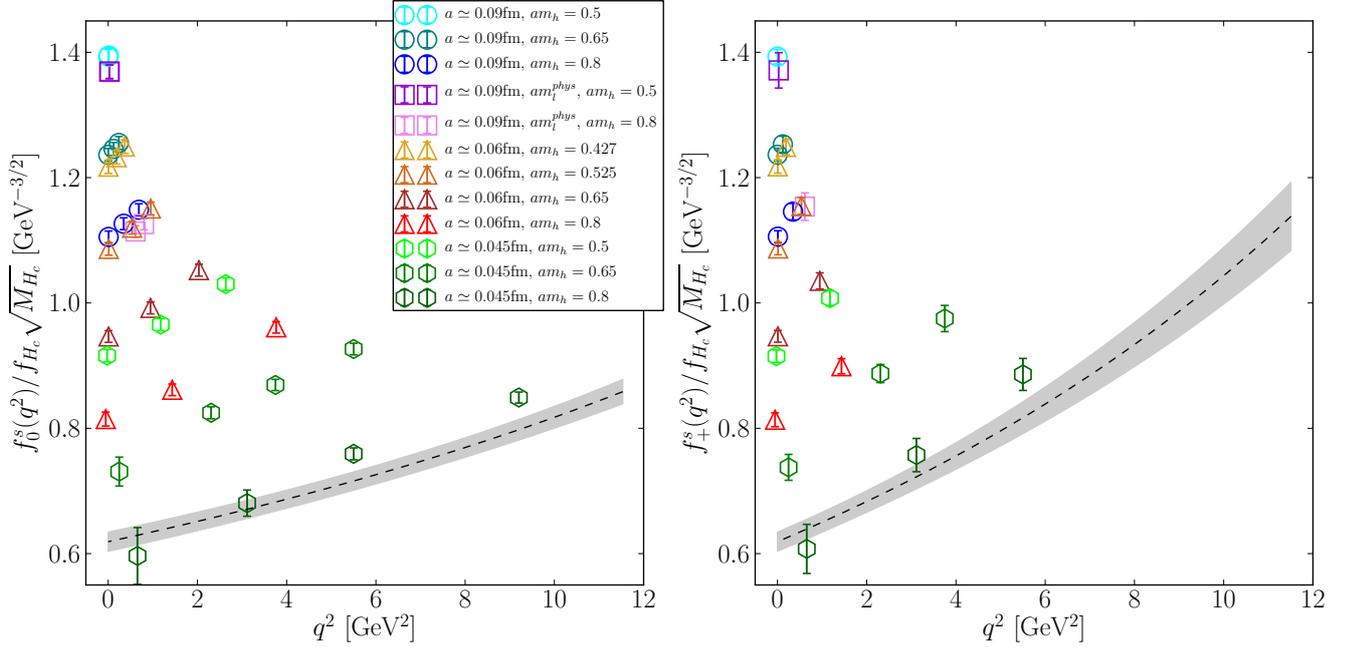


Figure 7.6: $R_{0,+}^s(q^2) = f_{0,+}^s(q^2)/(f_{H_c} \sqrt{M_{H_c}})$ against q^2 . The grey band shows the result of the extrapolation at $a = 0$ and physical quark masses. Sets listed in the legend follow the order of sets in table 4.1.

7.3.2 Direct Method

Zero Recoil

We performed an isolated extrapolation of $f_0^s(q_{\max}^2)$ to the physical point. Once again, this was performed using a fit function for $f_0^s(q_{\max}^2)$ consisting of the right hand side of Eq. (7.24) with the index n discarded. We find

$$f_0^s(q_{\max}^2) = 0.899(13). \quad (7.27)$$

The extrapolation against M_{η_h} is shown in Fig. 7.7. The error budget for this result is given in Table 7.5.

We include in Fig. 7.7 a previous lattice determination of this quantity [152], shown as a red triangle. Our two studies, containing largely independent systematic uncertainties, are in agreement. The previous study used the $N_f = 2 + 1$ MILC gluon ensembles, with HISQ s and c valence quarks, and an NRQCD b quark. Using NRQCD meant they could perform their simulation directly at the physical b mass. However, the matching of lattice NRQCD to continuum QCD causes their

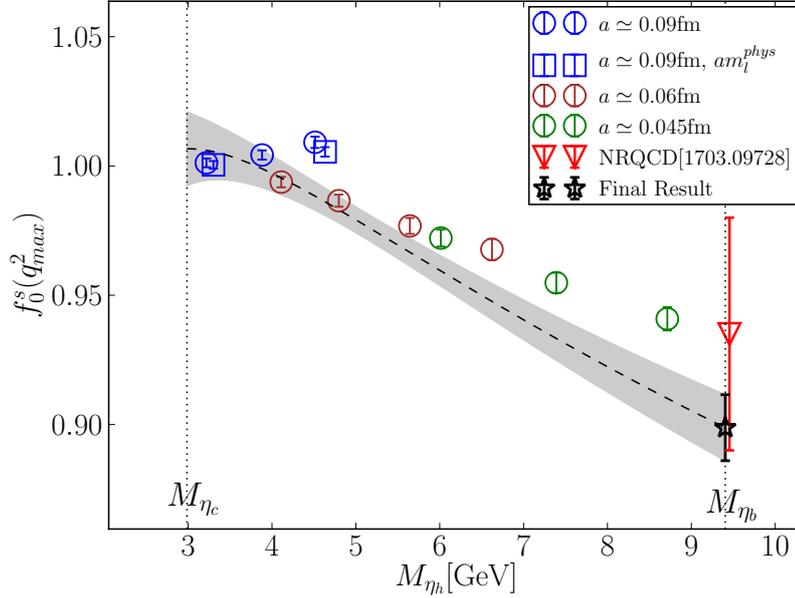


Figure 7.7: $f_0^s(q_{\max}^2)$ against M_{η_h} (a proxy for the heavy quark mass). The grey band shows the result of the extrapolation at $a = 0$ and physical quark masses. We also include the result from a previous lattice calculation, which used the NRQCD discretisation for the b quark [152]. Sets listed in the legend follow the order of sets in Table 4.1.

Source	% Fractional Error
Statistics	1.04
$m_h \rightarrow m_b$,	0.75
$a \rightarrow 0$	0.27
mistuning	0.40
Total	1.42

Table 7.5: Error budget for $f_0^s(q_{\max}^2)$ found via the direct method. These values are determined in the same way as for $R_0^s(q_{\max}^2)$, described in the caption of Table 7.4. No scale setting error is listed since this is less than 0.01% on account of $f_0^s(q_{\max}^2)$ being dimensionless.

dominant error.

As when using the ratio method, we performed a number of tests on this extrapolation at zero recoil, and present results in Fig. 7.8.

Non-Zero Recoil

Fig. 7.9 shows the result of the full extrapolation of the ratio throughout the q^2 range described in Sec. 7.2.4.

Fig. 7.10 shows the resulting form factors from the direct approach. In Fig. 7.11 we give an associated error budget for these throughout q^2 . Statistical errors in f_+^s grow in the $q^2 \rightarrow q_{\max}^2$ region due to the 'fine tuning' effect discussed in Sec. 5.3.5. The effect is not as severe as in the NRQCD case since the lattice data we obtain for $f_+^s(q^2)$ is much further away from the q_{\max}^2 point.

We take the results from the direct method as our final result, and supply the ratio method results as a consistency test, since the product of the direct method is more precise. Fig. 7.12 shows the form factors resulting from the two methods on top of each other. As one can see from this plot, the results are in good agreement for all physical q^2 values.

In Figure 7.13, we show our final results (direct approach) against lattice form factors determined from the NRQCD calculation mentioned in Sec. 7.3.2 [152]. Our results are in excellent agreement with the NRQCD calculation, and are more precise for both $f_0^s(q^2)$ and $f_+^s(q^2)$ throughout all q^2 .

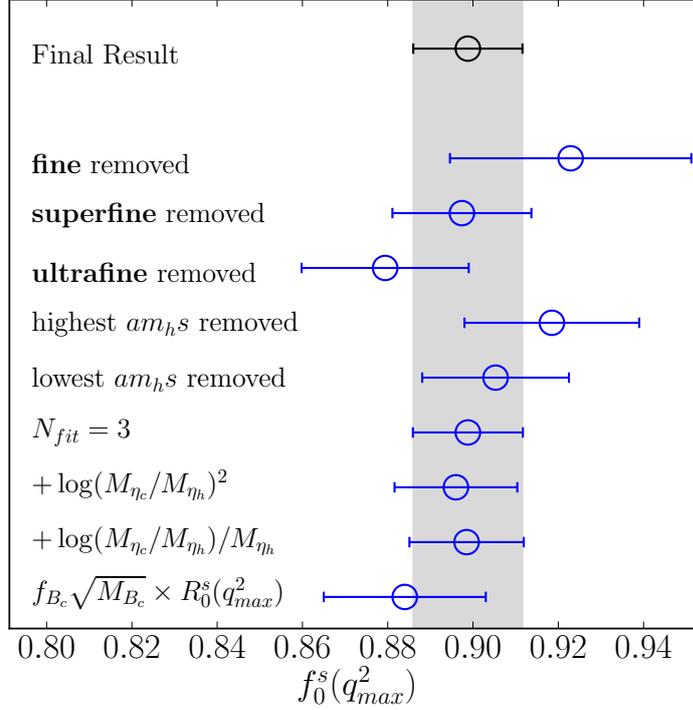


Figure 7.8: Results of tests of the $f_0^s(q_{max}^2)$ extrapolation. The top three blue points show $f_0^s(q_{max}^2)$ at continuum and physical quark mass, if data from the fine, superfine or ultrafine ensembles are not used in the fit. The fourth and fifth blue points show the result if data at the highest/lowest am_{h0}^{val} value on each ensemble are removed. The point labelled $N_{fit} = 3$ is the result of extending the sum in (7.24) such that it truncates at 3 rather than 2 in each of the i, j, k directions. The point labelled $+\log(M_{\eta_h}/M_{\eta_c})^2$ represents the result of adding a $\rho_2 \log(M_{\eta_h}/M_{\eta_c})^2$ term in the first set of brackets in (7.24), where $\rho_{2,n}$ are new fit parameters with the same prior distributions as ρ_n . Similarly, the point labelled $+\log(M_{\eta_c}/M_{\eta_h})/M_{\eta_h}$ gives the result of adding this term multiplied by $\rho_{2,n}$. The lowest point shows the result from the extrapolation of $R_0^s(q_{max}^2)$, multiplied by the PDG value for $\sqrt{M_{B_c}}$ [3] and the result of our extrapolation of f_{B_c} to the physical point detailed Sec. 6.3.4.

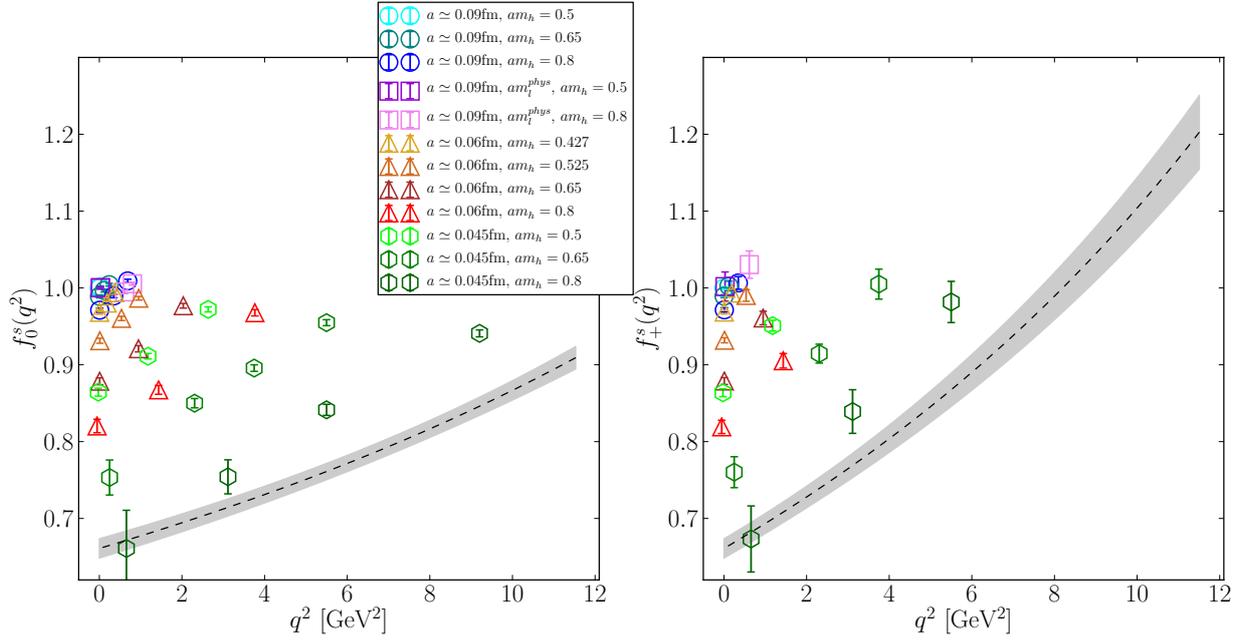


Figure 7.9: $f_{0,+}^s(q^2)$ against q^2 . The grey band shows the result of the extrapolation at $a = 0$ and physical quark masses. Sets listed in the legend follow the order of sets in Table 4.1.

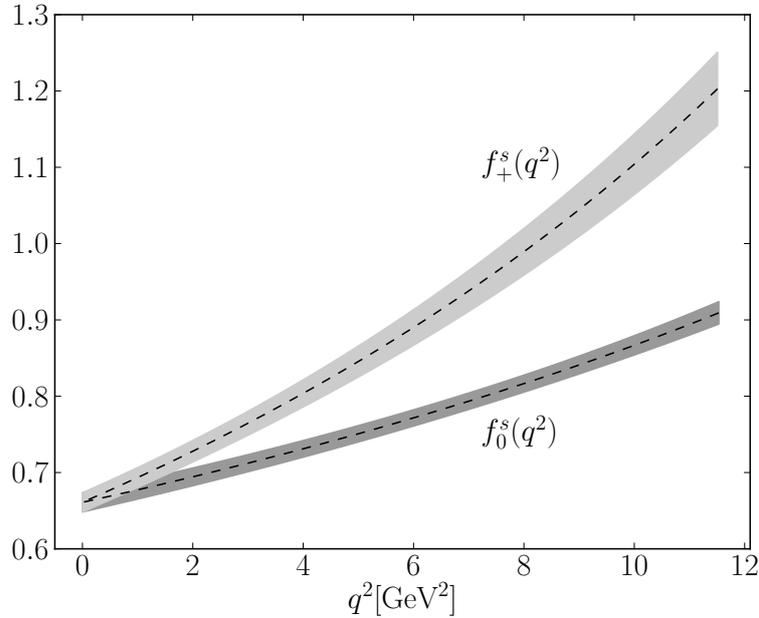


Figure 7.10: Final result for $f_{0,+}^s(q^2)$ against q^2 at the physical point .

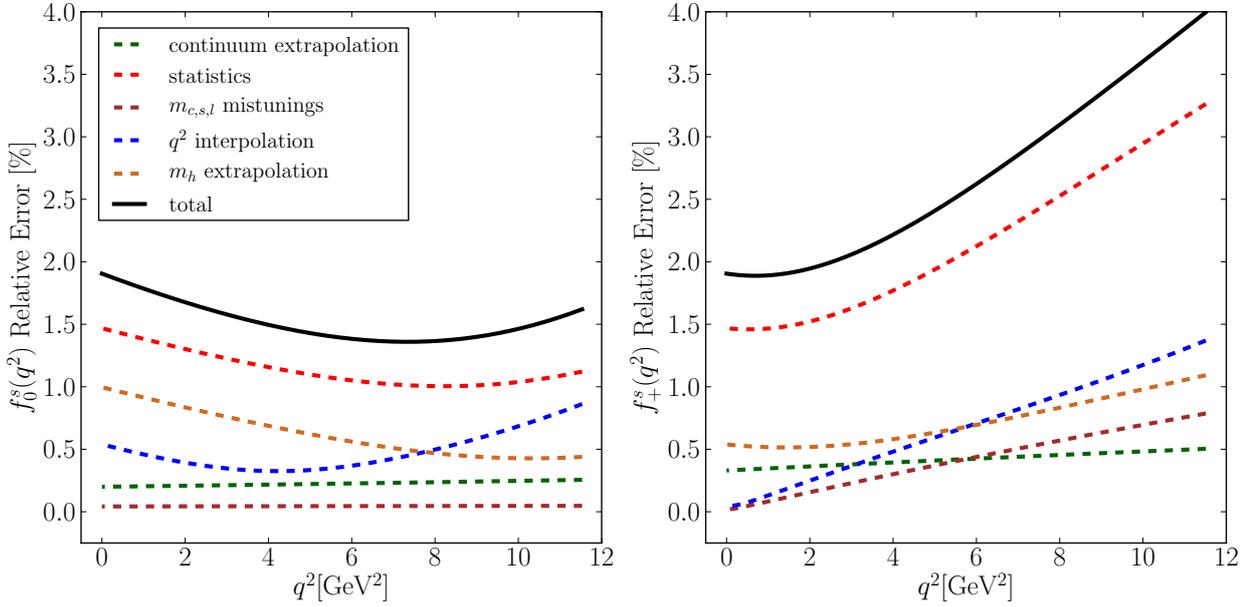


Figure 7.11: Error budget for $f_{0,+}^s(q^2)$ against q^2 . Values are determined via approach explained in the caption of Table 7.4.

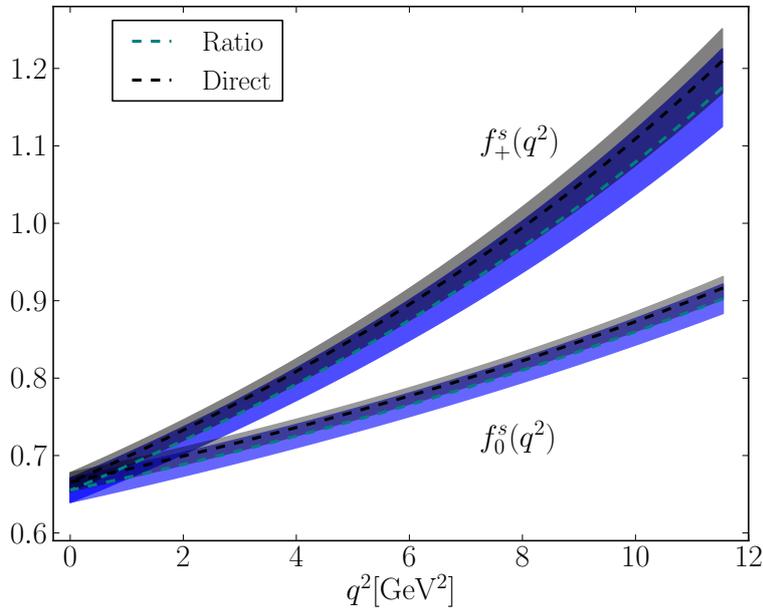


Figure 7.12: Results for $f_{0,+}^s(q^2)$ against q^2 at the physical point, from both the ratio method and the direct method.

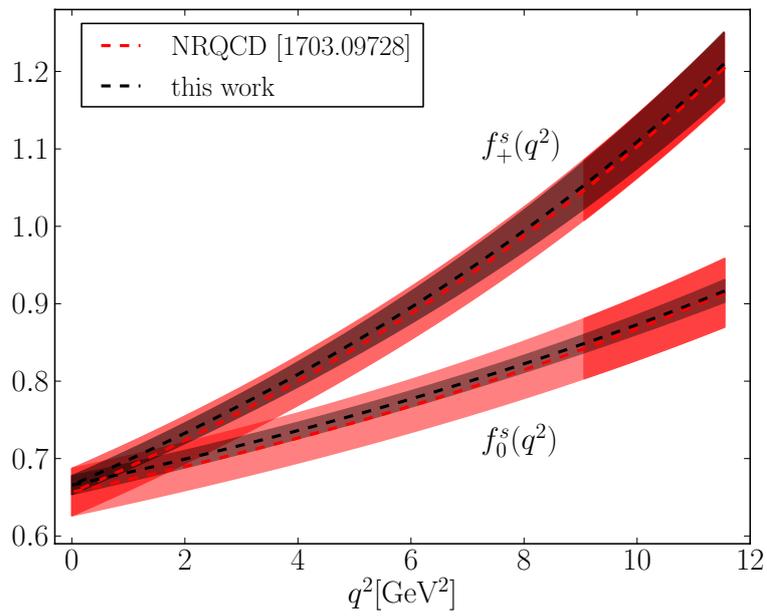


Figure 7.13: Our final result for $f_{0,+}^s(q^2)$ against form factors calculated from a previous study using the NRQCD action for the b quark [152]. The darker shaded region of the NRQCD band shows where lattice data was available in that study, the rest of the band shows the result of an extrapolation in q^2 using the BCL parameterization.

7.3.3 Unitarity Test

Unitarity and crossing symmetry impose bounds on the coefficients of the BCL parameterization of $f_{0,+}^s(q^2)$, $\{a_n\}$ [195, 196]. As another consistency test, we show here that the coefficients found in our fit satisfy these bounds.

To obtain bounds on the BCL coefficients, one must relate them to those of a different parameterization, that of Boyd, Grinstein and Lebed (BGL) [48]:

$$f^s(q^2) = \frac{1}{B(z)\phi(z)} \sum_{n \geq 0}^N b_n z^n. \quad (7.28)$$

$B(z)$ is known as the Blaschke factor:

$$B(z) = \frac{z - z_*}{1 - z z_*}, \quad (7.29)$$

where $z_* = z(M_{B_c}^2)$ for f_0^s , or $z(M_{B_c}^2)$ for f_+^s . $\phi(z)$ is the outer function:

$$\begin{aligned} \phi(z) &= M_{B_s}^{2-s} 2^{2+p} \sqrt{\kappa n_f} \left[\frac{M_{D_s}}{M_{B_s}} (1+z) \right]^{s-3/2} \\ &\times \left[(1-z) \left(1 + \frac{M_{D_s}}{M_{B_s}} \right) + 2 \sqrt{\frac{M_{D_s}}{M_{B_s}}} (1+z) \right]^{-s-p}. \end{aligned} \quad (7.30)$$

In the f_0^s case, $\kappa = 12\pi M_{B_s}^2 \chi_A$, $p = 1, s = 3$. In the f_+^s case, $\kappa = 6\pi M_{B_s}^2 \chi_V$, $p = 3, s = 2$. The quantities $\chi_{V,A}$ are the once-subtracted dispersion relations at $q^2 = 0$ for vector and axial $b \rightarrow c$ currents respectively, computed in [48] to be $\chi_V = 5.7 \times 10^{-3}/m_b^2$ and $\chi_A = 9.6 \times 10^{-3}/m_b^2$.

The BGL coefficients, $\{b_n\}$, obey the unitarity constraint

$$\sum_{m=0}^{\infty} |b_m|^2 \leq 1 \quad (7.31)$$

by construction of the parameterization. To see how this applies to the BCL coefficients $\{a_n\}$, one must relate them to $\{b_m\}$ by equating the two parameterizations to find

$$\sum_{m=0}^M b_m z^m = \psi(z) \sum_{n=0}^N a_n z^n, \quad (7.32)$$

where $\psi(z)$ is given by

$$\psi(z) = \frac{M_{\text{pole}}^2}{4(t_+ - t_0)} \phi(z) \frac{(1-z)^2 (1-z_*)^2}{(1-z z_*)^2}, \quad (7.33)$$

where $M_{\text{pole}} = M_{B_c^0}$ in the f_0^s case and $M_{B_c^*}$ in the f_+^s case. Expanding $\psi(z)$ around $z = 0$, comparing coefficients of z in (7.32), and imposing the constraint (7.31), we arrive at a constraint for the BCL coefficients

$$\mathcal{B} \equiv \sum_{j,k=0}^{L,L} B_{jk} a_j a_k \leq 1, \quad (7.34)$$

$$B_{jk} = \sum_{n=0}^{\infty} \eta_n \eta_{n+|j-k|}. \quad (7.35)$$

where $\{\eta_n\}$ are the Taylor coefficients of $\psi(z)$.

$\psi(z)$ is bounded on the closed disk $|z| < 1$, so its Taylor coefficients are rapidly decreasing. We computed values for B_{jk} by truncating the sum in its definition (7.35) at 100. These values are given in Table 7.6. With these B_{jk} values, and the a_n coefficients at the physical point from our fit (via the direct method), we find

$$\mathcal{B}_0 = 0.0008(15),$$

$$\mathcal{B}_+ = 0.0204(66).$$

These comfortably satisfy the unitarity bound. Additionally, as discussed in [197], the leading contributions to $\mathcal{B}_{0,+}$ are of order $(\Lambda_{\text{QCD}}/m_b)^3 \simeq 10^{-3}$ in HQET. This expectation is approximately fulfilled by our result.

	B_{00}	B_{01}	B_{02}
f_0^s	0.00186	-0.000258	-0.000703
f_+^s	0.00179	-0.000367	0.00108

Table 7.6: Numerical values for B_{jk} appearing in the unitarity bound for BCL coefficients, defined in (7.35), for the f_0^s and f_+^s cases. The rest of the elements can be obtained from these using the properties $B_{j(j+k)} = B_{0k}$ and $B_{jk} = B_{kj}$.

7.3.4 R_{D_s}

Using our calculated form factors $f_{0,+}^s(q^2)$, we can produce a new prediction for the quantity

$$R_{D_s} = \frac{\mathcal{B}(B_s \rightarrow D_s \tau \nu_\tau)}{\mathcal{B}(B_s \rightarrow D_s \ell \nu_\ell)}, \quad (7.36)$$

where $l = e$ or μ (the ambiguity between e and μ is negligible in comparison to the current precision on R_{D_s}). As mentioned in the introduction, the analogous quantities R_D and R_{D^*} are in tension between SM prediction and experimental measurement. There is, at time of writing, no published measurement of R_{D_s} , providing an opportunity for lattice QCD to give a clear prediction of the value of R_{D_s} expected by the SM.

Armed with form factors from a lattice QCD calculation, one can immediately produce an R_{D_s} determination by taking the ratio of SM branching fractions (7.2) between the $l = \tau$ and $l = \mu, e$ cases. $|V_{cb}|$ and η_{EW} cancel in the ratio.

A lattice prediction has already been made in [152] of $R_{D_s}|_{\text{SM}} = 0.301(6)$. We here report a new prediction:

$$R_{D_s}|_{\text{SM}} = 0.2985(43)_{\text{stat}}(27)_{\text{sys}}. \quad (7.37)$$

To arrive at this prediction we averaged over the $l = e$ and $l = \mu$ cases. We give an error budget for this result in terms of errors from our lattice calculation in table 7.7. As a check we also compute R_{D_s} using form factors from resulting from the ratio method to find $R_{D_s}|_{\text{SM}} = 0.2999(58)$.

Source	% Fractional Error
Statistics	1.27
Kinematic Interpolation	0.85
$m_h \rightarrow m_b$,	0.74
$a \rightarrow 0$	0.06
Quark Mass Mistuning	0.02
Total	1.70

Table 7.7: Error budget for $R_{D_s}|_{\text{SM}}$. Values are determined via approach explained in the caption of Table 7.4.

7.4 Conclusions

We have produced a fully non-perturbative lattice QCD prediction of the scalar and vector form factors for the $B_s \rightarrow D_s \ell \nu$ decay throughout the entire q^2 range (Fig.

7.10, see Sec. 7.5 to reconstruct), and a value for R_{D_s} (Eq. (7.37)). Our results are statistics dominated. In this calculation we used correlation functions from 3 lattice spacings, including an ensemble with an approximately physical light quark mass, and learned the b -mass dependence of the form factors by obtaining data at 12 different heavy quark masses.

Our results supply an independent check on the NRQCD formalism for computing pseudoscalar→pseudoscalar form factors. Our results validate the q^2 -extrapolation from high q^2 lattice data used in the NRQCD case since our formalism produced lattice data throughout all q^2 . We have also shown that the systematic error assigned to the NRQCD results to account for perturbative matching and truncation of the $1/m$ series in the NRQCD-HISQ current is sufficient. Our results are however more precise and do not rely on the assumptions implicit in the NRQCD formalism.

Our calculation has shown that a heavy-HISQ determination of the $B \rightarrow D\ell\nu$ form factors is very plausible. Such a calculation could use an essentially identical process as given here, with the strange valence quark simply replaced with a light one. Perhaps correlation functions from additional ensembles with smaller light quark masses would be necessary to resolve the dependence of the form factors on the light mass. Also, more statistics would likely be necessary, since the presence of a light valence quark increases the noise in the lattice data, and statistics is already the dominant uncertainty in this calculation.

7.5 Reconstructing Form Factors

This section gives the necessary information to reproduce the functional form of the form factors through q^2 reproduced in this work. We here express the form factors in terms of the BCL parameterization [193]:

$$f_0^s(q^2) = \frac{1}{1 - \frac{q^2}{M_{B_c^0}^2}} \sum_{n=0}^2 a_n^0 z^n(q^2), \quad (7.38)$$

$$f_+^s(q^2) = \frac{1}{1 - \frac{q^2}{M_{B_c^*}^2}} \sum_{n=0}^2 a_n^+ \left(z^n(q^2) - \frac{n}{3} (-1)^{n-3} z^3(q^2) \right),$$

where the function $z(q^2)$ is defined by defined by

$$z(q^2) = \frac{\sqrt{t_+ - q^2} - \sqrt{t_+}}{\sqrt{t_+ - q^2} + \sqrt{t_+}}, \quad (7.39)$$

and $t_+ = (M_{B_s} + M_{D_s})^2$ (one should take the PDG 2018 values for these masses). For the position of the poles, one can use $M_{B_c^0} = 6.70390(80)\text{GeV}$ and $M_{B_c^*} = 6.28030(80)\text{GeV}$. The coefficients $a_n^{0,+}$ found from our fit, along with their covariance, is given in table 7.8.

a_0^0	a_1^0	a_2^0	a_0^+	a_1^+	a_2^+
0.66097	-0.26421	-0.26158	0.66097	-3.17196	0.10935
0.00016	0.00217	0.00125	0.00016	0.00014	0.00001
	0.06838	0.18373	0.00217	0.01578	-0.00031
		3.47982	0.00125	0.18432	-0.00606
			0.00016	0.00014	0.00001
				0.27937	0.09825
					4.06414

Table 7.8: Our results for z -coefficients in the BCL parameterization (7.38). The first row gives mean values, and the rest of the table gives the covariance matrix associated with these parameters.

7.6 Numerical Values for Lattice Results

In this section we give two tables, consisting of all numerical results for form factors, ratios $R_{0,+}^s(q^2)$, masses, energies and decay constants required for the extrapolations performed to the physical point. Table 7.9 gives results relevant to all q^2 values, while 7.10 gives results that vary over q^2 .

Set	$am_{\mu_0}^{\text{val}}$	aM_{H_s}	aM_{D_s}	aM_{H_c}	$a f_{H_c}$	aM_{η_h}	aM_{η_c}	aM_{η_s}
2	0.5	0.95976(11)	0.90222(10)	1.419515(41)	0.186299(70)	1.471675(38)	1.367014(40)	0.313886(75)
	0.65	1.12519(15)		1.573302(40)	0.197220(77)	1.775155(34)		
	0.8	1.28139(19)		1.721226(39)	0.207068(78)	2.064153(30)		
3	0.5	0.95446(13)	0.87715(11)	1.40025(26)	0.183482(46)	1.470095(25)	1.329291(27)	0.304826(52)
	0.8	1.27560(24)		1.702438(24)	0.203382(50)	2.062957(19)		
4	0.427	0.77433(17)	0.59150(11)	1.067224(46)	0.126564(70)	1.233585(41)	0.896806(48)	0.207073(96)
	0.525	0.88447(21)		1.172556(46)	0.130182(72)	1.439515(37)		
	0.65	1.01931(29)		1.303144(46)	0.133684(75)	1.693895(33)		
	0.8	1.17383(42)		1.454205(46)	0.137277(79)	1.987540(30)		
5	0.5	0.80237(22)	0.439879(93)	1.011679(29)	0.099006(56)	1.342747(27)	0.666754(39)	0.153827(77)
	0.65	0.96357(30)		1.169775(31)	0.100553(66)	1.650264(23)		
	0.8	1.11734(41)		1.321656(35)	0.101720(82)	1.945763(21)		

Table 7.9: Values extracted from correlation function fits required for the extrapolation of $f_{0,+}^s(q^2)$, $R_{0,+}^s(q^2)$ to the physical point. $a f_{H_c}$ is found from fitting the H_c correlation functions to obtain $a_0^{H_c}$, and using Eq. (4.37). D_s energies away from zero recoil are given in Table 7.10.

Set	$am_{\ell 0}^{\text{val}}$	$q^2[\text{GeV}^2]$	aE_{D_s}	$f_0^s(q^2)$	$f_+^s(q^2)$	$R_0(q^2)[\text{GeV}^{-3/2}]$	$R_+(q^2)[\text{GeV}^{-3/2}]$	
2	0.5	0.016	0.90222(10)	1.0013(15)		1.410(12)		
		0.000	0.90391(11)	1.0001(15)	1.0001(15)	1.409(12)	1.409(12)	
		0.65	0.234	0.90222(10)	1.0044(16)		1.269(11)	
		0.118	0.91315(12)	0.9968(16)	1.0031(61)	1.260(11)	1.268(13)	
		0.003	0.92407(14)	0.9893(17)	0.9895(17)	1.250(11)	1.250(11)	
		0.8	0.678	0.90222(10)	1.0096(18)		1.162(10)	
		0.342	0.93000(15)	0.9904(20)	1.0069(88)	1.1397(99)	1.159(14)	
		0.008	0.95766(25)	0.9721(24)	0.9725(24)	1.1186(99)	1.1190(99)	
		3	0.5	0.030	0.87715(11)	1.0004(15)	1.369(11)	
		0.026	0.87759(11)	1.0001(15)	1.002(19)	1.369(11)	1.371(28)	
		0.8	0.801	0.87715(11)	1.0054(18)		1.1258(91)	
		0.613	0.89178(15)	0.9948(22)	1.030(18)	1.1139(91)	1.154(22)	
4	0.427	0.371	0.59150(11)	0.9938(20)		1.250(11)		
		0.188	0.60215(14)	0.9802(23)	0.9932(52)	1.233(11)	1.249(12)	
		0.006	0.61276(17)	0.9682(24)	0.9686(25)	1.217(11)	1.218(11)	
0.525		0.953	0.59150(11)	0.9866(23)		1.151(10)		
		0.536	0.61276(17)	0.9603(28)	0.9902(76)	1.120(10)	1.155(13)	
		0.012	0.63940(30)	0.9313(33)	0.9319(33)	1.086(10)	1.087(10)	
0.65		2.032	0.59150(11)	0.9768(31)		1.0523(95)		
		0.948	0.63940(30)	0.9208(41)	0.9609(88)	0.9920(95)	1.035(13)	
		0.008	0.68092(55)	0.8786(45)	0.8788(45)	0.9465(94)	0.9468(94)	
0.8		3.765	0.59150(11)	0.9677(41)		0.9611(91)		
		1.434	0.68092(55)	0.8673(57)	0.9052(94)	0.8614(93)	0.899(12)	
		-0.055	0.7381(13)	0.8203(87)	0.8192(86)	0.815(11)	0.814(11)	

Table 7.10: Parameters extracted from correlation function fits at varying q^2 points (from the fine and fine-physical ensembles, sets 2 and 3). $f_{0,+}^s(q^2)$ are extracted via (7.18) and (7.20), and $R_{0,+}^s(q^2)$ is defined in (7.21).

Set	am_{h0}^{val}	$q^2[\text{GeV}^2]$	aE_{D_s}	$f_0^s(q^2)$	$f_+^s(q^2)$	$R_0(q^2)[\text{GeV}^{-3/2}]$	$R_+(q^2)[\text{GeV}^{-3/2}]$
5	0.5	2.635	0.439879(93)	0.9721(33)		1.030(10)	
		1.177	0.48519(24)	0.9111(36)	0.9509(70)	0.9655(96)	1.008(12)
		-0.024	0.52252(60)	0.8644(52)	0.8636(51)	0.916(10)	0.9151(99)
0.65	5.500	0.439879(93)	0.9548(39)			0.9265(93)	
	3.749	0.48519(24)	0.8957(43)	1.005(20)		0.8691(89)	0.975(21)
	2.306	0.52252(60)	0.8500(61)	0.915(12)		0.8248(95)	0.887(14)
0.8	0.248	0.5758(32)	0.753(23)	0.760(20)		0.731(23)	0.738(21)
	9.204	0.439879(93)	0.9409(45)			0.8490(88)	
	5.500	0.52252(60)	0.8413(72)	0.982(27)		0.7592(95)	0.886(26)
0.656	3.114	0.5758(32)	0.754(22)	0.839(28)		0.681(21)	0.757(27)
	0.656	0.631(10)	0.661(50)	0.673(43)		0.596(45)	0.608(39)

Table 7.11: Parameters extracted from correlation function fits at varying q^2 points, from the ultrafine ensemble.

Conclusions

In this work, we produced the first published results of applying the heavy-HISQ approach to semileptonic form factors. We found a new determination of $h_{A_1}^s(1)$, twice as accurate as the previous result and containing considerably fewer assumptions -

$$\mathcal{F}^{B_s \rightarrow D_s^*}(1) = h_{A_1}^s(1) = 0.9020(96)_{\text{stat}}(90)_{\text{sys}}. \quad (8.1)$$

Future experimental data may be combined with this result to produce a new determination of $|V_{cb}|$.

We found $B_s \rightarrow D_s l \nu$ form factors using heavy-HISQ, also improving on the precision in comparison to previous lattice results. For the first time, we were able to obtain lattice data spanning the entire q^2 range, due to the properties of the heavy-HISQ method. From these form factors, we found

$$R_{D_s}|_{\text{SM}} = 0.2985(43)_{\text{stat}}(27)_{\text{sys}}. \quad (8.2)$$

which can be combined with future experimental data to provide a new test of the Standard Model, namely a new probe into the possibility of lepton flavour violation.

Besides the successes from the heavy-HISQ approach, I have discovered some problems with using NRQCD-HISQ currents to compute semileptonic $b \rightarrow c$ transitions on the lattice. Namely, it was discovered that the expansion of NRQCD-HISQ currents do not converge very fast. $\mathcal{O}(1/m_b)$ and possibly $\mathcal{O}(1/m_b^2)$ terms are important for the dispersion relation of heavy-light mesons. So-called negligible pieces of the spacial vector current have large magnitudes:

$$V_k^{(2)} \sim V_k^{(4)} \sim 0.35 \times V_k^{(0)}. \quad (8.3)$$

I attempted some approaches to non-perturbatively renormalizing the NRQCD-HISQ currents in order to account for these issues, with limited success.

Heavy-HISQ calculations are more computationally costly than their equivalent calculations using NRQCD for the b . Taking into account the need for finer lattices and multiple data with different m_h values, heavy-HISQ costs something of the

order of 20 times more than NRQCD. However, the cost in human time of NRQCD, via perturbative matching calculations and larger and less stable correlator fits, is clearly larger than for heavy-HISQ. Also, heavy-HISQ does not contain assumptions of negligible subleading terms in the relativistic expansion or validity of perturbation theory via the matching.

The dominant uncertainty in our heavy-HISQ results are statistical. Future calculations with this approach must increase the statistics of lattice data to improve on the results presented here. This means gaining data on more gauge configurations, and with more choices of source timeslice t_0 . This will simply require more computational resources to achieve.

These two results from heavy-HISQ are in strong agreement with all other recent lattice determinations of these form factors and the analogous $B \rightarrow D^* \ell \nu$ and $B \rightarrow D \ell \nu$ form factors. Other lattice determinations use very different methodologies to the work presented here, compare for example our $h_{A_1}^s(1)$ calculation to the Fermilab/MILC calculation of $h_{A_1}(1)$. That calculation used different gauge ensembles ($N_f = 2 + 1$ MILC), a different action for the b and c quarks (Fermilab action), a different approach to analyzing correlation functions (double-ratio approach), a different normalization of currents (perturbative normalization), and different continuum and light mass extrapolation. The combination of consistent results from independent studies makes the overall contribution of lattice QCD to $b \rightarrow c$ form factors extremely robust.

Further contributions from lattice QCD are necessary in $b \rightarrow c$ transitions. The current precision on $|V_{cb}|$ is limited in roughly equal part by theoretical and experimental errors, so more precision on $b \rightarrow c$ form factors is needed to more precisely determine $|V_{cb}|$, and understand the source of the tension in its exclusive/inclusive determinations. The SM predictions of $R(D_{(s)}^{(*)})$ are currently much more precise than the experimental measurements, however more independent SM calculations of these ratios are necessary to ensure no errors are being underestimated.

List of Figures

2.1	The flavor-changing charged current vertex.	7
2.2	A sketch of the unitarity triangle.	8
2.3	Exclusion regions for the vertices of the CKM triangle from various measurements, courtesy of the most recent PDG update [3].	9
2.4	Leptonic decay of meson M at tree level in the electroweak coupling.	10
2.5	Semileptonic decay, $M \rightarrow M' l \bar{\nu}$, at tree level in electroweak coupling.	12
2.6	Different determinations of $ V_{cb} $. Points labelled $w = 1$ are determinations from extrapolating measurements of decay rates to the zero recoil point and combining them with a lattice determination of the form factor at zero recoil. Points labelled $w \geq 1$ are results from using a combination of both branching fractions and lattice form factors through some range of w . The first name mentioned in the labels give the source of the lattice form factors, and the second gives the source of the experimental data (e.g. the HPQCD+BaBar point used form factors from the HPQCD collaboration and data from the BaBar experiment). The highest point in red is from [36], the second and third highest from [37], fourth from [38], fifth from [39]. The bottom point is from the PDG [3], using data from the ALPEPH [40], Belle [41], BaBar [42, 43], and CLEO [44] experiments.	15
2.7	$R(D^{(*)})$ determinations from SM and measurement [4]. One can see from this that the SM prediction (the small blue ellipse) is inconsistent with the experimental average (the filled red ellipse).	17
2.8	The relationship between scale Q and the strong coupling constant α_s , from the PDG [3].	20
3.1	Depiction of color spaces (fibres) at two points in spactime (base space) with the value of the quark field q represented at each point as a color vector, and the connection $W(x, y)$ needed to compare the two color vectors. A gauge transform changes the two vectors in different ways, so for the comparison to be gauge independent the connection must also transform appropriately.	34

3.2	Depiction of a gauge invariant quark bilinear, connected by a Wilson line made of gauge links.	34
3.3	The elementary plaquette.	35
3.4	Terms additional to the elementary plaquette in the improved pure QCD action.	37
3.5	Taste mixing at tree level.	43
3.6	Different scales relevant to non-perturbative physics, and brackets showing the range of scales that typical lattices can resolve. The larger the range of scales resolved, the more computationally expensive the calculation.	45
3.7	An extrapolation to $m_h = m_b$ of the H_s decay constant (where H_s is a pseudoscalar $\bar{h}s$ meson) [5]. The colorful points are measurements of f_{H_s} on the lattice, the color denotes lattice spacing. The grey band shows the result of a continuum extrapolation of this lattice data. The x axis, M_{η_h} , is a proxy for the h -quark mass.	46
5.1	B_c 2-point correlation functions, with the valence charm given momentum twists $\theta = 0, 3$ and 5	71
5.2	Decay amplitude ratios (colourful points) against the expected relativistic behaviour (grey dotted line and band). Adding the $A_{0,\text{lat}}^{(1)}$ piece of the current does not improve the relativistic behaviour of the ratio.	72
5.3	$B \rightarrow D\ell\nu$ form factors. The coloured points show lattice data, each color represents an ensemble. The grey band represents the continuum and kinematically extrapolated result. The red band shows the kinematically extrapolated result at $a \simeq 0.15\text{fm}$	79
5.4	$B_s \rightarrow D_s\ell\nu$ form factors. The coloured points show lattice data, each color represents an ensemble. The grey band represents the continuum and kinematically extrapolated result. The red band shows the kinematically extrapolated result at $a \simeq 0.15\text{fm}$. Note that the $a \simeq 0.09\text{fm}$ data is not included in the extrapolation, since this prevents a good fit from being possible, see Sec. 5.3.4.	80

- 5.5 Lattice results for $f_0^s(q_{\max}^2)$ against a^2 . The grey band shows the result for $f_0^s(q_{\max}^2)$ computed in chapter 7 using the Heavy-HISQ approach for comparison. Clearly the result on the fine ensemble (the black point), and possibly on the coarse ensemble (the blue point), contain large unknown systematic errors. 81
- 5.6 Ratios of (matrix elements of) the subleading NRQCD-HISQ currents $V_\mu^{(n>0)}$ to the leading order current $V_\mu^{(0)}$, in the $B_s \rightarrow D_s$ case. . . . 82
- 5.7 $B \rightarrow D$ form factors extracted from $\langle V_0 \rangle$ and $\langle S \rangle$. The statistical errors here are not well under control, a continuum and kinematic extrapolation is not possible using these form factors. 84
- 5.8 A schematic of the chain of steps towards normalizing the scalar, temporal vector and spatial vector NRQCD-HISQ currents. Details given in the next three sections; 5.4.1, 5.4.2 and 5.4.3. 85
- 5.9 Comparison of form factors from $\langle V_0 \rangle, \langle S \rangle$ and $\langle V_0 \rangle, \langle V_k \rangle$, no additional normalizations. The bands show the results of fitting the data to the BCL parameterization (5.23), where coefficients $a_n^{0,+}$ are fit parameters. These bands are intended simply to guide the eye. . . . 88
- 5.10 Comparison of form factors from $\langle V_0 \rangle, \langle S \rangle$ and $\langle V_0 \rangle, \langle V_k \rangle$, $\langle V_0 \rangle$ is multiplied by Z_{V_0} given in (5.39). The bands show the results of fitting the data to the BCL parameterization (5.23), where coefficients $a_n^{0,+}$ are fit parameters. 89
- 5.11 Z_{V_k} from constraining form factors to be the same from $\langle V_0 \rangle, \langle S \rangle$ and $\langle V_0 \rangle, \langle V_k \rangle$ 90
- 5.12 Comparison of form factors from $\langle V_0 \rangle, \langle S \rangle$ and $\langle V_0 \rangle, \langle V_k \rangle$, with $\langle V_0 \rangle$ normalized with Z_{V_0} and $\langle V_k \rangle$ normalized with Z_{V_k} . The bands show the results of fitting the data to the BCL parameterization (5.23), where coefficients $a_n^{0,+}$ are fit parameters. 90

- 6.1 Tests of the correlator fits on the fine ensemble. The left panel shows J_{00}^{nn} at the heavy mass $am_{h0}^{\text{val}} = 0.5$. The errors shown here as statistical, estimated by taking the second derivative of the best-fit χ^2 with respect to the fit parameter in question. At $N_{\text{test}} = 1$ we give the final accepted result. $N_{\text{test}} = 2$ gives the results when all priors are broadened by 50%. $N_{\text{test}} = 3$ and 4 gives the results of setting $N_{\text{exp}} = 4$ and 6 respectively. $N_{\text{test}} = 5, 6$ gives the results of setting $t_{\text{cut}} = 2, 4$ respectively for all correlators. $N_{\text{test}} = 7$ gives the result without marginalising out the $n = 5$ excited state. $N_{\text{test}} = 8$ gives the result of moving the SVD cut from 10^{-3} to 10^{-2} 100
- 6.2 Effective energies and amplitudes for H_s and D_s^* correlators on the fine ensemble. The energies are obtained from Eq. (6.14), and amplitudes from Eq. (6.15). The grey bands give the results of the full multiexponential fit. 102
- 6.3 Sanity check for fits to the 3-point correlation functions. This ratio should approach J_{00}^{nn} for $t \gg 0$ and $t \ll T$. The grey bands show the result for J_{00}^{nn} from the full Bayesian simultaneous fit for each am_h (here they are all bunched up so appear as a single band). If the T values were larger, one can envision the data reaching a plateau at the same height as the grey band. 103
- 6.4 The variation of the 1-loop expression for η_A (Eq. (6.23)) throughout the $m_c \leq m_h \leq m_b$ range. For m_b and m_c values we used $m_b^{\overline{\text{MS}}}(m_b^{\overline{\text{MS}}})$ and $m_c^{\overline{\text{MS}}}(m_c^{\overline{\text{MS}}})$. For the coupling constant we used $\alpha_s(\sqrt{m_b m_c})$ 107
- 6.5 $h_{A_1}^s(1)$ against M_{η_h} (a proxy for the heavy quark mass). The grey band shows the result of the extrapolation at $a = 0$ and physical l, s and c masses. Sets listed in the legend follow the order of sets in Table 6.1. The red point represents a determination of the same quantity from a previous study using the NRQCD action for the b [33]. . . . 113
- 6.6 $h_{A_1}(1)$ and $h_{A_1}^s(1)$ data on the fine ensemble. Note that on this ensemble the light quark is not physical, $m_l/m_s = 0.2$ 114

- 6.7 $h_{A_1}^{(s)}(1)$ from different calculations. Our result is marked (HISQ,HPQCD). Those marked (NRQCD,HPQCD) are from [33]. The quantity marked (HPQCD) is the result of multiplying our result for $h_{A_1}^s(1)$ with the ratio $h_{A_1}(1)/h_{A_1}^s(1)$ computed at [33]. The quantity marked (Fermilab,Fermilab/MILC) is from [38]. Note that our methodology is very different to that of Fermilab/MILC in a number of ways, so the comparison between our and their results is a very robust test. 115

- 6.8 Lattice data and continuum extrapolated data for three studies of $h_{A_1}(1)$ and $h_{A_1}^s$, against the pion mass. Points labeled FNAL/MILC are from [38], and those labeled NRQCD are from [33]. The x-axis must be taken with a pinch of salt, the points at $M_\pi = M_{\eta_s}$ have pions in the sea of smaller masses than M_{η_s} , but we place them here to signify that the spectator quark has the mass of a strange quark. 116

- 6.9 Results of $h_{A_1}^s(1)$ extrapolation tests. Points **1-3** show the final result if data from the fine, superfine or ultrafine ensembles are not used in the fit. Points **4 & 5** points show the result if data at the highest/lowest am_{h0}^{val} value on each ensemble are removed. Point **6**, ' $N_{\text{nuisance}} = 3$ ' shows the result of truncating each sum in $\mathcal{N}_{\text{disc}}$ (Eq. (6.25)) at 3 rather than 2. Point **7**, '+ $1/m_b^3$ ' results from adding an extra term to (6.24) of the form $p/M_{\eta_h}^3$ where p is a fit parameter with the same prior as $l_{V,A,P}^s$. In this case, the Bayes factor falls by a factor of 7, suggesting that the data does not contain a cubic dependence on the heavy mass. Points **8 & 9** show the results of the implementations of η_A described in Sec. 6.2.4. ρ is a fit parameter with prior distribution 0 ± 1 . Including this factor causes the Bayes factor to drop by a factor of 20, implying that the data cannot resolve logarithms in m_h . Point 9 shows the result of using the 1-loop expression for η_A (Eq. (6.23)), with m_c/m_h replaced with M_{η_c}/M_{η_h} . Point **10**, ' $A + 1/m_b m_c + 1/m_b^2$ ' is the result of replacing $1 + l_V/m_c^2$ in the fit with simply a fit parameter A with prior distribution 1 ± 1 . The fact that this does not affect the fit implies that charm mistuning does not strongly affect the extrapolation. Points **11 & 12** show the result of replacing the heavy mass proxy $M_{\eta_h}/2$ with M_{H_s} and Eq. (6.21) respectively. Point **13**, 'Ratio with f_{H_c} ' is the result of an alternative extrapolation described in Sec. 6.3.4. 117
- 6.10 $h_{A_1}^s(1)/(f_{H_c}\sqrt{M_{H_c}})$ against M_{η_h} (a proxy for the heavy quark mass). The grey band shows the result of the extrapolation at $a = 0$ and physical l,s and c masses. Sets listed in the legend follow the order of sets in Table 6.1. The black point shows our final result for $h_{A_1}^s(1)$ divided by $\sqrt{M_{B_c}}$ from the PDG [3] and f_{B_c} from our extrapolation of f_{H_c} to continuum and physical b mass. 119
- 6.11 f_{H_c} against M_{η_h} (a proxy for the heavy quark mass). The grey band shows the result of the extrapolation at $a = 0$ and physical l,s and c masses. Sets listed in the legend follow the order of sets in table 6.1. The red point shows the result from a previous heavy-HISQ determination of f_{B_c} on 2+1 gauge ensembles [5]. 120
- 6.12 Lattice results for $M_{D_s^*} - M_{\eta_c}/2$ on each ensemble. The grey band shows the PDG result [3]. 121

- 6.13 Extrapolation of $M_{H_s} - M_{\eta_h}/2$ to the physical point. The grey band shows the result at $a = 0$ and physical charm, strange and light masses. 122
- 7.1 Tests on the correlator fits on the fine ensemble. $N_{\text{test}} = 1$ gives the final accepted result. $N_{\text{test}} = 2$ and 3 gives the results of setting $N_{\text{exp}} = 4$ and 6 respectively. $N_{\text{test}} = 4$ gives the results when all priors are given a standard deviation one and a half times the size of the final fit. $N_{\text{test}} = 5$ gives the result of setting $t_{\text{cut}} = 4$ rather than 2 for all correlators. $N_{\text{test}} = 6$ gives the result without marginalising out the $n = 5$ excited state. $N_{\text{test}} = 7$ gives the result of moving the svd cut from 10^{-3} to 10^{-2} . $N_{\text{test}} = 8, 9, 10$ gives the result of using a chained fit described above with priors for the full correlated fit given additional fractional errors of 100%, 150% and 200% respectively. . . 132
- 7.2 Superaveraged effective energies (given by Eq. (6.14)) and amplitudes (given by Eq. (6.15)) for a selection of 2-point correlators on the fine ensemble. The grey bands show the corresponding results for these quantities from the simultaneous Bayesian fits. These plots supply a further check of our correlator fits - results of the Bayesian fits are in good agreement with the effective energies and amplitudes. 134
- 7.3 $\tilde{C}_3(t, T)/\tilde{C}^{H_s}(t)\tilde{C}^{D_s}(T - t)$, which should plateau at J_{00}^{nn} , in the fine ensemble for the $J = S$ and $J = V_0$ cases. The grey bands show the corresponding results for these quantities from the simultaneous Bayesian fits. Unfortunately, in the V_0 case the oscillating component dominates the correlation function, preventing any plateau from being visible. 135
- 7.4 $R_0^s(q_{\text{max}}^2) = f_0^s(q_{\text{max}}^2)/(f_{H_c}\sqrt{M_{H_c}})$ against M_{η_h} (a proxy for the heavy quark mass). The grey band shows the result of the extrapolation at $a = 0$ and physical l, s and c masses. Sets listed in the legend follow the order of sets in table 4.1. 140

- 7.5 Results of tests of the $R_0^s(q_{\max}^2)$ extrapolation. The top three blue points show $R_0^s(q_{\max}^2)$ at continuum and physical quark mass, if data from the fine, superfine or ultrafine ensembles are not used in the fit. The fourth and fifth blue points show the result if data at the highest/lowest am_{h0}^{val} value on each ensemble are removed. The point labelled $N_{fit} = 3$ is the result of extending the sum in (7.24) such that it truncates at 3 rather than 2 in each of the i, j, k directions. The point labelled $+\log(M_{\eta_h}/M_{\eta_c})^2$ represents the result of adding a $\rho_{2,n}\log(M_{\eta_h}/M_{\eta_c})^2$ term in the first set of brackets in (7.24), where $\rho_{2,n}$ are new fit parameters with the same prior distributions as ρ_n . Similarly $+\log(M_{\eta_h}/M_{\eta_c})/M_{\eta_h}$ shows the result of adding this term multiplied by $\rho_{2,n}$. The lowest point shows the result of our direct extrapolation of $f_0^s(q_{\max}^2)$ to the physical point, divided by the PDG value for $\sqrt{M_{B_c}}$ [3] and the result of our extrapolation of f_{B_c} to the physical point detailed in Sec. 6.3.4. 142
- 7.6 $R_{0,+}^s(q^2) = f_{0,+}^s(q^2)/(f_{H_c}\sqrt{M_{H_c}})$ against q^2 . The grey band shows the result of the extrapolation at $a = 0$ and physical quark masses. Sets listed in the legend follow the order of sets in table 4.1. 143
- 7.7 $f_0^s(q_{\max}^2)$ against M_{η_h} (a proxy for the heavy quark mass). The grey band shows the result of the extrapolation at $a = 0$ and physical quark masses. We also include the result from a previous lattice calculation, which used the NRQCD discretisation for the b quark [152]. Sets listed in the legend follow the order of sets in Table 4.1. 144

- 7.8 Results of tests of the $f_0^s(q_{\max}^2)$ extrapolation. The top three blue points show $f_0^s(q_{\max}^2)$ at continuum and physical quark mass, if data from the fine, superfine or ultrafine ensembles are not used in the fit. The fourth and fifth blue points show the result if data at the highest/lowest am_{h0}^{val} value on each ensemble are removed. The point labelled $N_{fit} = 3$ is the result of extending the sum in (7.24) such that it truncates at 3 rather than 2 in each of the i, j, k directions. The point labelled $+\log(M_{\eta_h}/M_{\eta_c})^2$ represents the result of adding a $\rho_2 \log(M_{\eta_h}/M_{\eta_c})^2$ term in the first set of brackets in (7.24), where $\rho_{2,n}$ are new fit parameters with the same prior distributions as ρ_n . Similarly, the point labelled $+\log(M_{\eta_c}/M_{\eta_h})/M_{\eta_h}$ gives the result of adding this term multiplied by $\rho_{2,n}$. The lowest point shows the result from the extrapolation of $R_0^s(q_{\max}^2)$, multiplied by the PDG value for $\sqrt{M_{B_c}}$ [3] and the result of our extrapolation of f_{B_c} to the physical point detailed Sec. 6.3.4. 146
- 7.9 $f_{0,+}^s(q^2)$ against q^2 . The grey band shows the result of the extrapolation at $a = 0$ and physical quark masses. Sets listed in the legend follow the order of sets in Table 4.1. 147
- 7.10 Final result for $f_{0,+}^s(q^2)$ against q^2 at the physical point 147
- 7.11 Error budget for $f_{0,+}^s(q^2)$ against q^2 . Values are determined via approach explained in the caption of Table 7.4. 148
- 7.12 Results for $f_{0,+}^s(q^2)$ against q^2 at the physical point, from both the ratio method and the direct method. 148
- 7.13 Our final result for $f_{0,+}^s(q^2)$ against form factors calculated from a previous study using the NRQCD action for the b quark [152]. The darker shaded region of the NRQCD band shows where lattice data was available in that study, the rest of the band shows the result of an extrapolation in q^2 using the BCL parameterization. 149

List of Tables

4.1	Parameters for the MILC gluon ensembles [7, 8]. a is the lattice spacing, determined from the Wilson flow parameter w_0 . Values for w_0/a are from: sets 0,1,2 [138], sets 3 and 4 [139], set 5 [140]. The physical value of w_0 was determined to be $w_0 = 0.1715(9)\text{fm}$ in [141]. Columns 5-7 give the masses used in the action for light, strange and charm quarks in the sea.	55
5.1	Parameters used in our calculation. am_{s0}^{val} and am_{c0}^{val} are the bare masses of the strange and charm valence quarks, tuned in [153]. am_{b0}^{val} is the bare mass of the valence bottom quark, tuned in [119]. $u_0 = (\text{Re Tr}(\square)/3)^{1/4}$ is the ‘tadpole improvement parameter’ as used in [119] via a perturbative calculation of the plaquette \square . $\{c_i\}$ are the coefficients for the kinetic and chromomagnetic terms in the NRQCD action (Eq. (3.60)) [154]. $\{T\}$ is the set of temporal separations between source ($B_{(s)}$ creation operator) and sink ($D_{(s)}$ annihilation operator). a_{sm} are the radii of the exponential smearing function applied to the $B_{(s)}$ and $D_{(s)}$ creation operators.	74
5.2	Momentum twists (and corresponding momenta and q^2 values) given to the charm propagator on each ensemble.	76
6.1	Parameters relevant to our calculation. Columns 3 and 4 give the s and c valence quark masses, these values were tuned in [139] to reproduce the correct η_s and η_c masses. We used a number of heavy quark masses to assist the extrapolation to the physical b mass, given in column 5. Column 6 gives the number of gauge configurations (n_{cfg}) and the number of t_0 choices (n_{src}) used. Column 7 gives the temporal separations between B_s source and D_s^* sink, T/a , of the 3-point correlation functions computed on each ensemble.	96
6.2	t_{cut} values used for each correlator on each ensemble. A_0 denotes the 3-point correlators. The rest are for 2-point correlators.	101
6.3	Normalization constants applied to the lattice axial vector current in Eq. (6.18). Z_A is found from Eq. (6.16) and Z_{disc} from Eq. (6.17). . .	104

- 6.4 Values extracted from correlation function fits for $h_{A_1}^s(1)$, along with masses of the two mesons on either end of the transition. $h_{A_1}^s(1)$ values are found from Eq. (6.18). Errors are statistical. 111
- 6.5 Values extracted from correlation function fits. f_{H_c} is the H_c decay constant derived from Eq. (4.37). 112
- 6.6 Error budget for $h_{A_1}^s(1)$. The value for statistics & Z_A is given by the partial standard deviation of $h_{A_1}^s(1)$ with respect to the lattice data. The value for $a \rightarrow 0$ is the partial standard deviation of $h_{A_1}^s(1)$ with respect to priors of the fit parameters in $\mathcal{N}_{\text{disc}}$. Similarly for $m_h \rightarrow m_b, c$ -mistuning the value is the partial standard deviation with respect to priors of $l_{V,A,P}$, and the mistuning value comes from priors of parameters in $\mathcal{N}_{\text{mistuning}}$ 112
- 7.1 Simulation details. Columns 2 and 3 give the s and c valence quark masses, which were tuned in [139]. Column 4 gives the bare heavy quark masses, we use a number of heavy quark masses to assist the extrapolation to the physical b mass. Column 5 gives the absolute value of the spatial momentum given to the D_s meson, using a momentum twist, in lattice units. These values were chosen with the following rationale: when only 2 are used, these correspond to the $q^2 = 0$ and q_{max}^2 points (except on the fine-physical ensemble, where we compute at the points q_{max}^2 and $q_{\text{max}}^2/2$). When 3 twists are used, the momenta correspond to $q^2 = 0$, $q^2 = q_{\text{max}}^2/2$, and q_{max}^2 points. When 4 are used, these are points for q_{max}^2 , $3q_{\text{max}}^2/4$, $q_{\text{max}}^2/2$, $q_{\text{max}}^2/4$, $q^2 = 0$. To give the D_s meson these spatial momenta we gave the charm an appropriate momentum twist in the $(1, 1, 1)$ direction. Column 6 gives the temporal separations between source and sink, T , of the 3-point correlation functions computed on each ensemble. 127
- 7.2 t_{cut} values used for each correlator on each ensemble. S and V_0 denote the 3-point correlators with the corresponding current S or V_0 . The rest are for 2-point correlators. There is one exception to the values here: on the 3-point correlators on the ultrafine ensemble with $am_{h_0}^{\text{val}} = 0.8$ and $q^2 = 0$, the t_{cut} given here is replaced with 8, 10 and 12 in the $T/a = 31, 36, 41$ cases respectively. This is due to the signal/noise degradation from the large D_s momentum causing noise that makes data close to the H_s source useless to the fit. 130

7.3	Normalization constants applied to the lattice axial vector current in (7.20). Z_V is found from (7.19) and Z_{disc} from (6.17).	136
7.4	Error budget for $R_0^s(q_{\text{max}}^2)$. The value for scale setting is given by the partial standard deviation of $R_0^s(q_{\text{max}}^2)$ with respect to w_0 and w_0/a . The value for statistics is given by the partial standard deviation of $R_0^s(q_{\text{max}}^2)$ with respect to the lattice data (with scale setting error subtracted in quadrature). The value for $a \rightarrow 0$ is the partial standard deviation of $R_0^s(q_{\text{max}}^2)$ with respect to priors of the fit parameters d_{0jk}^0 for $i+j \neq 0$. Similarly for $m_h \rightarrow m_b$, the value is the partial standard deviation with respect to priors of d_{ijk}^0 for $k \neq 0$. The mistuning value comes from priors of parameters in $\mathcal{N}_{\text{mistuning}}$	141
7.5	Error budget for $f_0^s(q_{\text{max}}^2)$ found via the direct method. These values are determined in the same way as for $R_0^s(q_{\text{max}}^2)$, described in the caption of Table 7.4. No scale setting error is listed since this is less than 0.01% on account of $f_0^s(q_{\text{max}}^2)$ being dimensionless.	144
7.6	Numerical values for B_{jk} appearing in the unitarity bound for BCL coefficients, defined in (7.35), for the f_0^s and f_+^s cases. The rest of the elements can be obtained from these using the properties $B_{j(j+k)} = B_{0k}$ and $B_{jk} = B_{kj}$	151
7.7	Error budget for $R_{D_s} _{\text{SM}}$. Values are determined via approach explained in the caption of Table 7.4.	152
7.8	Our results for z -coefficients in the BCL parameterization (7.38). The first row gives mean values, and the rest of the table gives the covariance matrix associated with these parameters.	154
7.9	Values extracted from correlation function fits required for the extrapolation of $f_{0,+}^s(q^2)$, $R_{0,+}^s(q^2)$ to the physical point. af_{H_c} is found from fitting the H_c correlation functions to obtain $a_0^{H_c}$, and using Eq. (4.37). D_s energies away from zero recoil are given in Table 7.10.	155
7.10	Parameters extracted from correlation function fits at varying q^2 points (from the fine and fine-physical ensembles, sets 2 and 3). $f_{0,+}^s(q^2)$ are extracted via (7.18) and (7.20), and $R_{0,+}^s(q^2)$ is defined in (7.21).	156
7.11	Parameters extracted from correlation function fits at varying q^2 points, from the ultrafine ensemble.	157

Bibliography

- [1] E. McLean, C. T. H. Davies, A. T. Lytle, and J. Koponen. Lattice QCD form factor for $B_s \rightarrow D_s^* \ell \nu$ at zero recoil with non-perturbative current renormalisation. 2019.
- [2] E. McLean, C. T. H. Davies, J. Koponen, and A. T. Lytle. $B_s \rightarrow D_s \ell \nu$ Form Factors for the full q^2 range from Lattice QCD with non-perturbatively normalized currents. 2019.
- [3] M. et. al. Tanabashi. Review of particle physics. *Phys. Rev. D*, 98:030001, Aug 2018.
- [4] Y. Amhis et al. Averages of b -hadron, c -hadron, and τ -lepton properties as of summer 2016. *Eur. Phys. J.*, C77:895, 2017. updated results and plots available at <https://hflav.web.cern.ch><https://hflav.web.cern.ch>.
- [5] C. McNeile, C. T. H. Davies, E. Follana, K. Hornbostel, and G. P. Lepage. Heavy meson masses and decay constants from relativistic heavy quarks in full lattice QCD. *Phys. Rev.*, D86:074503, 2012.
- [6] Brian Colquhoun, Christine Davies, Jonna Koponen, Andrew Lytle, and Craig McNeile. B_c decays from highly improved staggered quarks and NRQCD. *PoS, LATTICE2016*:281, 2016.
- [7] A. Bazavov et al. Lattice QCD ensembles with four flavors of highly improved staggered quarks. *Phys. Rev.*, D87(5):054505, 2013.
- [8] A. Bazavov et al. Scaling studies of QCD with the dynamical HISQ action. *Phys. Rev.*, D82:074501, 2010.
- [9] Sheldon L. Glashow. Partial-symmetries of weak interactions. *Nuclear Physics*, 22(4):579 – 588, 1961.
- [10] Steven Weinberg. A model of leptons. *Phys. Rev. Lett.*, 19:1264–1266, Nov 1967.
- [11] Abdus Salam. Weak and Electromagnetic Interactions. *Conf. Proc.*, C680519:367–377, 1968.

-
- [12] Matthew D. Schwartz. *Quantum Field Theory and the Standard Model*. Cambridge University Press, 2014.
- [13] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.*, 13:321–323, Aug 1964.
- [14] Peter W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.*, 13:508–509, Oct 1964.
- [15] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. Global conservation laws and massless particles. *Phys. Rev. Lett.*, 13:585–587, Nov 1964.
- [16] M.C. Gonzalez-Garcia and Michele Maltoni. Phenomenology with massive neutrinos. *Physics Reports*, 460(1):1 – 129, 2008.
- [17] Katherine Freese. Review of Observational Evidence for Dark Matter in the Universe and in upcoming searches for Dark Stars. *EAS Publ. Ser.*, 36:113–126, 2009.
- [18] Adam G. Riess et al. Observational evidence from supernovae for an accelerating universe and a cosmological constant. *Astron. J.*, 116:1009–1038, 1998.
- [19] S. Perlmutter et al. Measurements of Omega and Lambda from 42 high redshift supernovae. *Astrophys. J.*, 517:565–586, 1999.
- [20] Laurent Canetti, Marco Drewes, and Mikhail Shaposhnikov. Matter and Antimatter in the Universe. *New J. Phys.*, 14:095012, 2012.
- [21] Kenneth G. Wilson. The Renormalization Group and Strong Interactions. *Phys. Rev.*, D3:1818, 1971.
- [22] Eldad Gildener. Gauge-symmetry hierarchies. *Phys. Rev. D*, 14:1667–1672, Sep 1976.
- [23] Eldad Gildener and Steven Weinberg. Symmetry breaking and scalar bosons. *Phys. Rev. D*, 13:3333–3341, Jun 1976.
- [24] Nicola Cabibbo. Unitary symmetry and leptonic decays. *Phys. Rev. Lett.*, 10:531–533, Jun 1963.
- [25] Makoto Kobayashi and Toshihide Maskawa. CP-Violation in the Renormalizable Theory of Weak Interaction. *Progress of Theoretical Physics*, 49(2):652–657, 02 1973.

- [26] Michael Gronau and Daniel Wyler. On determining a weak phase from charged b decay asymmetries. *Physics Letters B*, 265(1):172 – 176, 1991.
- [27] Michael Gronau and David London. How to determine all the angles of the unitarity triangle from $b_d^0 \rightarrow dk_s$ and $b_s^0 \rightarrow d\phi$. *Physics Letters B*, 253(3):483 – 488, 1991.
- [28] Robert Fleischer. Flavour Physics and CP Violation. In *High-energy physics. Proceedings, European School, Kitzbuehel, Austria, August 21-September, 2005*, 2006.
- [29] Large m_w , m_z behaviour of the $\mathcal{O}(a)$ corrections to semileptonic processes mediated by w . *Nuclear Physics B*, 196(1):83 – 92, 1982.
- [30] *Phys. Rev.*, 171(1):83 – 92. [Erratum: *Phys. Rev.* 174, 2169 (1968)].
- [31] David Atwood and William J. Marciano. Radiative corrections and semileptonic b decays. *Phys. Rev. D*, 41:1736–1740, Mar 1990.
- [32] V. V. Kiselev, A. K. Likhoded, and A. I. Onishchenko. Semileptonic B_c meson decays in sum rules of QCD and NRQCD. *Nucl. Phys.*, B569:473–504, 2000.
- [33] Judd Harrison, Christine Davies, and Matthew Wingate. Lattice QCD calculation of the $B_{(s)} \rightarrow D_{(s)}^* \ell \nu$ form factors at zero recoil and implications for $|V_{cb}|$. *Phys. Rev.*, D97(5):054502, 2018.
- [34] Ikaros I. Y. Bigi, Mikhail A. Shifman, N. Uraltsev, and Arkady I. Vainshtein. High power n of m_b in beauty widths and $n = 5 \rightarrow \infty$ limit. *Phys. Rev.*, D56:4017–4030, 1997. [,205(1996)].
- [35] Andre H. Hoang, Zoltan Ligeti, and Aneesh V. Manohar. B decays in the epsilon expansion. *Phys. Rev.*, D59:074017, 1999.
- [36] Heechang Na, Chris M. Bouchard, G. Peter Lepage, Chris Monahan, and Junko Shigemitsu. $B \rightarrow D \ell \nu$ form factors at nonzero recoil and extraction of $|V_{cb}|$. *Phys. Rev.*, D92(5):054510, 2015. [Erratum: *Phys. Rev.* D93,no.11,119906(2016)].
- [37] Jon A. Bailey et al. $B \rightarrow D \ell \nu$ form factors at nonzero recoil and $|V_{cb}|$ from 2+1-flavor lattice QCD. *Phys. Rev.*, D92(3):034506, 2015.

- [38] Jon A. Bailey et al. Update of $|V_{cb}|$ from the $\bar{B} \rightarrow D^* \ell \bar{\nu}$ form factor at zero recoil with three-flavor lattice QCD. *Phys. Rev.*, D89(11):114504, 2014.
- [39] Benjamin Grinstein and Andrew Kobach. Model-Independent Extraction of $|V_{cb}|$ from $\bar{B} \rightarrow D^* \ell \bar{\nu}$. *Phys. Lett.*, B771:359–364, 2017.
- [40] D Buskulic et. al. A measurement of $|V_{cb}|$ from $B^0 \rightarrow D^* l^- \nu_l$. *Physics Letters B*, 359(1):236 – 248, 1995.
- [41] Kazuo Abe et al. Measurement of $\mathcal{B}(\bar{B}^0 \rightarrow D^+ l^- \bar{\nu})$ and determination of $|V_{cb}|$. *Phys. Lett.*, B526:258–268, 2002.
- [42] Bernard Aubert et al. Measurements of the Semileptonic Decays $\bar{B} \rightarrow D \ell \bar{\nu}$ and $\bar{B} \rightarrow D^* \ell \bar{\nu}$ Using a Global Fit to $DX_l \bar{\nu}$ Final States. *Phys. Rev.*, D79:012002, 2009.
- [43] Bernard Aubert et al. Measurement of $|V_{cb}|$ and the Form-Factor Slope in $B \rightarrow D \ell^- \bar{\nu}$ Decays in Events Tagged by a Fully Reconstructed B Meson. *Phys. Rev. Lett.*, 104:011802, 2010.
- [44] John E. Bartelt et al. Measurement of the $B \rightarrow D$ lepton neutrino branching fractions and form-factor. *Phys. Rev. Lett.*, 82:3746, 1999.
- [45] Florian U. Bernlochner, Zoltan Ligeti, Michele Papucci, and Dean J. Robinson. Combined analysis of semileptonic B decays to D and D^* : $R(D^{(*)})$, $|V_{cb}|$, and new physics. *Phys. Rev.*, D95(11):115008, 2017. [erratum: *Phys. Rev.* D97,no.5,059902(2018)].
- [46] Dante Bigi, Paolo Gambino, and Stefan Schacht. A fresh look at the determination of $|V_{cb}|$ from $B \rightarrow D^* \ell \nu$. *Phys. Lett.*, B769:441–445, 2017.
- [47] Irinel Caprini, Laurent Lellouch, and Matthias Neubert. Dispersive bounds on the shape of $\bar{B} \rightarrow D^* \ell \bar{\nu}$ form-factors. *Nucl. Phys.*, B530:153–181, 1998.
- [48] C. Glenn Boyd, Benjamin Grinstein, and Richard F. Lebed. Model-independent determinations of $B \rightarrow D \ell \nu, D^* \ell \nu$ form factors. *Nuclear Physics B*, 461(3):493 – 511, 1996.
- [49] Jon A. Bailey, Sunkyoo Lee, Weonjong Lee, Jaehoon Leem, and Sungwoo Park. Updated evaluation of ϵ_K in the standard model with lattice QCD inputs. *Phys. Rev.*, D98(9):094505, 2018.

- [50] J. P. Lees et al. Evidence for an excess of $\bar{B} \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau$ decays. *Phys. Rev. Lett.*, 109:101802, 2012.
- [51] J. P. Lees et al. Measurement of an Excess of $\bar{B} \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau$ Decays and Implications for Charged Higgs Bosons. *Phys. Rev.*, D88(7):072012, 2013.
- [52] M. Huschle et al. Measurement of the branching ratio of $\bar{B} \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau$ relative to $\bar{B} \rightarrow D^{(*)}\ell^-\bar{\nu}_\ell$ decays with hadronic tagging at Belle. *Phys. Rev.*, D92(7):072014, 2015.
- [53] Y. Sato et al. Measurement of the branching ratio of $\bar{B}^0 \rightarrow D^{*+}\tau^-\bar{\nu}_\tau$ relative to $\bar{B}^0 \rightarrow D^{*+}\ell^-\bar{\nu}_\ell$ decays with a semileptonic tagging method. *Phys. Rev.*, D94(7):072007, 2016.
- [54] S. Hirose et al. Measurement of the τ lepton polarization and $R(D^*)$ in the decay $\bar{B} \rightarrow D^*\tau^-\bar{\nu}_\tau$. *Phys. Rev. Lett.*, 118(21):211801, 2017.
- [55] S. Hirose et al. Measurement of the τ lepton polarization and $R(D^*)$ in the decay $\bar{B} \rightarrow D^*\tau^-\bar{\nu}_\tau$ with one-prong hadronic τ decays at Belle. *Phys. Rev.*, D97(1):012004, 2018.
- [56] Roel Aaij et al. Measurement of the ratio of branching fractions $\mathcal{B}(\bar{B}^0 \rightarrow D^{*+}\tau^-\bar{\nu}_\tau)/\mathcal{B}(\bar{B}^0 \rightarrow D^{*+}\mu^-\bar{\nu}_\mu)$. *Phys. Rev. Lett.*, 115(11):111803, 2015. [Erratum: *Phys. Rev. Lett.*115,no.15,159901(2015)].
- [57] R. Aaij et al. Measurement of the ratio of the $B^0 \rightarrow D^{*-}\tau^+\nu_\tau$ and $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$ branching fractions using three-prong τ -lepton decays. *Phys. Rev. Lett.*, 120(17):171802, 2018.
- [58] R. Aaij et al. Test of Lepton Flavor Universality by the measurement of the $B^0 \rightarrow D^{*-}\tau^+\nu_\tau$ branching fraction using three-prong τ decays. *Phys. Rev.*, D97(7):072013, 2018.
- [59] Dante Bigi and Paolo Gambino. Revisiting $B \rightarrow D\ell\nu$. *Phys. Rev.*, D94(9):094008, 2016.
- [60] Sneha Jaiswal, Soumitra Nandi, and Sunando Kumar Patra. Extraction of $|V_{cb}|$ from $B \rightarrow D^{(*)}\ell\nu_\ell$ and the Standard Model predictions of $R(D^{(*)})$. *JHEP*, 12:060, 2017.
- [61] Dante Bigi, Paolo Gambino, and Stefan Schacht. $R(D^*)$, $|V_{cb}|$, and the Heavy Quark Symmetry relations between form factors. *JHEP*, 11:061, 2017.

- [62] Wolfgang Altmannshofer, Peter Stangl, and David M. Straub. Interpreting Hints for Lepton Flavor Universality Violation. *Phys. Rev.*, D96(5):055008, 2017.
- [63] Christoph Bobeth, Gudrun Hiller, and Giorgi Piranishvili. Angular distributions of $\bar{B} \rightarrow \bar{K}\ell^+\ell^-$ decays. *JHEP*, 12:040, 2007.
- [64] Chris Bouchard, G. Peter Lepage, Christopher Monahan, Heechang Na, and Junko Shigemitsu. Standard Model Predictions for $B \rightarrow K\ell^+\ell^-$ with Form Factors from Lattice QCD. *Phys. Rev. Lett.*, 111(16):162002, 2013. [Erratum: *Phys. Rev. Lett.*112,no.14,149902(2014)].
- [65] Roel Aaij et al. Test of lepton universality using $B^+ \rightarrow K^+\ell^+\ell^-$ decays. *Phys. Rev. Lett.*, 113:151601, 2014.
- [66] Marzia Bordone, Gino Isidori, and Andrea Pattori. On the Standard Model predictions for R_K and R_{K^*} . *Eur. Phys. J.*, C76(8):440, 2016.
- [67] Sébastien Descotes-Genon, Lars Hofer, Joaquim Matias, and Javier Virto. Global analysis of $b \rightarrow s\ell\ell$ anomalies. *JHEP*, 06:092, 2016.
- [68] Bernat Capdevila, Sebastien Descotes-Genon, Joaquim Matias, and Javier Virto. Assessing lepton-flavour non-universality from $B \rightarrow K^*\ell\ell$ angular analyses. *JHEP*, 10:075, 2016.
- [69] Bernat Capdevila, Sebastien Descotes-Genon, Lars Hofer, and Joaquim Matias. Hadronic uncertainties in $B \rightarrow K^*\mu^+\mu^-$: a state-of-the-art analysis. *JHEP*, 04:016, 2017.
- [70] Nicola Serra, Rafael Silva Coutinho, and Danny van Dyk. Measuring the breaking of lepton flavor universality in $B \rightarrow K^*\ell^+\ell^-$. *Phys. Rev.*, D95(3):035029, 2017.
- [71] Aoife Bharucha, David M. Straub, and Roman Zwicky. $B \rightarrow V\ell^+\ell^-$ in the Standard Model from light-cone sum rules. *JHEP*, 08:098, 2016.
- [72] Wolfgang Altmannshofer, Christoph Niehoff, Peter Stangl, and David M. Straub. Status of the $B \rightarrow K^*\mu^+\mu^-$ anomaly after Moriond 2017. *Eur. Phys. J.*, C77(6):377, 2017.

- [73] Sebastian Jäger and Jorge Martin Camalich. Reassessing the discovery potential of the $B \rightarrow K^* \ell^+ \ell^-$ decays in the large-recoil region: SM challenges and BSM opportunities. *Phys. Rev.*, D93(1):014028, 2016.
- [74] R. Aaij et al. Test of lepton universality with $B^0 \rightarrow K^{*0} \ell^+ \ell^-$ decays. *JHEP*, 08:055, 2017.
- [75] Kazuo Abe et al. An Upper bound on the decay $\tau \rightarrow \mu \gamma$ from Belle. *Phys. Rev. Lett.*, 92:171802, 2004.
- [76] A. M. Baldini et al. Search for the lepton flavour violating decay $\mu^+ \rightarrow e^+ \gamma$ with the full dataset of the MEG experiment. *Eur. Phys. J.*, C76(8):434, 2016.
- [77] Ilaria Brivio and Michael Trott. The Standard Model as an Effective Field Theory. *Phys. Rept.*, 793:1–98, 2019.
- [78] Marat Freytsis, Zoltan Ligeti, and Joshua T. Ruderman. Flavor models for $\bar{B} \rightarrow D^{(*)} \tau \bar{\nu}$. *Phys. Rev.*, D92(5):054018, 2015.
- [79] Martin Bauer and Matthias Neubert. Minimal Leptoquark Explanation for the $R_{D^{(*)}}$, R_K , and $(g-2)_g$ Anomalies. *Phys. Rev. Lett.*, 116(14):141802, 2016.
- [80] Andreas Crivellin, Giancarlo D’Ambrosio, and Julian Heeck. Explaining $h \rightarrow \mu^\pm \tau^\mp$, $B \rightarrow K^* \mu^+ \mu^-$ and $B \rightarrow K \mu^+ \mu^- / B \rightarrow K e^+ e^-$ in a two-Higgs-doublet model with gauged $L_\mu - L_\tau$. *Phys. Rev. Lett.*, 114:151801, 2015.
- [81] I.S. Altarev, Yu.V. Borisov, N.V. Borovikova, S.N. Ivanov, E.A. Kolomensky, M.S. Lasakov, V.M. Lobashev, V.A. Nazarenko, A.N. Pirozhkov, A.P. Serebrov, Yu.V. Sobolev, E.V. Shulgina, and A.I. Yegorov. New measurement of the electric dipole moment of the neutron. *Physics Letters B*, 276(1):242 – 246, 1992.
- [82] David J. Gross and Frank Wilczek. Ultraviolet behavior of non-abelian gauge theories. *Phys. Rev. Lett.*, 30:1343–1346, Jun 1973.
- [83] Stefan Scherer. Introduction to chiral perturbation theory. *Adv. Nucl. Phys.*, 27:277, 2003. [,277(2002)].
- [84] S. Fubini and G. Furlan. Renormalization effects for partially conserved currents. *Physics Physique Fizika*, 1(4):229–247, 1965.

- [85] Howard Georgi. An Effective Field Theory for Heavy Quarks at Low-energies. *Phys. Lett.*, B240:447–450, 1990.
- [86] Matthias Neubert. Effective field theory and heavy quark physics. In *Physics in $D \geq 4$. Proceedings, Theoretical Advanced Study Institute in elementary particle physics, TASI 2004, Boulder, USA, June 6-July 2, 2004*, pages 149–194, 2005.
- [87] Richard F. Lebed and Mahiko Suzuki. Current algebra and the Ademollo-Gatto theorem in spin flavor symmetry of heavy quarks. *Phys. Rev.*, D44:829–836, 1991.
- [88] Michael E. Luke. Effects of subleading operators in the heavy quark effective theory. *Physics Letters B*, 252(3):447 – 455, 1990.
- [89] M. Ademollo and R. Gatto. Nonrenormalization theorem for the strangeness-violating vector currents. *Phys. Rev. Lett.*, 13:264–266, Aug 1964.
- [90] Adam F. Falk and Matthias Neubert. Second order power corrections in the heavy quark effective theory. 1. Formalism and meson form-factors. *Phys. Rev.*, D47:2965–2981, 1993.
- [91] W.E. Caswell and G.P. Lepage. Effective lagrangians for bound state problems in qed, qcd, and other field theories. *Physics Letters B*, 167(4):437 – 442, 1986.
- [92] Geoffrey T. Bodwin, Eric Braaten, and G. Peter Lepage. Rigorous QCD analysis of inclusive annihilation and production of heavy quarkonium. *Phys. Rev.*, D51:1125–1171, 1995. [Erratum: *Phys. Rev.* D55,5853(1997)].
- [93] Nora Brambilla. A Short introduction to nonrelativistic effective field theories. In *Proceedings, 23rd International Workshop on the Fundamental Problems of High Energy Physics and Field Theory: Protvino, Russia, June 21-23, 2000*, 2000.
- [94] Leslie L. Foldy and Siegfried A. Wouthuysen. On the dirac theory of spin 1/2 particles and its non-relativistic limit. *Phys. Rev.*, 78:29–36, Apr 1950.
- [95] G. Peter Lepage, Lorenzo Magnea, Charles Nakhleh, Ulrika Magnea, and Kent Hornbostel. Improved nonrelativistic QCD for heavy quark physics. *Phys. Rev.*, D46:4052–4067, 1992.

-
- [96] Michael E. Peskin and Daniel V. Schroeder. *An Introduction to quantum field theory*. 1995.
- [97] Thomas DeGrand and Carleton E. Detar. *Lattice methods for quantum chromodynamics*. 2006.
- [98] Philippe de Forcrand. Simulating QCD at finite density. *PoS*, LAT2009:010, 2009.
- [99] G. C. Wick. Properties of bethe-salpeter wave functions. *Phys. Rev.*, 96:1124–1134, Nov 1954.
- [100] K. Symanzik. Some Topics in Quantum Field Theory. In *Mathematical Problems in Theoretical Physics. Proceedings, 6th interational conference on mathematical physics, West Berlin, Germany, August 11-20, 1981*, pages 47–58, 1981.
- [101] P. Weisz. Continuum limit improved lattice action for pure yang-mills theory (i). *Nuclear Physics B*, 212(1):1 – 17, 1983.
- [102] M. Lüscher and P. Weisz. On-shell improved lattice gauge theories. *Comm. Math. Phys.*, 97(1-2):59–77, 1985.
- [103] P. Weisz and R. Wohlert. Continuum limit improved lattice action for pure yang-mills theory (ii). *Nuclear Physics B*, 236(2):397 – 422, 1984.
- [104] A. Hart, G. M. von Hippel, and R. R. Horgan. Radiative corrections to the lattice gluon action for HISQ improved staggered quarks and the effect of such corrections on the static potential. *Phys. Rev.*, D79:074008, 2009.
- [105] E. Follana, Q. Mason, C. Davies, K. Hornbostel, G. P. Lepage, J. Shigemitsu, H. Trotter, and K. Wong. Highly improved staggered quarks on the lattice, with applications to charm physics. *Phys. Rev.*, D75:054502, 2007.
- [106] Kenneth G. Wilson. Confinement of Quarks. *Phys. Rev.*, D10:2445–2459, 1974. [,319(1974)].
- [107] John B. Kogut and Leonard Susskind. Hamiltonian Formulation of Wilson’s Lattice Gauge Theories. *Phys. Rev.*, D11:395–408, 1975.

- [108] Christopher Monahan, Junko Shigemitsu, and Ron Horgan. Matching lattice and continuum axial-vector and vector currents with nonrelativistic QCD and highly improved staggered quarks. *Phys. Rev.*, D87(3):034017, 2013.
- [109] I. Allison et al. High-Precision Charm-Quark Mass from Current-Current Correlators in Lattice and Continuum QCD. *Phys. Rev.*, D78:054513, 2008.
- [110] C. T. H. Davies, C. McNeile, E. Follana, G. P. Lepage, H. Na, and J. Shigemitsu. Update: Precision D_s decay constant from full lattice QCD using very fine lattices. *Phys. Rev.*, D82:114504, 2010.
- [111] G. C. Donald, C. T. H. Davies, R. J. Dowdall, E. Follana, K. Hornbostel, J. Koponen, G. P. Lepage, and C. McNeile. Precision tests of the J/ψ from full lattice QCD: mass, leptonic width and radiative decay rate to η_c . *Phys. Rev.*, D86:094501, 2012.
- [112] Sasa Prelovsek. Lattice QCD review of charmonium and open-charm spectroscopy. In *Proceedings, 6th International Workshop on Charm Physics (Charm 2013): Manchester, UK, August 31-September 4, 2013*, 2013.
- [113] C. McNeile, C. T. H. Davies, E. Follana, K. Hornbostel, and G. P. Lepage. High-Precision f_{B_s} and HQET from Relativistic Lattice QCD. *Phys. Rev.*, D85:031503, 2012.
- [114] A. Bazavov et al. B - and D -meson leptonic decay constants from four-flavor lattice QCD. 2017.
- [115] C. McNeile, C. T. H. Davies, E. Follana, K. Hornbostel, and G. P. Lepage. High-Precision c and b Masses, and QCD Coupling from Current-Current Correlators in Lattice and Continuum QCD. *Phys. Rev.*, D82:034512, 2010.
- [116] Andrew Lytle, Brian Colquhoun, Christine Davies, Jonna Koponen, and Craig McNeile. Semileptonic B_c decays from full lattice QCD. *PoS, BEAUTY2016:069*, 2016.
- [117] A. Gray, I. Allison, C. T. H. Davies, Emel Dalgic, G. P. Lepage, J. Shigemitsu, and M. Wingate. The Upsilon spectrum and $m(b)$ from full lattice QCD. *Phys. Rev.*, D72:094507, 2005.

-
- [118] Christine T. H. Davies, Judd Harrison, Ciaran Hughes, Ronald R. Horgan, Georg M. von Hippel, and Matthew Wingate. Improving the Kinetic Couplings in Lattice Non-Relativistic QCD. 2018.
- [119] R. J. Dowdall et al. The Upsilon spectrum and the determination of the lattice spacing from lattice QCD including charm quarks in the sea. *Phys. Rev.*, D85:054509, 2012.
- [120] Steven Gottlieb, W. Liu, D. Toussaint, R. L. Renken, and R. L. Sugar. Hybrid-molecular-dynamics algorithms for the numerical simulation of quantum chromodynamics. *Phys. Rev. D*, 35:2531–2542, Apr 1987.
- [121] David J. E. Callaway and Aneesur Rahman. Microcanonical ensemble formulation of lattice gauge theory. *Phys. Rev. Lett.*, 49:613–616, Aug 1982.
- [122] David J. E. Callaway and Aneesur Rahman. Lattice gauge theory in the microcanonical ensemble. *Phys. Rev. D*, 28:1506–1514, Sep 1983.
- [123] Simon Duane and John B. Kogut. Hybrid stochastic differential equations applied to quantum chromodynamics. *Phys. Rev. Lett.*, 55:2774–2777, Dec 1985.
- [124] Simon Duane and John Kogut. The theory of hybrid stochastic algorithms. *Nuclear Physics B*, 275:398–420, 11 1986.
- [125] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092, 1953.
- [126] Simon Duane, A.D. Kennedy, Brian J. Pendleton, and Duncan Roweth. Hybrid monte carlo. *Physics Letters B*, 195(2):216 – 222, 1987.
- [127] K. Jansen. Actions for dynamical fermion simulations: are we ready to go? *Nuclear Physics B - Proceedings Supplements*, 129-130:3 – 16, 2004. Lattice 2003.
- [128] Michael Creutz. Chiral anomalies and rooted staggered fermions. *Physics Letters B*, 649(2):230 – 234, 2007.
- [129] Michael Creutz. Why rooting fails. *PoS*, LATTICE2007:007, 2007.

-
- [130] Andreas S. Kronfeld. Lattice Gauge Theory with Staggered Fermions: How, Where, and Why (Not). *PoS*, LATTICE2007:016, 2007.
- [131] Stephen R. Sharpe. Rooted staggered fermions: Good, bad or ugly? *PoS*, LAT2006:022, 2006.
- [132] E. Follana, A. Hart, and C. T. H. Davies. The Index theorem and universality properties of the low-lying eigenvalues of improved staggered quarks. *Phys. Rev. Lett.*, 93:241601, 2004.
- [133] Gordon C. Donald, Christine T. H. Davies, Eduardo Follana, and Andreas S. Kronfeld. Staggered fermions, zero modes, and flavor-singlet mesons. *Phys. Rev.*, D84:054504, 2011.
- [134] Stephan Durr. Physics of η' with rooted staggered quarks. *Phys. Rev.*, D85:114503, 2012.
- [135] M. A. Clark and A. D. Kennedy. Accelerating dynamical fermion computations using the rational hybrid Monte Carlo (RHMC) algorithm with multiple pseudofermion fields. *Phys. Rev. Lett.*, 98:051601, 2007.
- [136] Andreas Frommer, Bertold Nockel, Stephan Gusken, Thomas Lippert, and Klaus Schilling. Many masses on one stroke: Economic computation of quark propagators. *Int. J. Mod. Phys.*, C6:627–638, 1995.
- [137] Beat Jegerlehner. Krylov space solvers for shifted linear systems. 1996.
- [138] Bipasha Chakraborty, C. T. H. Davies, P. G. de Oliveira, J. Koponen, G. P. Lepage, and R. S. Van de Water. The hadronic vacuum polarization contribution to a_μ from full lattice QCD. *Phys. Rev.*, D96(3):034516, 2017.
- [139] Bipasha Chakraborty, C. T. H. Davies, B. Galloway, P. Knecht, J. Koponen, G. C. Donald, R. J. Dowdall, G. P. Lepage, and C. McNeile. High-precision quark masses and QCD coupling from $n_f = 4$ lattice QCD. *Phys. Rev.*, D91(5):054508, 2015.
- [140] C. McNeile. private communication, 2015.
- [141] R. J. Dowdall, C. T. H. Davies, G. P. Lepage, and C. McNeile. V_{us} from π and K decay constants in full lattice QCD with physical u , d , s and c quarks. *Phys. Rev.*, D88:074504, 2013.

- [142] M. R. Hestenes and E. Stiefel. Methods of conjugate gradients for solving linear systems. *Journal of research of the National Bureau of Standards*, 49:409–436, 1952.
- [143] Jonathan Richard Shewchuk. An introduction to the conjugate gradient method without the agonizing pain. Technical report, 1994.
- [144] D. Guadagnoli, F. Mescia, and S. Simula. Lattice study of semileptonic form-factors with twisted boundary conditions. *Phys. Rev.*, D73:114504, 2006.
- [145] Marc Wagner, Stefan Diehl, Till Kuske, and Johannes Weber. An introduction to lattice hadron spectroscopy for students without quantum field theoretical background. 2013.
- [146] G.P.Lepage. Corrfitter: <https://github.com/gplepage/corrfitter>, 2012.
- [147] G. P. Lepage, B. Clark, C. T. H. Davies, K. Hornbostel, P. B. Mackenzie, C. Morningstar, and H. Trottier. Constrained curve fitting. *Nucl. Phys. Proc. Suppl.*, 106:12–20, 2002. [,12(2001)].
- [148] G. Parisi. The Strategy for Computing the Hadronic Mass Spectrum. *Phys. Rept.*, 103:203–211, 1984.
- [149] G. Peter Lepage. The Analysis of Algorithms for Lattice Field Theory. In *Boulder ASI 1989:97-120*, pages 97–120, 1989.
- [150] Colin J. Morningstar and J. Shigemitsu. Perturbative matching of lattice and continuum heavy light currents with NRQCD heavy quarks. *Phys. Rev.*, D59:094504, 1999.
- [151] C. Hughes, C. T. H. Davies, and C. J. Monahan. New methods for B meson decay constants and form factors from lattice NRQCD. *Phys. Rev.*, D97(5):054509, 2018.
- [152] Christopher J Monahan, Heechang Na, Chris M Bouchard, G Peter Lepage, and Junko Shigemitsu. $B_s \rightarrow D_s \ell \nu$ Form Factors and the Fragmentation Fraction Ratio f_s/f_d . 2017.
- [153] Bipasha Chakraborty, C. T. H. Davies, B. Galloway, P. Knecht, J. Koponen, G. C. Donald, R. J. Dowdall, G. P. Lepage, and C. McNeile. High-precision quark masses and qcd coupling from $n_f = 4$ lattice qcd. *Phys. Rev. D*, 91:054508, Mar 2015.

- [154] T. C. Hammant, A. G. Hart, G. M. von Hippel, R. R. Horgan, and C. J. Monahan. Radiative improvement of the lattice nonrelativistic QCD action using the background field method with applications to quarkonium spectroscopy. *Phys. Rev.*, D88(1):014505, 2013. [Erratum: *Phys. Rev.*D92,no.11,119904(2015)].
- [155] Claude Bourrely, Laurent Lellouch, and Irinel Caprini. Model-independent description of $B \rightarrow \pi l \nu$ decays and a determination of $|V_{ub}|$. *Phys. Rev. D*, 79:013008, Jan 2009.
- [156] H. Schroder. New results on heavy quarks from ARGUS and CLEO. In *Heavy quark physics. Proceedings, 138th WE-Heraeus Seminar, Bad Honnef, Germany, December 14-16, 1994*, pages 9–22, 1994.
- [157] D. Bortoletto et al. Exclusive and inclusive decays of b mesons into d_s mesons. *Phys. Rev. Lett.*, 64:2117–2120, Apr 1990.
- [158] R. Fulton et al. Exclusive and inclusive semileptonic decays of b mesons to d mesons. *Phys. Rev. D*, 43:651–663, Feb 1991.
- [159] H. Albrecht et al. Measurement of the decay $B^0 \rightarrow D^{*0} l^- \nu$. *Physics Letters B*, 275(1):195 – 201, 1992.
- [160] B. Barish et al.
- [161] D. Buskulic et al. Strange b baryon production and lifetime in z decays. *Physics Letters B*, 384(1):449 – 460, 1996.
- [162] D. Buskulic et al. Heavy quark tagging with leptons in the ALEPH detector. *Nucl. Instrum. Meth.*, A346:461–475, 1994.
- [163] G. Abbiendi et al. Measurement of $|V_{cb}|$ using $\bar{B}^0 \rightarrow D^{*+} \ell^- \bar{\nu}$. *Phys. Lett.*, B482:15–30, 2000.
- [164] P. Abreu et al. Measurement of V_{cb} from the decay process $\bar{B}^0 \rightarrow D^{*+} \ell^- \bar{\nu}$. *Phys. Lett.*, B510:55–74, 2001.
- [165] N. E. Adam et al. Determination of the $\bar{B} \rightarrow D^* \ell \bar{\nu}$ width and $|V_{cb}|$. *Phys. Rev.*, D67:032001, 2003.
- [166] J. Abdallah et al. Measurement of $|V_{cb}|$ using the semileptonic decay $\bar{B}_d^0 \rightarrow D^{*+} \ell^- \bar{\nu}_l$. *Eur. Phys. J.*, C33:213–232, 2004.

- [167] Bernard Aubert et al. Determination of the form-factors for the decay $B^0 \rightarrow D^{*-}\ell^+\nu_\ell$ and of the CKM matrix element $|V_{cb}|$. *Phys. Rev.*, D77:032002, 2008.
- [168] Bernard Aubert et al. Measurement of the Decay $B^- \rightarrow D^{*0} e^- \bar{\nu}(e)$. *Phys. Rev. Lett.*, 100:231803, 2008.
- [169] W. Dungel et al. Measurement of the form factors of the decay $B^0 \rightarrow D^{*-}\ell^+\nu$ and determination of the CKM matrix element $|V_{cb}|$. *Phys. Rev.*, D82:112007, 2010.
- [170] A. Abdesselam et al. Precise determination of the CKM matrix element $|V_{cb}|$ with $\bar{B}^0 \rightarrow D^{*+}\ell^- \bar{\nu}_\ell$ decays with hadronic tagging at Belle. 2017.
- [171] A. Abdesselam et al. Measurement of CKM Matrix Element $|V_{cb}|$ from $\bar{B} \rightarrow D^{*+}\ell^- \bar{\nu}_\ell$. 2018.
- [172] Jack Laiho and Ruth S. Van de Water. $B \rightarrow D^*\ell\nu$ and $B \rightarrow D\ell\nu$ form factors in staggered chiral perturbation theory. *Phys. Rev.*, D73:054501, 2006.
- [173] Jonathan Flynn, Taku Izubuchi, Andreas Juttner, Taichi Kawanai, Christoph Lehner, Edwin Lizarazo, Amarjit Soni, Justus Tobias Tsang, and Oliver Witzel. Form factors for semi-leptonic B decays. *PoS, LATTICE2016:296*, 2016.
- [174] Jon A. Bailey, Tanmoy Bhattacharya, Rajan Gupta, Yong-Chull Jang, Weonjong Lee, Jaehoon Leem, Sungwoo Park, and Boram Yoon. Calculation of $\bar{B} \rightarrow D^*\ell\bar{\nu}$ form factor at zero recoil using the Oktay-Kronfeld action. *EPJ Web Conf.*, 175:13012, 2018.
- [175] K. G. Wilson. Quarks and Strings on a Lattice. *New Phenomena in Subnuclear Physics*, Part A:69–143, 1977.
- [176] B. Sheikholeslami and R. Wohlert. Improved continuum limit lattice action for qcd with wilson fermions. *Nuclear Physics B*, 259(4):572 – 596, 1985.
- [177] Aida X. El-Khadra, Andreas S. Kronfeld, and Paul B. Mackenzie. Massive fermions in lattice gauge theory. *Phys. Rev.*, D55:3933–3957, 1997.
- [178] skopt: <https://github.com/scikit-optimize/scikit-optimize>.
- [179] G. C. Donald, C. T. H. Davies, J. Koponen, and G. P. Lepage. V_{cs} from $D_s \rightarrow \phi\ell\nu$ semileptonic decay and full lattice QCD. *Phys. Rev.*, D90(7):074506, 2014.

- [180] Matthias Neubert. B decays and the heavy quark expansion. *Adv. Ser. Direct. High Energy Phys.*, 15:239–293, 1998. [,239(1997)].
- [181] G. 't Hooft. *Can We Make Sense Out of “Quantum Chromodynamics”?*, pages 943–982. Springer US, Boston, MA, 1979.
- [182] A. Bazavov et al. Up-, down-, strange-, charm-, and bottom-quark masses from four-flavor lattice QCD. *Phys. Rev.*, D98(5):054517, 2018.
- [183] Andrzej Czarnecki. Two-loop qcd corrections to $b \rightarrow c$ transitions at zero recoil. *Phys. Rev. Lett.*, 76:4124–4127, May 1996.
- [184] F.E. Close, G.J. Gounaris, and J.E. Paschalis. The ratio $\Gamma(B \rightarrow D\ell\nu)/\Gamma(B \rightarrow D^*\ell\nu)$ in the quark model. *Physics Letters B*, 149(1):209 – 212, 1984.
- [185] Nathan Isgur, Daryl Scora, Benjamin Grinstein, and Mark B. Wise. Semileptonic b and d decays in the quark model. *Phys. Rev. D*, 39:799–818, Feb 1989.
- [186] D. Buskulic et al. Measurements of $|V_{cb}|$, form-factors and branching fractions in the decays $\bar{B}^0 \rightarrow D^{*+}\ell^-\bar{\nu}_\ell$ and $\bar{B}^0 \rightarrow D^+\ell^-\bar{\nu}_\ell$. *Phys. Lett.*, B395:373–387, 1997.
- [187] R. Glattauer et al. Measurement of the decay $B \rightarrow D\ell\nu_\ell$ in fully reconstructed events and determination of the Cabibbo-Kobayashi-Maskawa matrix element $|V_{cb}|$. *Phys. Rev.*, D93(3):032006, 2016.
- [188] Vardan Khachatryan et al. Observation of the rare $B_s^0 \rightarrow \mu^+\mu^-$ decay from the combined analysis of CMS and LHCb data. *Nature*, 522:68–72, 2015.
- [189] Jon A. Bailey, A. Bazavov, C. Bernard, C. M. Bouchard, C. DeTar, Daping Du, A. X. El-Khadra, J. Foley, E. D. Freeland, E. Gámiz, Steven Gottlieb, U. M. Heller, Jongjeong Kim, A. S. Kronfeld, J. Laiho, L. Levkova, P. B. Mackenzie, Y. Meurice, E. Neil, M. B. Oktay, Si-Wei Qiu, J. N. Simone, R. Sugar, D. Toussaint, R. S. Van de Water, and Ran Zhou. $B_s \rightarrow D_s/b \rightarrow d$ semileptonic form-factor ratios and their application to $\text{BR}(B_s^0 \rightarrow \mu^+\mu^-)$. *Phys. Rev. D*, 85:114502, Jun 2012.
- [190] Mariam Atoui, Vincent Morénas, Damir Bečirević, and Francesco Sanfilippo. $B_s \rightarrow D_s\ell\nu_\ell$ near zero recoil in and beyond the Standard Model. *The European Physical Journal C*, 74(5):2861, May 2014.

-
- [191] Witzel O., Flynn J., Juettner A., Soni A., and Tsang J. T. Semi-leptonic form factors for $B_s \rightarrow Kl\nu$ and $B_s \rightarrow D_s l\nu$. 36th Annual International Symposium on Lattice Field Theory, July 22-28, 2018, East Lansing, MI, USA, 2018.
- [192] J. Koponen, C. T. H. Davies, G. C. Donald, E. Follana, G. P. Lepage, H. Na, and J. Shigemitsu. The shape of the $D \rightarrow K$ semileptonic form factor from full lattice QCD and V_{cs} . 2013.
- [193] Claude Bourrely, Irinel Caprini, and Laurent Lellouch. Model-independent description of $B \rightarrow \pi l\nu$ decays and a determination of $|V_{ub}|$. *Phys. Rev.*, D79:013008, 2009. [Erratum: *Phys. Rev.*D82,099902(2010)].
- [194] R. J. Dowdall, C. T. H. Davies, T. C. Hammant, and R. R. Horgan. Precise heavy-light meson masses and hyperfine splittings from lattice QCD including charm quarks in the sea. *Phys. Rev.*, D86:094510, 2012.
- [195] Susumu Okubo. New improved bounds for K_{l3} parameters. *Phys. Rev. D*, 4:725–733, Aug 1971.
- [196] Susumu Okubo. Exact bounds for K_{l3} decay parameters. *Phys. Rev. D*, 3:2807–2813, Jun 1971.
- [197] Thomas Becher and Richard J. Hill. Comment on form-factor shape and extraction of $|V_{ub}|$ from $B \rightarrow \pi l\nu$. *Physics Letters B*, 633(1):61 – 69, 2006.