The strange decays of heavy quarks

The long and winding road that leads to new physics

William George Parrott

Supervised by Dr. Chris Bouchard

Submitted in fulfillment of the requirements for the degree of
Doctor of Philosophy
June 2022

School of Physics and Astronomy
College of Science and Engineering
University of Glasgow

Abstract

This thesis presents the results of lattice Quantum Chromodynamics (QCD) calculations of hadronic matrix elements for pseudoscalar to pseudoscalar semileptonic decays involving $c \to s$ and $b \to s$ weak interaction transitions. We calculate scalar, vector and tensor form factors, and use them to test the Standard Model. Most prominently, we use $D \to K\ell\nu$ form factors, calculated on the lattice in the Highly Improved Staggered Quark formalism, to calculate a much more precise determination of the Cabbibo-Kobayashi-Maskawa (CKM) matrix element $|V_{cs}|$, allowing for more stringent tests of CKM unitarity. We also calculate branching fractions and other experimentally measurable quantities for the rare $B \to K\ell\ell$ decay, finding significant tension with experiment in this one-loop level process, which is sensitive to physics beyond the Standard Model.

Acknowledgements

Firstly I would like to thank my supervisors, Chris Bouchard and Christine Davies, for their patience, support, and their resolve to weather the blizzard of my stupid questions. Thanks also to Dan Hatton, Judd Harrison, Peter Lepage, Laurence Cooper and Euan McLean for help with all aspects of lattice calculations. Thanks to Stephen for his long arms and shallow pockets, Panos for his cooking tips, and of course Hugh Tennent Tennent (1863-1890) for his contribution to inspirational beverages.

Declaration of originality

Aside from the well established introductory material in Chapters 2 and 3, the work presented here is my own, working in collaboration with others in the HPQCD collaboration. The content of Chapters 5, 6 and 7 is based on work published in [1, 2, 3, 4]. Many of the figures and captions in this thesis appear as in those papers, and are also my own work.

"I'm not claiming that what I did earlier wasn't incredibly stupid, but I don't think the situation can be improved by further bad decisions." - Ed Thomas

Contents

1 Introduction 1

2 The Standard Model of particle physics 3

2.1 The basis of the Standard Model . . . 3

2.1.1 Spin-half fields . . . 4

2.1.2 Gauge field theories . . . 5

2.1.3 Quantum Chromodynamics . . . 7

2.1.4 The electroweak sector . . . 8

2.2 The field content of the Standard Model . . . 12

2.3 Weak Decays . . . 14

2.3.1 Flavour changing charged currents . . . 14

2.3.2 Flavour changing neutral currents . . . 17

2.3.3 Semileptonic decays . . . 20

2.4 The path integral formalism . . . 21

2.4.1 Correlation functions . . . 23

3 Lattice Quantum Chromodynamics 25

3.1 Motivation . . . 25

3.2 Basic principles . . . 25
3.2.1 Discretising the path integral . . . 26
3.2.2 The gluon action . . . 28
3.3 Staggered quarks and the HISQ formalism . . . 30
3.3.1 The naive quark action and doubling . . . 30
3.3.2 Symanzik improvement at tree level . . . 32
3.3.3 Symanzik improvement at one loop . . . 33
3.3.4 Staggered quarks . . . 34
3.3.5 Twisted boundary conditions . . . 35
3.3.6 Quarks on the lattice . . . 36
3.4 HISQ ensembles . . . 39

4 Methods for lattice calculations of form factors **41**
4.1 A note on units . . . 41
4.2 Calculating correlation functions . . . 42
4.2.1 Phases in staggered quarks . . . 43
4.2.2 Sources and random walls . . . 52
4.3 Fitting correlation functions . . . 55
4.3.1 Fit forms . . . 55
4.3.2 Bayesian fitting . . . 57
4.3.3 Prior selection . . . 59
4.3.4 SVD cuts . . . 60
4.3.5 Goodness of fit . . . 63
4.4 Form factors . . . 64

4.4.1 Z_V and Z_T . . . 66
4.4.2 The modified z expansion . . . 67
4.5 Heavy-HISQ . . . 73
4.5.1 The problem . . . 73
4.5.2 The solution . . . 74
4.5.3 Heavy-HISQ in action . . . 75

5 $D \to K$ form factors: an improved determination of V_{cs} 76
5.1 Introduction . . . 76
5.1.1 Previous status of V_{cs} . . . 77
5.2 Lattice calculation . . . 80
5.2.1 Simulation details . . . 80
5.2.2 Correlation functions . . . 80
5.2.3 Correlator fits . . . 81
5.2.4 z expansion and continuum extrapolation . . . 85
5.2.5 A more agnostic approach; cubic splines . . . 88
5.3 Form factor results . . . 89
5.3.1 Form factor shape . . . 91
5.4 Extracting the differential decay rate and V_{cs} determination . . . 96
5.4.1 Isospin breaking, $\eta_{\rm EW}$ and $\delta_{\rm EM}$ . . . 96
5.4.2 Tests of flavour universality . . . 97
5.4.3 V_{cs} determination . . . 100
5.5 Discussion and conclusions . . . 108

6 $B_s \to \eta_s$ form factors **112**

6.1 Introduction . . . 112

6.2 Calculation details . . . 113

6.2.1 Simulation details . . . 113

6.2.2 Correlation functions . . . 114

6.2.3 Correlator fits . . . 114

6.2.4 z expansion: incorporating heavy mass dependence . . . 118

6.3 Results and discussion . . . 124

6.3.1 Form factor results . . . 124

6.3.2 Tests of HQET . . . 129

6.4 Conclusions . . . 133

7 $B \to K$ form factors **134**

7.1 Introduction . . . 134

7.2 Calculation details . . . 136

7.2.1 Simulation details . . . 136

7.2.2 Correlation functions . . . 137

7.2.3 Spatial and temporal vector components . . . 137

7.2.4 Correlator fits . . . 138

7.2.5 z expansion . . . 144

7.3 Form Factor results . . . 154

7.3.1 $B \to K$ form factors . . . 154

7.3.2 Form factors and uncertainties across the M_H range . . . 157

7.3.3 Spectator quark dependence of form factors . . . 163

7.4 Key phenomenology 165
7.4.1 $B \to K\ell^+\ell^-$ 165
7.4.2 $B \to K\nu\bar{\nu}$ 176
7.5 Numerical results 177
7.6 Conclusions 180

8 Conclusions **181**

List of Figures

2.1 Unitarity triangle of the CKM matrix (courtesy of CKMfitter [31]). 16

2.2 Feynman diagram for a FCCC semileptonic weak decay $M_1 \to M_2 \ell \nu$. 17

2.3 Example of a FCNC Feynman diagram for $B^+ \to K^+ \ell^+ \ell^-$. 17

4.1 Schematic of a two-point correlation function. 46

4.2 Schematic of a three-point correlation function. 50

4.3 Example of two- and three-point correlator data on the ultrafine 5 (set 8) ensemble. In both cases error bars are too small to be visible. The left hand plot is the log of the (folded) two-point data for a heavy-light H meson with $am_h = 0.194$ (actually a D meson). 'BG' is a Goldstone pseudoscalar with spin-taste $\gamma_5 \otimes \xi_5$ and 'BNG' a non-Goldstone pseudoscalar ($\gamma_5\gamma_0 \otimes \xi_5\xi_0$). Oscillations are much larger in the latter case, causing some negative values which are ignored by the log. On the right is the log of a three-point function for a $H \to K$ decay with a scalar current insertion, for $am_h = 0.6$ and twist $\theta = 1.529$. Three different T values are shown. 57

4.4 Example [2] of estimating priors for the ground state mass, and V_{00}^{nn} using Equations 4.44 and 4.46. Priors in green are based on the data points, whilst the fit result (including error band) is in blue. 59

4.5 SVD diagnosis plot for the $D \to K$ coarse 5 (set 5) fit described in the text. The horizontal black dotted line indicates $1 - \sqrt{\frac{2}{N_G}}$, whilst the vertical red dotted lines gives the $\kappa = \lambda_n / \lambda_n^{\max}$ value at which this line is crossed. 62

4.6 f_+ for the $t_0 = 0$ (f_+^0), 'minimum' ($f_+^{\rm min}$) and 'reverse' ($f_+^{\rm rev}$) choice (see text), as well as the weighted average of the three. In each case, the result is divided by the mean on the $t_0 = 0$ case. 69

4.7 The relative error on f_+ for the $t_0 = 0$ (f_+^0), 'minimum' ($f_+^{\rm min}$) and 'reverse' ($f_+^{\rm rev}$) choice (see text). In each case, the data points indicate the z values for the set 8 data points in [3], with lines joining points of the same mass, which share a symbol. 70

4.8 Scalar, vector and tensor form factors across the whole physical q^2 range evaluated at different masses M_H between M_D and M_B. This plot contains results from [3] which is the subject of Chapter 7. 75

5.1 Feynman diagram for $D^- \to K^0 \ell^- \bar{\nu}$ 77

5.2 Left: stability plot for our fit on the set 5 ($a = 0.12$ fm) lattice, with our preferred fit using $N = 4$ exponentials, shown as the value at location 0. The different panels show (from the top) the mass of the D (parameter $E_0^{D,{\rm n}}$), the ground state energy of the K (parameter $E_0^{K,{\rm n}}$) with the largest twist for this set of 2.646, and current matrix element for the temporal vector current $V_{00}^{\rm nn}$ for twist 0.441. Tests 1 and 2 give the results from including one fewer and one more exponential respectively. Test 3 increases $t_{\rm min}$ by 1 across the whole fit. Tests 4 and 5 double and halve the svd cut and tests 6 and 7 double and halve all prior widths. The final test, 8, shows the results when the single correlator is fit on its own or, in the case of $V_{00}^{\rm nn}$, just with the D and K two-point correlation functions required, rather than as part of one big simultaneous fit. Right: for each ensemble, we plot the ratio $(E_K^2 - M_K^2)/|\vec{p}_K|^2$ from our fit results against $|a\vec{p}_K|^2$ to check that the K meson energy agrees with that expected from the spatial momentum given to the meson in the lattice calculation. The points for gluon field configurations with physical sea u/d quark mass are in black. The ratio agrees with the expected value of 1 throughout the range of momenta and lattice spacing values. The purple wedge shows $1 \pm |a\vec{p}_K/\pi|^2$. 83

5.3 Left: the difference between the non-Goldstone $\hat{D}$ and Goldstone D meson masses, from our fit results, as a function of lattice spacing. The results show clearly that the splitting is a discretisation effect and is only a few MeV even on the coarsest lattices. Right: the renormalisation factor for the local temporal vector current, Z_V, plotted as a function of lattice spacing. The purple hexagons give results for Z_V values for the local vector current determined in a symmetric momentum-subtraction scheme on the lattice [95]. The two sets of Z_V values differ at finite lattice spacing by discretisation effects. In both cases the points in black correspond to gluon field configurations with physical u/d sea quark mass. 84

5.4 $D \to K$ f_0 and f_+ form factors in q^2 space using the z expansion, our preferred fit. Data points from the eight gluon ensembles are included, and are joined by lines corresponding to the fit at that lattice spacing (Equations (5.8) and (5.9)). The continuum, physical point results are the coloured bands. 90

5.5 Comparison between our preferred z expansion fit described in Section 5.2.4 (solid bands) and the spline fit described in Section 5.2.5 (hatched bands). 91

5.6 Stability test of the z-expansion fit; 0 marks our final result. Test 1 removes all the results from gluon field configurations with $m_s/m_l = 5$, so that only sets 1, 2 and 3 remain. Test 2 removes the results from sets 1, 2 and 3 and fits the others. Test 3 takes t_0 in the q^2 to z mapping to the 'minimum' prescription described in Section 4.4.2. Test 4 sets t_0 to t_-. Test 5 includes an extra term in the sums over n up to N and over j up to N_j (Equations (5.8), (5.9) and (5.11)). Test 6 removes the highest momentum data point for each gluon field ensemble (and highest two on set 7 so that there are no results included with $q^2 < 0$). Test 7 doubles the width of all 'd' priors (this decreased the Gaussian Bayes Factor), and 8 halves them. Test 9 sets the logarithmic factor $L(m_l)$ to zero (Equation (5.10)). Test 10 shows the results of a completely different kind of fit, a cubic spline fit in q^2 discussed in Section 5.2.5. Test 11 removes the $f_0(0) = f_+(0)$ constraint, in this case the black point is $f_0(0)$ and the red is $f_+(0)$. 92

5.7 Comparison of our lattice form factors at $q^2 = 0$ and q^2_{max} with earlier lattice QCD calculations. The points marked 'HPQCD '10' are from [94]; the points marked 'HPQCD '13' from [138] and the points marked 'ETMC '17' from [139, 140]. A preliminary analysis of the scalar form factor in [145] gives $f_0(0) = 0.768(16)$, but we have not plotted that point. Our new results are labelled 'HPQCD '21' and demonstrate a significant improvement in uncertainty over earlier values. 93

5.8 Left: uncertainties for f_0 and f_+ (for $m_u = m_d$) as a function of q^2. The red line 'Inputs' shows the uncertainties coming from fixed inputs, such as experimental meson masses used in the analysis. The purple line 'q mistunings' adds in uncertainties arising from mistuning of valence and sea quark masses. The blue 'Statistics' line further adds the statistical uncertainties from the lattice results (correlator fits). Finally, the black line ('Discretisation') gives the total uncertainty, now including the contribution from discretisation effects. These uncertainties add in quadrature, so we plot the squared percentage error and include an axis showing the corresponding percentage error on the right for clarity. Right: breakdown of the contributions to statistical uncertainty of the form factors at extremal q^2 values, from each of the eight ensembles (Table 3.1). The variance σ_i^2 from each ensemble is added linearly, and normalised by the total variance. . . . 94

5.9 Left: the original f_+ form factor, as well as the result using the refitting procedure described in Section 5.3.1. The black line (and grey uncertainty band, barely visible) is the ratio of the two. We see that the refitting reproduces the original form factor and its uncertainty accurately. Right: comparison of the shape of $f_+(q^2)$ for $D \to K$ expressed in terms of ratios of the z-expansion coefficients a_1^+ and a_2^+ to a_0^+. Ellipses give the 68% confidence limits ($\Delta\chi^2 = 2.3$). Experimental results are from [150, 151, 152, 153]. CLEO results are for $D^0 \to K^- e^+ \nu_e$ (dark blue) and $D^+ \to \bar{K}^0 e^+ \nu_e$ (light blue); all other experimental data is for $D^0 \to K^- e^+ \nu_e$. The HFLAV experimental average [153] is given as the red ellipse. Our results here are given by the black ellipse, showing good agreement. 96

5.10 Left: lepton flavour universality tests in $D \to K$ decay. The solid black curve as a function of q^2 shows the Standard Model ratio of branching fractions for a muon in the final state to that for an electron obtained from our form factors using Equation (5.19). The width of the curve gives the (very small) uncertainty from our results. Possible QED effects are not included here. The points, with error bars, are from the BES experiment [156]. For illustration the red and blue dashed lines show what the curve would look like in the presence of a new physics scalar coupling for the μ case (see Equation (5.23) for definition of ζ_S). Right: the forward-backward asymmetry of the muon produced in $D \to K$ decay. This is defined with respect to the angle θ_ℓ in the W rest-frame. The solid black line shows the SM result derived from our form factors, including the lattice QCD uncertainty but ignoring any uncertainty from possible QED corrections. For illustration the red and blue dashed lines show what the curve would look like in the presence of a new physics scalar coupling for the μ case (see text and Equation (5.23)). . . . 99

5.11 Plots of the determination of $|\eta_{\rm EW} V_{cs}|^2(1 + \delta_{\rm EM})$ per q^2 bin for CLEO D^0 and D^+ results [150]. The total uncertainty for each bin is given in black and this is broken down into experimental (blue) and theoretical (red) contributions, the latter coming from our form factors. Each data point is centred on the q^2 bin it corresponds to. Note that the uncertainties are correlated between q^2 bins. The purple band gives the weighted average for these data points, with all correlations included. 102

5.12 Plots of the determination of $|\eta_{\rm EW} V_{cs}|^2(1 + \delta_{\rm EM})$ per q^2 bin for BaBar D^0 [151] and BES D^0 [152] results. The total uncertainty for each bin is given in black and this is broken down into experimental (blue) and theoretical (red) contributions, the latter coming from our form factors. Each data point is centred on the q^2 bin it corresponds to. Note that the uncertainties are correlated between q^2 bins. The purple band gives the weighted average for these data points, with all correlations included. 103

5.13 Plot of the determination of $|\eta_{\rm EW}V_{cs}|^2(1+\delta_{\rm EM})$ per q^2 bin for BES D^+ results [158]. The total uncertainty for each bin is given in black and this is broken down into experimental (blue) and theoretical (red) contributions, the latter coming from our form factors. Each data point is centred on the q^2 bin it corresponds to. Note that the uncertainties are correlated between q^2 bins. The purple band gives the weighted average for these data points, with all correlations included. 103

5.14 Left: plot of $|V_{cs}|$ per bin for CLEO, BaBar and BES results from [150, 151, 152]. Each data point is centred on the q^2 bin it corresponds to and the error bars plotted include the uncertainties from $\eta_{\rm EW}$ and $\delta_{\rm EM}$. The purple line and band give the result from our total weighted average for $|V_{cs}|^2$, with all correlations included. Right: comparison plot of the determination of $|V_{cs}|$ using the differential decay rate for CLEO, BaBar and BES results from [150, 151, 152, 158] for D^0 and D^+ decays. The purple band gives the total weighted average for V_{cs}, not including the BES '17 result. 104

5.15 Left: comparison plot of the determination of $|V_{cs}|$ using the total branching fraction. Experimental results are from [160] for $D^+ \to \overline{K}^0\mu^+\nu_\mu$, [150, 161] for $D^+ \to \overline{K}^0e^+\nu_e$, from [156, 162] for $D^0 \to K^-\mu^+\nu_\mu$ and from [150, 151, 161, 162] for $D^0 \to K^-e^+\nu_e$ decays. Note that the BES results for final state e are the quoted averages for results from [161] and [152, 158]. The purple band gives the total average for V_{cs}, assuming 100% correlation of systematic uncertainties for results from a given experiment. The width of the purple band encompasses all uncertainties, including those from $\eta_{\rm EW}$ and $\delta_{\rm EM}$. Right: comparison plot of the determination of $|V_{cs}|$ using the extrapolation of experimental results to $q^2 = 0$. Experimental results are from [158, 163] for $D^+ \to \overline{K}^0e^+\nu_e$, from [156] for $D^0 \to K^-\mu^+\nu_\mu$ and [150, 151, 152, 162] for $D^0 \to K^-e^+\nu_e$. The purple band gives the weighted average result for V_{cs} obtained from the HFLAV weighted average [129] of the experimental results but including a correction for η_{EW} and an additional uncertainty from QED corrections. 107

5.16 Our $|V_{cs}|$ result compared with other $N_f = 2+1+1$ and $N_f = 2+1$ results using lattice QCD. Different symbols indicate different lattice calculations, whilst different colours indicate the method used. Blue indicates use of the differential rate in q^2 bins, red indicates use of the $f_+(0)$ method and green indicates use of the total branching fraction for the decay. Points marked 'HPQCD '21' come from this work, 'ETMC '17' is from [139, 140], 'HPQCD '13' is from [138], 'HPQCD '10' is from [94] and 'Fermilab/MILC '04' is from [137]. For comparison we give at the bottom the value currently quoted in the Particle Data Tables [29] from semileptonic $D \to K$ decay. The blue band carries our preferred result, $V_{cs}^{\mathrm{d}\Gamma/\mathrm{d}q^2}$, down the plot. 109

5.17 Left: a comparison of constraints on V_{cs} and V_{cd} with the expectation from CKM unitarity. Red bands show the $\pm 1\sigma$ range for the determination of V_{cs} and V_{cd} from leptonic decays of D_s and D^+ combined with decay constants from lattice QCD. The diagonal red band is the constraint from the ratio of leptonic rates for D_s and D^+ combined with the lattice QCD ratio of decay constants. The solid light blue band shows the result for V_{cd} from the $D \to \pi\ell\overline{\nu}$ decay combined with lattice QCD form factor results. See text for a discussion of the values used. The darker blue band shows our new determination here of V_{cs} from $D \to K\ell\overline{\nu}$ with $\pm 1\sigma$ uncertainties. For comparison the black dashed line gives the unitarity constraint curve of $|V_{cd}|^2 + |V_{cs}|^2 + |V_{cb}|^2 = 1$. Right: a similar comparison to above, this time of constraints on V_{cs} and V_{us}. Red bands show the $\pm 1\sigma$ range for the determination of V_{cs} and V_{us} from leptonic decays of D_s and K^+ combined with decay constants from lattice QCD. The light blue band shows the result for V_{us} from $K \to \pi\ell\overline{\nu}$ decay combined with lattice QCD form factor results. See text for a discussion of the values used. The darker blue band again shows our new determination of V_{cs} and the black dashed line gives the unitarity constraint curve, $|V_{us}|^2 + |V_{cs}|^2 + |V_{ts}|^2 = 1$. 110

6.1 Left: stability tests of the chained, marginalised fit used on a typical three-point correlator. Test 0, the final result, shows the value of V_{00}^{nn} for $am_h = 0.45$, $a|\vec{p}| = 0.1430$ on set 7, with $N_{\mathrm{exp}}^{\mathrm{3pts}} = 3$ exponential terms and three additional states marginalised (as discussed in the text), with $t_{\mathrm{min}} = 2$, the number of data points removed from the fit at the start and end of the data. Tests 1 and 2 show the effects of increasing and decreasing the number of fitted exponentials by 1, tests 3 and 4 show the effect of doubling and halving the SVD cut, and 5 and 6 show the effect of doubling and halving the error on all priors. Test 7 shows the effect of an increase on t_{min} by 1, and test 8 shows the reduction of the marginalised exponentials from 6 to 5. Finally, test 9 shows the result of just fitting the vector 3 point correlator for this mass and twist, and the relevant 2 points; this gives a reduced error since the smaller fit requires a smaller SVD cut. Fitting like this does not preserve correlations, however, so we use a global fit. Other two and three-point correlators behaved similarly well under the same tests. The χ^2/d.o.f. values (purple $\times$s) are also plotted for reference. Note that these are the raw values and hence artificially small (see text) and the degrees of freedom are not the same across all tests. Right: we plot the ratio $(E_{\eta_s}^2 - M_{\eta_s}^2)/|\vec{p}_{\eta_s}|^2$ from our fit results against $|a\vec{p}_{\eta_s}|^2$ to check that the η_s energy in our final fit results agrees with the momentum given to the meson in the lattice calculation. 117

6.2 $\left(1 - \frac{q^2}{M_{H_{s0}^*}^2}\right) f_0(z)$ (top) and $\left(1 - \frac{q^2}{M_{H_s^*}^2}\right) f_+(z)$ (bottom) data points and final result at the physical point (coloured bands). Data points are labelled by mass for sets 6, 7 and 8 respectively, where e.g. m0.8 indicates $am_h = 0.8$ on that ensemble. Lines between data points of a given heavy mass over the full z range are there to guide the eye. For f_0, the additional continuum data point from [171] is shown in purple, and helps to pin down the form factor in the high q^2 limit. 125

6.3 Top: final form factor results for $f_0(q^2)$ and $f_+(q^2)$. Results from [104] at $q^2 = 0$ and $q^2 = q^2_{\text{max}}$ are also shown. Bottom: stability tests of the fit of the form factors $f_{0,+}(0)$, $f_0(q^2_{\text{max}})$ and $f_+(q^2_{\text{max}})$. Test 0 is the final result, shown throughout by the blue band. Tests 1, 2 and 3 are the results if the fine, superfine and ultrafine data are removed respectively. Test 4 is the fit without the data point from [171]. Test 5 adds a cubic term in the z expansion (Equations (6.1) and (6.2)). Test 6 shows the effect of extending the i, j, k sum in Equation (6.4) Tests 7 and 8 remove the highest masses and momenta for all lattice spacings respectively. Test 9 is without the log term in Equation (6.4), here we find that d_{i000} terms change to mimic the Taylor expansion of the log, and we require much larger priors (0.0(5.0)) to account for this. Test 10 shows the effect of doubling the width of all d_{ijkn} priors. We see that our extrapolation is stable to all of the above modifications. GBF values again tell us that our priors are conservative. 126

6.4 The contributions to the total percentage error (black line) of $f_0(q^2)$ (top) and $f_+(q^2)$ (bottom) from different sources, shown as an accumulating error. The red dashed line ('inputs') includes values for masses taken from the PDG [29] and used in the fit as described above. The purple dotted line ('q mistunings') adds, negligibly, to the inputs the error contribution from the quark mistunings associated with c fit parameters, whilst the solid green line ('statistics') further adds the error from our correlator fits. The blue dot-dash line ('HQET') includes the contribution from the expansion in the heavy quark mass, and, finally, the thick black line ('Discretisation'), the total error on the form factor, also includes the discretisation errors. The percentage variance adds linearly and the scale for this is given on the left hand axis. The percentage standard deviation, the square root of this, can be read from the scale on the right-hand side. 127

6.5 The form factors $f_{0,+}(0)$, $f_0(q^2_{\text{max}})$ and $f_+(q^2_{\text{max}})$ over the range of heavy masses from the physical D_s to the physical B_s. Results are included for $f_{0,+}(0)$, $f_0(q^2_{\text{max}})$ and $f_+(q^2_{\text{max}})$ (in their respective colours) for several other decays related by SU(3) flavour symmetry [104, 138, 139, 172, 173]. Data points are plotted at the x axis values corresponding to their physical heavy meson mass, not the mass that would result from their heavy quark and a strange quark (which would put them all at M_{D_s} or M_{B_s}). In the case of M_B and M_D some of the points are offset slightly either side of the mass for clarity. 128

6.6 Left: the quantities α, β^{-1} and δ, defined in Equations (6.6), (6.7) and (6.8), over the range of heavy masses from the physical D_s to the physical B_s. Right: the form factor ratio, $\frac{f_0(q^2)}{f_+(q^2)}\left(1-\frac{q^2}{M^2_{B_s^*}}\right)^{-1}$ over the range $0 \leq q^2 \leq M^2_{B_s}$ (blue band), as compared with the HQET expectation in the limit $q^2 \to M^2_{B_s}$ (red band), defined in Equation (6.9). 130

6.7 Left: the form factor ratio $\frac{f_0^{H_s}(q^2(E))}{f_0^{B_s}(q^2(E))}$ evaluated at η_s energy $E = E_{\text{min}} = M_{\eta_s} = 0.6885(22)\,\text{GeV}$ (blue line and error band) and at E_{max} corresponding to the largest energy available to the η_s in a D_s decay (red line and error band). Both ratios are plotted over a range of inverse heavy meson masses up to $M_{D_s}^{-1}$. The black dashed line marks $M_{H_s}^{-1} = M_{B_s}^{-1}$. Results are compared with the expectation of $\sqrt{\frac{M_{B_s}}{M_{H_s}}}$ [174], given by the black band (see text). Right: the form factor ratio $\frac{f_+^{B_s}(0)}{f_+^{H_s}(0)}$ plotted against the meson mass ratio M_{H_s}/M_{B_s} in a log-log plot. Our results are shown as a blue curve with error band. The HQET expectation that the form factor ratio should depend on the 3/2 power of the mass ratio is shown as a black dashed line. In contrast, the red dashed line shows linear dependence on the mass. Results for the D_s meson correspond to the left-hand end of the plot, $\log(M_{D_s}/M_{B_s}) = -1.003$. 132

7.1 Example of a Feynman diagram for $B^+ \to K^+\ell^+\ell^-$. 134

7.2 Left: a comparison of values and their statistical errors for the vector form factor derived from matrix elements for the spatial and temporal vector currents on ensembles where both are available. The filled symbols are the temporal vector results and the open symbols the spatial vector results. We have offset spatial vector results slightly on the q^2-axis for clarity. Right: the ratio of the f_+ values for the spatial and temporal vector cases. We see no evidence of any differences between them (within our uncertainties) that would indicate discretisation effects. 138

7.3 Comparison of raw correlation function data from binned and unbinned data on set 8. In each case, data is normalised by e^{-Mt}, offset from integer t/a values and has errors inflated 500 times for clarity. Clockwise from top left, we have plots for $H(am_h = 0.194)$, $\hat{H}(am_h = 0.8)$, $K(\theta = 0)$, $\hat{K}(\theta = 4.705)$. 140

7.4 f_0,f_+ and f_T results in z space for $B \to K$ on set 8. In red are the results we use, from the unbinned ('ub') data. For comparison in black are the results from the binned ('b') data. 141

7.5 Correlations between matrix elements for a simultaneous fit (left) and one which is split up according to heavy mass (right) on set 8. Correlations between different twists at fixed mass are shown in blue, and different masses at fixed twist in red. The bars are additive (i.e. blue sits above red, not behind it). See text for more details. 143

7.6 Stability tests for the z expansion fit evaluated at the physical B mass. Test 0 is the final result, 1 and 2 take different priors for ζ_0, test 3 allows ζ_0 to vary between the form factors and test 4 drops the term containing ζ entirely. Test 5 increases N by 1 (to 4) and test 6 does the same for each component of N_{ijkl}. Test 7 doubles the width of ζ_n and all d and ρ priors, and 8 halves them. Test 9 removes the chiral logarithm term by setting $\mathcal{L} = 1$, and 10 tightens the prior on the ρ coefficients considerably (from 0(1)). Test 11 removes the $f_0(0) = f_+(0)$ constraint; here the black point is $f_0(0)$ and the red is $f_+(0)$. Tests 12, 13 and 14 remove all the lattices with physical light masses, all of set 8 data, and results with $m_l = m_s$ respectively. Test 15 removes the spatial vector data, and 16 removes the largest mass from all ensembles with multiple masses. The χ^2 per degree of freedom and log(GBF) value for each test are shown in the bottom pane in blue and red respectively. For the latter tests, data is removed from the fit, resulting in a log(GBF) which is not comparable with others and so not displayed. As discussed, χ^2 values are only meaningful relatively. χ^2 values for tests 7 and 8, which change widths on many priors, are not directly comparable with other tests. 153

7.7 $\left(1 - \frac{q^2}{M^2_{H^*_{s0}}}\right) f_0(z)$ data points and final result at the physical point (blue band). Data points are labelled by heavy quark mass, where e.g. m0.8 indicates $am_h = 0.8$ on that ensemble. Lines between data points of a given heavy mass are the result of the fit evaluated on this ensemble and mass with all lattice artefacts present. Sets 9 and 10 are the $H_s \to \eta_s$ data from sets 1 and 2 in [2], which were fitted simultaneously with sets 6 and 7 respectively. 154

7.8 $\left(1-\frac{q^2}{M_{H_s^*}^2}\right) f_+(z)$ data points and final result at the physical point (red band). Data points are labelled by heavy quark mass, where e.g. m0.8 indicates $am_h = 0.8$ on that ensemble. Lines between data points of a given heavy mass are the result of the fit evaluated on this ensemble and mass with all lattice artefacts present. Sets 9 and 10 are the $H_s \to \eta_s$ data from sets 1 and 2 in [2], which were fitted simultaneously with sets 6 and 7 respectively. At large $|z|$ (large q^2), data obtained from both temporal and spatial components of V^μ are shown, the latter with end caps specifying the associated uncertainty. As discussed in Section 7.2.3, errors for f_+ at large q^2 are significantly smaller when obtained from spatial vector components. 155

7.9 $\left(1-\frac{q^2}{M_{H_s^*}^2}\right) f_T(z)$ data points and final result at the physical point (green band). Data points are labelled by heavy quark mass, where e.g. m0.8 indicates $am_h = 0.8$ on that ensemble. Lines between data points of a given heavy mass are the result of the fit evaluated on this ensemble and mass with all lattice artefacts present. . 156

7.10 The contributions to the total percentage error (black line) of $B \to K$ form factors $f_0(q^2)$ (top) and $f_+(q^2)$ (middle) and $f_T(q^2)$ (bottom) from different sources, shown as an accumulating error. The red dashed line ('inputs') includes values for parameters, such as masses, taken from the PDG [29] and used in the fit as described above. The purple dotted line ('q mistunings') adds, negligibly, to the inputs the error contribution from the quark mistunings associated with c fit parameters and errors from the light quark chiral extrapolation, whilst the solid blue line ('statistics') further adds the error from our correlator fits. The green dot-dash line ('HQET') includes the contribution from the expansion in the heavy quark mass, and, finally, the thick black line ('Discretisation'), the total error on the form factor, also includes the discretisation errors. In the case of the tensor form factor, the difference here is so small as to obscure the HQET line. The percentage variance adds linearly and the scale for this is given on the left hand axis. The percentage standard deviation, the square root of this, can be read from the scale on the right-hand side. 158

7.11 Final $B \to K$ form factor results at the physical point across the full q^2 range. . 159

7.12 The form factors at q^2_{max} and $q^2 = 0$ evaluated across the range of physical heavy masses from the D to the B. Other lattice studies [140, 173, 182, 212] of both $D \to K$ and $B \to K$ are shown for comparison. We also include some $B \to K$ results at $q^2 = 0$ from Gubernari et al. [213], a calculation using light cone sum rules. We do not include Chapter 5's $D \to K$ results that share data with our calculation here. At the B end, data points are offset from M_B for clarity. Note that we have run Z_T to scale μ in this plot, where μ is defined linearly between 2 GeV and $m_b = 4.8$ GeV, according to Equation (7.9). 160

7.13 Breakdown of the contributions to the statistical uncertainty of the form factors at their extremes from data on each ensemble. Uncertainty from each ensemble σ_i is added in quadrature, normalised by the total uncertainty squared $\sum_i \sigma_i^2$. Sets 6 and 7 include contributions from $H_s \to \eta_s$ data. Left: evaluated at $M_H = M_B$ ($B \to K$). Right: evaluated at $M_H = M_D$ ($D \to K$). 162

7.14 The green band gives our $D \to K$ tensor form factor at $\mu = 2$ GeV, across the physical q^2 range. Results from [212] are included for comparison. 163

7.15 Comparison of our $B \to K$ scalar and vector form factors with those of $B_s \to \eta_s$ [2] and $B_c \to D_s$ [214] to show the impact of changing the spectator quark mass. In the right pane, we have multiplied the form factors by their common pole factors for clarity. We take $P^0(q^2) = 1 - \frac{q^2}{M^2_{B^*_{s0}}}$, $P^+(q^2) = 1 - \frac{q^2}{M^2_{B^*_s}}$, using PDG masses [29]. 164

7.16 Left: comparison of our $B \to K$ tensor form factor (at $\mu = 4.8$ GeV) with those of $B_c \to D_s$ [214] to show the impact of changing the spectator quark mass. Right: a similar comparison of our $D \to K$ scalar and vector form factors with those for $D_s \to \eta_s$ [2] and $B_c \to B_s$ [98]. 165

7.17 The real and imaginary parts of δC_7^{eff} and of the $\mathcal{O}(\alpha_s)$ and $\mathcal{O}(\lambda_u^{(s)})$ contributions to δC_9^{eff}, as defined in Equations (7.21) and (7.20). 171

7.18 The real (left) and imaginary (right) parts of the contributions to $f_+ \Delta C_9^{\text{eff}}$ from each of the three non-zero terms, $T^{(0)}_{K,-}$ and $T^{(\text{nf})}_{K,\pm}$ (Equation (7.22)). A decay channel-specific factor of $e_q \in \{2/3, -1/3\}$ is removed from $T^{(0)}_{K,-}$ and $T^{(\text{nf})}_{K,-}$. . . . 172

7.19 Left: the real and imaginary parts of $f_+\Delta C_9^{\text{eff}}$ for both the B^0 and B^+ cases. Right: the real and imaginary parts of $C_7^{\text{eff},0}$ and $C_7^{\text{eff},1}$ (see Equation (7.18)), showing the combined effect of the nonfactorisable and $\mathcal{O}(\alpha_s)$ corrections. Corrected values are shown with solid lines and dark fill colour, while uncorrected values are shown with dotted lines and light fill colour. 173

7.20 The real and imaginary parts of $C_9^{\text{eff},0}$ and $C_9^{\text{eff},1}$ (see Equation (7.19)), for $B^0 \to K^0\ell^+\ell^-$ (left) and $B^+ \to K^+\ell^+\ell^-$ (right), showing the combined effect of the nonfactorisable and $\mathcal{O}(\alpha_s)$ corrections. Corrected values are shown with solid lines and dark fill colour, while uncorrected values are shown with dotted lines and light fill colour. $C_9 = C_9^{\text{eff},0} - Y(q^2) = 4.114(14)$ [40] is included for comparison. 173

7.21 Left: differential branching fraction for $B^+ \to K^+\ell^+\ell^-$, with our result in blue, compared with experimental results [9, 193, 194, 200, 202, 204]. Note that Belle '19, and LHCb '14C and '21 have $\ell = e$, whilst otherwise $\ell = \mu$. Horizontal error bars indicate bin widths. Right: the total branching fraction for $B^0 \to K^0\ell^+\ell^-$. Our result (HPQCD '22) is given by the black band, as compared with experimental results [193, 194, 196, 199, 200, 201]. Dashed lines indicate the effect of adding QED uncertainty (see [4]) to our result. 174

7.22 Comparison of branching fractions with recent experimental results [9, 194] in well behaved regions of q^2. Here we show the ratio of the experimental branching fraction to our results, meaning our results correspond to the line at 1. The error bars are 5σ long, with markers at 1, 3 and 5 σ. Note that the σ here are different from those calculated in Table 7.11. On the right, labels indicate q^2 bins in units of GeV^2. 176

7.23 Left: the short distance contribution to $d\mathcal{B}(B \to K\nu\bar{\nu})/dq^2$, for both the charged and neutral cases. Right: our result for $B^+ \to K^+\nu\bar{\nu}$, including long distance effects, as compared with current 90% confidence limits from BaBar [247] and Belle II [248]. Theoretical uncertainties achieved by Belle II with 5 ab^{-1} and 50 ab^{-1} [241], assuming our central value, are also included. 177

List of Tables

2.1 The fields of the Standard Model and their $SU(3)_c \times SU(2)_L \times U(1)_Y$ representations. 13

2.2 Wilson coefficients used in Chapter 7, as well as their corresponding operators [39]. Central values and uncertainties are quoted at $\mu_b = 4.2\,\text{GeV}$ and taken from [40]. Note that the operators given for $C_7^{\text{eff},0}$ and C_8^{eff} are those for C_7 and C_8, with the effective coefficient containing contributions from other operators (see text). The sums over p run over all quarks except the top. 19

3.1 Gluon ensembles used in this work. In column 3, β is the same one defined in Section 3.2.2 and is discussed further in the text. The Wilson flow parameter, $w_0 = 0.1715(9)\text{fm}$ (also discussed in the text), is determined in [75], and is used to calculate the lattice spacing a via values for w_0/a [72] in column 4. Column 5 gives the spatial (N_x) and temporal (N_t) dimensions of each lattice in lattice units whilst columns 6-10 give the masses of the valence and sea quarks, noting that these are the same in the case of the light quark. 40

4.1 δ_{FV} values used on each ensemble. 73

5.1 Details of the statistics, T values and K meson momenta used on each ensemble. The number of configurations n_{cfg} and time sources n_{src} are given, and momenta can be obtained from twist, θ, via $\theta = |a\vec{p}_K|N_x/(\sqrt{3}\pi)$, where N_x is the spatial dimension of the lattice, given in Table 3.1. 81

5.2 Priors used in the fit on each set. $A_{i\neq 0}^{D,K}$ indicates the amplitudes for non-oscillating and oscillating D mesons and for non-oscillating K mesons. $A_i^{K,\mathrm{o}}$ is the amplitude for oscillating K, which we expect to be smaller because the oscillation vanishes at zero momentum when the quark masses are the same. Parameters denoted S and V refer to the J_{ij}^{kl} parameters (see Equation (4.40)) for the scalar and temporal vector currents respectively. Columns 4 and 5 then give the priors for the ground-state to ground-state parameter cases where at least one of the states is an oscillating state. For the cases where at least one state is an excited state, $\mathcal{P}[S_{ij\neq 00}^{kl}] = \mathcal{P}[V_{ij\neq 00}^{kl}] = 0.0(5)$ in all cases. 82

5.3 Priors and their posterior values for the z expansion fit. 87

5.4 Values and uncertainties for the fit coefficients $a_n^{0,+}$, pole masses, and chiral logarithmic term $(1 + L(m_l))$ for the reconstruction of our form factors at the physical point as a function of q^2 from Equations (5.8) and (5.9). The correlation matrix between these parameters is given below the row with their values. The pole masses are in GeV. The pole masses and $L(m_l)$ are very slightly correlated due to the way the fit function is constructed. These correlations are too small to have any meaningful effect on the fit, but we include them for completeness in reconstructing our results. 93

5.5 Total width for $D \to K$ semileptonic decay up to a factor of $|\eta_{EW} V_{cs}|^2(1 + \delta_{EM})$, determined from our form factors. We give results for all 4 modes that we consider. They differ slightly in the parent and daughter meson masses and in the mass of the lepton in the final state; these affect the kinematic factors in the differential rate and the end-points of integration for the total width. These values can be combined with experimental values of the relevant branching fraction and D meson lifetime to determine $|V_{cs}|$. 105

6.1 Details of the statistics, T values, heavy masses, and η_s meson twists used on each ensemble. The number of configurations n_{cfg} and time sources n_{src} are given, and momenta can be obtained from twist, θ, via $\theta = |a\vec{p}_{\eta_s}|N_x/(\sqrt{3}\pi)$, where N_x is the spatial dimension of the lattice, given in Table 3.1. In each case, the smallest am_h value corresponds to the tuned am_c mass on that ensemble. 114

6.2 Priors used in the fit on each set. Priors are based on previous experience of similar fits [147] (see text) and given conservative widths. In some places, adjustment is made for lattice spacing, and priors are tuned using an increase in the GBF. The effect of doubling and halving the standard deviation on all priors on the final fit result is shown in Figure 6.1. 116

6.3 Priors and their posterior values for the z expansion fit. 121

6.4 Priors and their posterior values for the z expansion fit. 122

6.5 Priors and their posterior values for the z expansion fit. 123

6.6 Values of fit coefficients $a_n^{0,+}$ and pole masses at the physical point for the $B_s \to \eta_s$ decay with correlation matrix are given below. Form factors can be reconstructed by evaluating Equations (6.1) and (6.2) using these coefficients and pole masses. Note that $M_{B_{s0}}$ is set to $M_{B_s} + 0.4\,\text{GeV}$. Masses are in GeV. The pole masses are very slightly correlated due to the way the fit function is constructed. These correlations are too small to have any meaningful effect on the fit, but we include them for completeness in reconstructing our results. 129

7.1 Details of the statistics, T values, heavy masses, and K meson twists used on each ensemble. The number of configurations n_{cfg} and time sources n_{src} are given, and momenta can be obtained from twist, θ, via $\theta = |a\vec{p}_K|N_x/(\sqrt{3}\pi)$, where N_x is the spatial dimension of the lattice, given in Table 3.1. In each case, the smallest am_h value corresponds to the tuned am_c mass on that ensemble, i.e. the $D \to K$ data in Chapter 5. 136

7.2 Priors and their posterior values for the z expansion fit. 148

7.3 Priors and their posterior values for the z expansion fit. 149

7.4 Priors and their posterior values for the z expansion fit. 150

7.5 Priors and their posterior values for the z expansion fit. 151

7.6 Priors and their posterior values for the z expansion fit. 152

7.7 Values of fit coefficients $a_n^{0,+,T}$, pole masses, and the $\mathcal{L}$ term with correlation matrix below, evaluated at the physical point and the B mass. Note that $a_0^+ = a_0^0$. Masses are in GeV. The pole masses and $\mathcal{L}$ are very slightly correlated due to the way the fit function is constructed. These correlations are too small to have any meaningful effect on the fit, but we include them for completeness. 157

7.8 Form factor results at the q^2 extremes. As described in the text, the $f_0^{D\to K}$ and $f_+^{D\to K}$ share data with the results in [1] (included for comparison) so should not be viewed as an independent calculation. 161

7.9 $B_s \to \eta_s$ form factor results at the q^2 extremes. As described in the text, the f_0 and f_+ share data with the results in [2] (included for comparison) so should not be viewed as an independent calculation. 161

7.10 Input parameters used to calculate SM observables. Details of the error on the Wilson coefficients are given in [40], and these are quoted at $\mu_b = 4.2\,\text{GeV}$. For more details on Wilson coefficients, see Section 2.3.2. $m_c^{\overline{\text{MS}}}(m_c^{\overline{\text{MS}}})$ is obtained using $m_c^{\overline{\text{MS}}}(3\text{ GeV}) = 0.9841(51)$ [219] and running the scale to its own mass. m_c and m_b are the c and b quark pole masses obtained from the masses in the $\overline{\text{MS}}$ scheme at three loops (see Section 7.4.1 for details). 168

7.11 Comparison of branching fractions with recent experimental results [9, 194] in well behaved regions of q^2. Note that the $B^+ \to K^+e^+e^-$ result quoted here from LHCb '21 is obtained using the $B^+ \to K^+\mu^+\mu^-$ result from LHCb '14A, combined with the ratio determined in LHCb '21. In the fifth column, the tension with our result is given by mean(Experiment - HPQCD)/sdev(Experiment - HPQCD). . . 175

7.12 Short distance contributions to branching fractions $\mathcal{B}^{(0/+)} = \mathcal{B}(B^{0/+} \to K^{0/+}\nu\bar{\nu})_{\text{SD}}$ for the rare decay $B \to K\nu\bar{\nu}$ (Equation (7.23)) integrated over various q^2 bins. Numerical values for q^2 bins are in units of GeV^2. 177

7.13 Branching fractions integrated over some commonly used q^2 binning schemes for the electron, muon and ℓ, which is the average of the two. In the first bin of the top panel, giving the branching fraction integrated over the full q^2 range, the first number is the whole integral, whilst the one which follows in square brackets is the result when the ranges 8.68-10.11 GeV^2 and 12.86-14.18 GeV^2 are excluded. Numerical values for q^2 bins are in units of GeV^2. QED corrections are not included here. 178

7.14 The ratio for $B \to K\mu^+\mu^-$ and $B \to Ke^+e^-$ branching fractions, R_e^μ, integrated over some commonly used q^2 binning schemes. We give results for the charged meson case, the neutral meson case and the charge-averaged case (defined as the ratio of the two charge-averaged integrals). Numerical values for q^2 bins are in units of GeV^2. A 1% uncertainty from QED effects [239] is not included in these numbers. 179

7.15 Branching fractions for decays to τ leptons, integrated over some commonly used q^2 bins. Numerical values for q^2 bins are in units of GeV^2. No uncertainty to allow for QED is included in these numbers. 179

7.16 The ratio of branching fractions, R for the τ case to the electron, muon and ℓ, (which is the average of the two) cases, integrated over some commonly used q^2 bins. Numerical values for q^2 bins are in units of GeV^2. No uncertainty to allow for QED effects is included in these numbers. 179

Chapter 1

Introduction

The Standard Model (SM) of particle physics is often lauded for its remarkable accuracy in its description of nearly everything we observe in the universe. If one turns a blind eye to the apparent incompatibility with that other great triumph, General Relativity, it really is rather deluxe. Experiments have confirmed many of its predictions to incredible levels of accuracy, and discovered the charm and bottom quarks, as well as the Higgs boson. The number of constraints which can be placed on the unitarity of the CKM matrix (Figure 2.1) is a testament to this success. And yet all is not as ship shape and Bristol fashion as it might appear. The SM provides no explanation of dark matter, or dark energy, for a start. And deep down, everyone knows that it is nothing more than the precocious little brother of some much more handsome high energy theory, the real McCoy.

After the success of the Higgs Boson discovery, the LHC has plodded along, slowly reducing the available phase space for new physics, pushing the energies of squarks and all manner of other hypothetical particles higher and higher - it seems unlikely that there will be any golden bullet shaped direct detection any time soon. So we find ourselves working in the so called intensity frontier, striving for ever greater precision in theoretical calculations and experimental measurements, trying to find the smallest crack in the SM, into which we might drive a wedge. Fortunately for us, there exist some places where cracks are beginning to appear, notably in the muon $g-2$ [5], and in the flavour sector, which shall be the focus of this thesis.

Flavour anomalies [6, 7], which point to new physics, currently present themselves most openly in the decays of B mesons [8]. In particular, the apparent violation of Lepton Flavour Uni-

versality (LFU) in rare $B \to K\ell^+\ell^-$ decays has now reached a significance of 3.1σ [9]. Better experimental precision is promised by the high luminosity LHC upgrade, but we also require improved theoretical inputs, often in the form of hadronic form factors from lattice QCD, in order to convincingly demonstrate such anomalies.

In this thesis, we will make use of the Highly Improved Staggered Quark (HISQ) action [10] to enable precise determination of hadronic form factors from lattice QCD. We will develop further the heavy-HISQ technique for $b \to s$ decays, first establishing its efficacy with a calculation of the unphysical $B_s \to \eta_s$ decay, and then building on this with a full calculation of the scalar, vector and tensor form factors of the semileptonic $B \to K$ decay. This first fully relativistic calculation will improve the theoretical perspective on the flavour anomalies observed in this decay.

Another place to hunt for SM deficiencies is in the unitarity of the CKM matrix. Checking this requires independent determinations of each matrix element. From an improved calculation of the $D \to K\ell\nu$ form factors, we will provide an improvement on the semileptonic determination of the matrix element V_{cs}, obtaining a value of $|V_{cs}| = 0.9663(80)$, the first time this quantity has been shown to be significantly less than unity, and with a sub percent uncertainty.

Chapter 2

The Standard Model of particle physics

In the following chapter we will outline the Standard Model (SM) of particle physics, focusing particularly on areas which will be relevant to the rest of the thesis. We do not intend to provide a first principles introduction to quantum field theory, for which much literature already exists e.g. [11, 12]. Below we will broadly follow the treatments of Mark Srednicki in his book *Quantum Field Theory* [12], as well as the lecture notes of David Tong [13] and those of Ben Gripaios [14].

2.1 The basis of the Standard Model

As is well documented, the Standard Model is best described as a Lorentz invariant gauge quantum field theory, in which fermionic matter fields interact via the exchange of vector bosons. We begin with a description of spinors, and gauge field theories, which will give us the tools to develop the SM further. We then progress to a more detailed discussion of the weak interactions which form the backbone of this thesis.

The basic features of the SM are two types of matter (fermions) and three forces (carried by bosons). The six flavours of quark, arranged into three generations, interact via the strong force (Quantum Chromodynamics), weak force and electromagnetism (Quantum Electrodynamics), which are carried by the gluons, $\{W^+, W^-, Z\}$ and γ bosons respectively. The remaining matter

takes the form of three generations of leptons. Of these, the three charged leptons, the e, μ and τ interact via electromagnetism and the weak force, whilst their three neutrino siblings are only weakly interacting.

2.1.1 Spin-half fields

Consider a field which transforms under the Lorentz group such that $\phi^a(x) \to D[\Lambda]^a{}_b \phi^b(\Lambda^{-1}x)$ for some representation $D[\Lambda]$ of the Lorentz group, which must satisfy $D[\Lambda_1]D[\Lambda_2] = D[\Lambda_1\Lambda_2]$. We can write an infinitesimal Lorentz transformation [15] as

$$\Lambda^\mu{}_\nu = e^{\frac{1}{2}\Omega_{\rho\sigma}(\mathcal{M}^{\rho\sigma})^\mu{}_\nu}, \tag{2.1}$$

where $(\mathcal{M}^{\rho\sigma})^\mu{}_\nu = \eta^{\rho\mu}\delta^\sigma{}_\nu - \eta^{\sigma\mu}\delta^\rho{}_\nu$ (with Minkowski metric $\eta^{\mu\nu}$) are (with anti-symmetric labelling $\rho\sigma$) the six generators of rotations and boosts in the Lorentz group, and (anti-symmetric) $\Omega_{\rho\sigma}$ are the six real numbers controlling the magnitude of each boost or rotation.

The generators $\mathcal{M}$ satisfy the Lie algebra of the Lorentz group,

$$[\mathcal{M}^{\rho\sigma}, \mathcal{M}^{\tau\nu}] = \eta^{\sigma\tau}\mathcal{M}^{\rho\nu} - \eta^{\rho\tau}\mathcal{M}^{\sigma\nu} + \eta^{\rho\nu}\mathcal{M}^{\sigma\tau} - \eta^{\sigma\nu}\mathcal{M}^{\rho\tau}. \tag{2.2}$$

Consider the four 4×4 Dirac matrices γ^μ, which satisfy the Clifford Algebra $\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu} I_4$. The commutator $S^{\mu\nu} = \frac{1}{4}[\gamma^\mu, \gamma^\nu]$ satisfies the Lie algebra of Equation (2.2) and thus forms a representation of the Lorentz algebra.

Putting all of this together, we introduce the four component Dirac spinor [16] field $\psi^\alpha(x)$, which the α, β components of $(S^{\mu\nu})^\alpha{}_\beta$ act upon. Under infinitesimal Lorentz transformations,

$$\psi^\alpha(x^\mu) \to S[\Lambda]^\alpha{}_\beta \psi^\beta((\Lambda^{-1})^\mu{}_\nu x^\nu) = e^{\frac{1}{2}\Omega_{\rho\sigma}(S^{\rho\sigma})^\alpha{}_\beta}\psi^\beta(e^{-\frac{1}{2}\Omega_{\tau\epsilon}(M^{\tau\epsilon})^\mu{}_\nu}x^\nu). \tag{2.3}$$

What's important here is that spacetime transforms under the representation of the Lorentz group ($SO(3,1)$), generated by $\mathcal{M}^{\mu\nu}$, whilst the components of our spin-half fields transform under a different, spinor representation, generated by $S^{\mu\nu}$. Demanding that the numbers $\Omega_{\mu\nu}$ are the same across both transformations ensures that we are doing the same Lorentz transformation.

In the Standard Model, spin half fermions (comprising the quarks and the leptons) will be described by four component spinors, transforming as in Equation 2.3.

2.1.2 Gauge field theories

The main source of the rich soup of phenomenology in which our universe slowly poaches is the presence of non-Abelian, that is to say non-commutative, gauge theories in the Standard Model.

Consider the set of $N \times N$ unitary matrices $U(N)$, which can be written as exponentiated $N \times N$ Hermitian matrices; e^{iH}. Expanding H in a basis of Hermitian matrices $\{T^a\}$ with real weighting ϵ^a, we find that in the neighbourhood of the identity, i.e. small ϵ^a, $e^{i\epsilon^a T^a} = 1 + i\epsilon^a T^a + ...$, where summation over a is implied. Adding a different small ϵ', we can write,

$$\begin{aligned} e^{i\epsilon^a T^a} e^{i\epsilon'^b T^b} e^{-i\epsilon^c T^c} e^{-i\epsilon'^d T^d} &= 1 - \epsilon^c \epsilon'^d T^c T^d + \epsilon^c \epsilon'^b T^b T^c + \mathcal{O}((\epsilon + \epsilon')^3) \\ &= 1 - \epsilon^c \epsilon'^d [T^c, T^d] + \mathcal{O}((\epsilon + \epsilon')^3), \end{aligned} \tag{2.4}$$

where we have expanded to second order in ϵ and ϵ', and relabelled dummy indices in the second line. This product of group elements must also lie in the group, and so can be expanded in terms of the basis $\{T^a\}$. The result is that

$$[T^a, T^b] = i f^{abc} T^c, \tag{2.5}$$

where f^{abc} are real constants (structure functions), which are manifestly anti-symmetric in the first two indices, but may be taken to be anti-symmetric in all three without loss of generality. This is the Lie algebra of the unitary matrices, and applies equally well to the subgroup $SU(N)$ in which we shall be interested. These special unitary matrices, which satisfy $U^\dagger U = \delta$ and have unit determinant, will be key to the group structure of the SM. The aim of the game is to find sets of matrices, or representations, which satisfy the Lie algebra of Equation (2.5)[1]. Matter in the SM then transforms in some representation of the gauge group.

Quantum Electrodynamics

As a brief exploration of a commuting, or Abelian, gauge theory, we give the familiar example of Quantum Electrodynamics, QED. Here we have a spin-half Dirac field ψ, the electron, and a

[1] It's worth mentioning that the Lorentz group, SO(3,1), cannot be represented by finite dimensional unitary matrices.

vector field A, the photon, which posses a local (gauge) symmetry,

$$\begin{aligned} \psi &\to e^{ie\alpha(x)}\psi \\ A^\mu &\to A^\mu - \partial^\mu \alpha. \end{aligned} \tag{2.6}$$

Allowing terms which are at most cubic in the fields, this symmetry, together with Lorentz invariance, fixes the Lagrangian,

$$\mathcal{L}_{\text{QED}} = \bar{\psi}(i\not{D} - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}, \tag{2.7}$$

where $\bar{\psi} \equiv \psi^\dagger\gamma^0$, $\not{X} \equiv \gamma_\mu X^\mu$ and the covariant derivative $D_\mu = \partial_\mu + ieA_\mu$. The field strength tensor is given by $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. Putting on our big and shiny group theory hat, we can say that in fact $U = e^{ie\alpha}$ is a 1×1 unitary matrix, and as such, QED is a $U(1)$ gauge theory.

Non-Abelian gauge theory

Let's go back to first principles and build a non-Abelian gauge theory, where matter (say fermions) transforms in a representation r of gauge group G. We shall see that we can recover QED as above, along with much more.

Under a global G transformation, our fermions transform as,

$$\psi \to U\psi \equiv e^{ig\alpha^a T_r^a}\psi, \tag{2.8}$$

where we insert constant g, (which will become the coupling) for convenience. Here T_r^a is an $n_r \times n_r$ matrix, and it acts on the length n_r vector, ψ. Being a fermion, ψ also carries a spinor index, which we suppress here. We want to promote G to a local (gauge) symmetry, such that $\alpha^a \to \alpha^a(x)$ which means that we require a derivative which transforms covariantly. To do this we follow the principle of minimal coupling and define the covariant derivative,

$$D_\mu = \partial_\mu + igA_\mu, \tag{2.9}$$

where we have introduced the vector gauge field A which transforms as $A \to A'$. We require that

$D_\mu\psi \to UD_\mu\psi$, (such that $\bar{\psi}\not{D}\psi \to \bar{\psi}U^{-1}U\not{D}\psi = \bar{\psi}\not{D}\psi$) which gives us

$$\begin{aligned} \partial_\mu + igA'_\mu &= U(\partial_\mu + igA_\mu)U^{-1} \\ \implies A'_\mu &= UA_\mu U^{-1} - \frac{i}{g}U\partial_\mu U^{-1} \\ &= UA_\mu U^{-1} + \frac{i}{g}(\partial_\mu U)U^{-1}, \end{aligned} \tag{2.10}$$

where we note that $\partial_\mu U^{-1} = U^{-1}\partial_\mu + (\partial_\mu U^{-1})$ because we are dealing with operators. We can write this in a way which does not involve the representation by noting that each A_μ must be an $n_r \times n_r$ matrix, and expanding it in terms of the representation $A_\mu \equiv A^a_\mu T^a_r$.

$$\begin{aligned} A'^a_\mu T^a_r &= A^c_\mu T^c_r + ig(\alpha^b T^b_r A^c_\mu T^c_r - A^c_\mu T^c_r \alpha^d T^d_r) - \partial_\mu \alpha^e T^e_r \\ &= A^a_\mu T^a_r + ig\alpha^b A^c_\mu [T^b_r, T^c_r] - \partial_\mu \alpha^a T^a_r \\ \implies A'^a_\mu &= A^a_\mu - gf^{bca}\alpha^b A^a_\mu - \partial_\mu \alpha^a. \end{aligned} \tag{2.11}$$

In analogy with QED, to find a dynamical term for the gauge field in the action, we must look for a gauge invariant combination. From the covariance property $D_\mu \to UD_\mu U^{-1}$, we find $igF^a_{\mu\nu}T^a_r \equiv [D_\mu, D_\nu] \to U[D_\mu, D_\nu]U^{-1}$, where we have defined the field strength tensor $F_{\mu\nu} = F^a_{\mu\nu}T^a_r$. Expanding out the commutator gives

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu - gf^{abc}A^b_\mu A^c_\nu. \tag{2.12}$$

This is not gauge invariant, but $\frac{1}{2g^2}tr[D_\mu, D_\nu][D^\mu, D^\nu] = -\frac{1}{4}F^a_{\mu\nu}F^{a\mu\nu}$ is, giving us the desired term in the Lagrangian.

For the case of Abelian gauge theories, such as QED, $f^{abc} = 0$ and we recover the familiar expressions for the transformation of the gauge field and the field strength tensor. The richness of non-Abelian gauge theories lies in the extra term in Equation (2.12), which gives rise to terms in the Lagrangian which are cubic and quartic in the gauge field, and thus self interactions.

2.1.3 Quantum Chromodynamics

Quantum Chromodynamics (QCD) [17, 18] is a non-Abelian $SU(3)$ gauge theory of the form discussed in Section 2.1.2. There are $N^2 - 1 = 8$ gauge bosons, the gluons, which couple to

fermions, the quarks. Quarks transform in the defining[2] 3-dimensional representation of $SU(3)$, the three indices (or colours) of which can be called red, green and blue (r, g and b). All of the 6 flavours of quark transform as colour triplets.

The QCD Lagrangian is given by

$$\mathcal{L}_{\text{QCD}} = -\frac{1}{4}G^a_{\mu\nu}G^{a\mu\nu} + \sum_{f\in\{u,d,s,c,t,b\}} \bar{\psi}\left(i\not{\partial} - g_s\not{A}^a\frac{\lambda^a}{2} - m_f\right)\psi. \tag{2.13}$$

As is often done, we denote the field strength (generically $F^a_{\mu\nu}$) as $G^a_{\mu\nu}$ here, denoting gluons. g_s is the strong coupling constant, and λ^a are the eight 3×3 Gell-Mann matrices, which provide an explicit basis for the defining triplet representation.

In Nature, all hadrons are colourless, which means they must comprise a colour singlet. Taking the example of mesons, we have a quark and antiquark triplet, labelled by quantum numbers Y^c, the colour hypercharge and I^c_3, the colour isospin. Combined, these give rise to a coloured octet and a colourless singlet $3\otimes\bar{3} = 8\oplus 1$. The singlet has colour combination $\frac{1}{\sqrt{3}}(r\bar{r}+g\bar{g}+b\bar{b})$. The gluons, themselves carrying colour charge, obey the same mathematics, but they comprise the octet $r\bar{g}$, $r\bar{b}$, $g\bar{r}$, $g\bar{b}$, $b\bar{r}$, $b\bar{g}$, $\frac{1}{\sqrt{2}}(r\bar{r}-g\bar{g})$ and $\frac{1}{\sqrt{6}}(r\bar{r}+g\bar{g}-2b\bar{b})$. There are only 8 gluons because of the $SU(3)$ structure of QCD. Had Nature chosen a $U(3)$ structure, the 9th gluon would be a colour singlet, and QCD would be an unconfined, long range force.

2.1.4 The electroweak sector

The electroweak sector [19, 20] is the basis for the semileptonic decays studied in this thesis, so we will introduce it in slightly more detail than QED or QCD. It is also the most nuanced of the gauge theories we shall discuss, and so it is instructive to piece it together more gradually, as it was uncovered historically.

Let's play the same game as with QCD, and try to find a gauge theory that describes the weak interaction. As we suspect the weak interaction is responsible for the quark flavour changing observed in β decay, we want a representation with at least two elements, which couples to quarks and leptons, so we choose the fundamental representation of the simplest non-Abelian Lie group: $SU(2)$. One set of possible generators for this representation are the Pauli matrices, $\sigma^i/2$, which

[2] In $SU(N)$ we have a defining representation carried by $\mathbb{C}^N$ vectors, on which $N\times N$ matrices act by multiplication.

satisfy the Lie algebra $[\frac{\sigma^i}{2}, \frac{\sigma^j}{2}] = i\epsilon_{ijk}\frac{\sigma^k}{2}$. For a 3 component $SU(2)$ gauge field W^i_μ, the covariant derivative is given by,

$$D_\mu = \partial_\mu + i\frac{g}{2}W^i_\mu\sigma^i = \partial_\mu + i\frac{g}{2}\begin{pmatrix} W^3_\mu & \sqrt{2}W^+_\mu \\ \sqrt{2}W^-_\mu & -W^3_\mu \end{pmatrix}, \tag{2.14}$$

where we have defined $W^\pm_\mu \equiv \frac{1}{\sqrt{2}}(W^1_\mu \mp iW^2_\mu)$. We package the quarks and leptons into doublets $l \equiv (\nu, e)^T$ and $q \equiv (u, d)^T$ (one for each generation), that the derivative acts upon. Now our weak Lagrangian can be constructed, containing terms like

$$\mathcal{L} \supset \bar{l}i\not{D}l + \bar{q}i\not{D}q. \tag{2.15}$$

Except that weak interactions have been shown experimentally [21] not to conserve parity[3], which is not permitted in this formalism. A simple fix is to introduce the projection operators $P_{L,R} \equiv \frac{1}{2}(1 \mp \gamma^5)$, for $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$. The effect of this is to project Dirac fermions into right (R) and left (L) handed components $\psi_{L,R} \equiv P_{L,R}\psi$, which can be separated in the Dirac Lagrangian (noting that $\overline{\psi_L} \equiv (P_L\psi)^\dagger\gamma^0 = \overline{\psi}P_R \equiv \overline{\psi}_R$).

$$\mathcal{L} = i(\overline{\psi_L}\not{\partial}\psi_L + \overline{\psi_R}\not{\partial}\psi_R) - m(\overline{\psi_L}\psi_R + \overline{\psi_R}\psi_L). \tag{2.16}$$

In such a theory, containing left and right handed components, we can assign the different components to different representations of the gauge group, but the mass term will no longer be gauge invariant.

Let's dictate that only the left handed parts of quarks and leptons couple to W_μ via $SU(2)$, creating an issue for these fermions having mass, which we shall have to resolve later. Now Equation (2.15) contains terms like $\frac{g}{\sqrt{2}}\overline{\nu_L}\not{W}^+e_L$ which we can identify with the $W^\pm$ boson involved in beta decay, for example. What about W^3_μ? This appears in terms like $i\frac{g}{2}\not{W}^3(\overline{\nu_L}\nu_L - \overline{e_L}e_L)$, which looks like the behaviour of a Z boson; a neutral boson which couples to neutrinos as well as charged leptons. Except we observe that the Z couples to both right and left handed fermions.

The solution to this next conundrum is to involve the other neutral boson we have discussed, the photon, which also couples to left and right handed fermions. We are about to unify the

[3]The parity operator reflects space: $(t, x, y, z) \to (t, -x, -y, -z)$.

weak force with electromagnetism. To see how this unfolds, we introduce a new $U(1)$ boson, B_μ. As before, we have our left handed $SU(2)$ doublets, q_L and l_L, with coupling constant g. These doublets now have a weak hypercharge $Y_{q,l}$ under the $U(1)$ gauge transformation gauged by B_μ, with coupling constant g'. Now the right handed fermions, u_R, d_R and e_R[4] are $SU(2)$ singlets, with weak hypercharge $Y_{u,d,e}$. Finally, we write the physical boson eigenstates A_μ and Z_μ as linear combinations of W^3_μ and B_μ,

$$\begin{aligned} W^3_\mu &= \cos\theta_W Z_\mu + \sin\theta_W A_\mu \\ B_\mu &= -\sin\theta_W Z_\mu + \cos\theta_W A_\mu, \end{aligned} \tag{2.17}$$

introducing the Weinberg angle, θ_W.

Now let's look at the covariant derivative part of the Lagrangian,

$$\mathcal{L} \supset -\overline{\psi_L}(\frac{g}{2}\not{W}^i\sigma^i + g'Y_\psi\not{B})\psi_L - \overline{\psi_R}g'Y_\psi\not{B}\psi_R. \tag{2.18}$$

Focusing on the second part, and expanding B_μ as per Equation (2.17), we find a term which corresponds to QED; $-\overline{\psi_R}g'Y_\psi cos\theta_W\not{A}\psi_R$; meaning $g'\cos\theta_W = |e|$ and Y_ψ takes the value of the electric charge in units of $|e|$.

The first part contains $X = -\overline{\psi_L}(g\sin\theta_W\frac{\sigma^3}{2} + g'\cos\theta_W Y_\psi)\not{A}\psi_L$, which we can rework,

$$\begin{aligned} X &= -\frac{1}{2}g\sin\theta_W\overline{\psi_L}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\not{A}\psi_L - \overline{\psi_L}|e|Y_\psi\not{A}\psi_L \\ &= -\frac{1}{2}g\sin\theta_W(\overline{\psi_L}^1\not{A}\psi^1_L - \overline{\psi_L}^2\not{A}\psi^2_L) - |e|Y_\psi(\overline{\psi_L}^1\not{A}\psi^1_L + \overline{\psi_L}^2\not{A}\psi^2_L) \\ &= -(\frac{g}{2}\sin\theta_W + |e|Y_\psi)\overline{\psi_L}^1\not{A}\psi^1_L + (\frac{g}{2}\sin\theta_W - |e|Y_\psi)\overline{\psi_L}^2\not{A}\psi^2_L. \end{aligned} \tag{2.19}$$

In both q_L and l_L, the charge of ψ^1_L is one greater than the charge of ψ^2_L. Using this, we can see that $g\sin\theta_W = |e|$. Once we have this, we must choose $Y_q = +\frac{1}{6}$ and $Y_e = -\frac{1}{2}$ to give us the familiar lepton charges of $+2/3(-1/3)$ for the up(down) type quarks and -1 for the charged leptons.

Magic - we have created the electroweak theory. Problems remain, however. Firstly, we have put left and right handed fermions in different representations of $SU(2)\times U(1)$, forbidding us

[4]There is no need for a ν_R in the SM, as this would not appear in any Lagrangian.

from writing a mass term for them. We also require masses for our W and Z bosons, which are forbidden by gauge invariance, but we observe in Nature. Enter Peter Higgs.

The Higgs mechanism

Our salvation comes in the form of the Higgs mechanism [22, 23, 24], by which we shall break electroweak symmetry and restore boson and fermion masses.

We introduce a scalar Higgs field, H which is an $SU(2)$ doublet with hypercharge $Y_H = \frac{1}{2}$ and a Lagrangian,

$$\mathcal{L}_H = (D_\mu H)^\dagger D^\mu H + \mu^2 H^\dagger H - \lambda (H^\dagger H)^2. \tag{2.20}$$

The potential part of this forms the celebrated Mexican hat, which is minimised when $|H| = \frac{1}{\sqrt{2}}\frac{\mu}{\sqrt{\lambda}} \equiv \frac{v}{\sqrt{2}}$, the vacuum expectation value (VEV). Without loss of generality, we may choose $\langle H \rangle = (0, \frac{v}{\sqrt{2}})^T$, for real v. Expanding out the covariant derivative part of $\mathcal{L}_H$ and focusing on the mass terms, we have,

$$\begin{aligned}
(D_\mu H)^\dagger D^\mu H &= \left(\left(\partial_\mu + ig\frac{\sigma^i}{2}W^i_\mu + i\frac{g'}{2}B_\mu\right)H\right)^\dagger \left(\partial^\mu + ig\frac{\sigma^j}{2}W^{j\mu} + i\frac{g'}{2}B^\mu\right)H \\
&\supset \frac{1}{8}(0, v)\begin{pmatrix} gW^3_\mu + g'B_\mu & \sqrt{2}gW^+_\mu \\ \sqrt{2}gW^-_\mu & -gW^3_\mu + g'B_\mu \end{pmatrix}\begin{pmatrix} gW^{3\mu} + g'B^\mu & \sqrt{2}gW^{+\mu} \\ \sqrt{2}gW^{-\mu} & -gW^{3\mu} + g'B^\mu \end{pmatrix}\begin{pmatrix} 0 \\ v \end{pmatrix} \\
&\supset \frac{v^2}{8}\Big(2g^2 W^-_\mu W^{+\mu} + (-gW^3_\mu + g'B_\mu)(-gW^{3\mu} + g'B^\mu)\Big),
\end{aligned} \tag{2.21}$$

$$\begin{aligned}
\frac{v^2}{8}\Big(2g^2 W^-_\mu W^{+\mu} + (gW^3_\mu - g'B_\mu)(gW^{3\mu} - g'B^\mu)\Big) &= \frac{(vg)^2}{4}W^+_\mu W^{-\mu} \\
&\quad + \frac{v^2}{8}(g\cos\theta_W + g'\sin\theta_W)^2 Z_\mu Z^\mu \\
&= \frac{(vg)^2}{4}W^+_\mu W^{-\mu} + \frac{v^2}{8}(g^2 + g'^2)Z_\mu Z^\mu,
\end{aligned} \tag{2.22}$$

where we have used $\sin\theta_W = \frac{g'}{\sqrt{g^2+g'^2}}$ and $\cos\theta_W = \frac{g}{\sqrt{g^2+g'^2}}$ in the last line, and there is no $A_\mu A^\mu$ term, giving us the massless photons we require. We can read off the mass terms for the

three bosons:

$$M_W = \frac{vg}{2}, \quad M_Z = \frac{\sqrt{g^2+g'^2}v}{2} = \frac{M_W}{\cos\theta_W}, \quad M_A = 0. \tag{2.23}$$

With the gauge boson masses fixed, let's turn our attention to the fermions. Adding the Higgs field allows for ($SU(2)$ and $U(1)$ invariant) Yukawa terms [20, 25] in the Lagrangian,

$$\mathcal{L} \supset -\lambda^u \overline{q_L} i\sigma^2 H^* u_R - \lambda^d \overline{q_L} H d_R - \lambda^e \overline{l_L} H e_R + h.c., \tag{2.24}$$

from which we can read off the fermion masses

$$m_u = \frac{v\lambda^u}{\sqrt{2}}, \quad m_d = \frac{v\lambda^d}{\sqrt{2}}, \quad m_e = \frac{v\lambda^e}{\sqrt{2}}, \quad m_\nu = 0. \tag{2.25}$$

The interactions of the Higgs boson h itself can also be derived by sending $v \to v + h(x)$ and looking at the terms that appear. These will not be relevant to this thesis so we simply mention the above for completeness.

2.2 The field content of the Standard Model

We have completed our whistle stop tour of the Standard Model. Summarising everything above, the fields of the SM belong to representations of

$$SU(3)_c \times SU(2)_L \times U(1)_Y, \tag{2.26}$$

where the c denotes colour - the charge of Quantum Chromodynamics (QCD), L denotes the left handed particles which experience the weak nuclear force, and Y denotes hypercharge. We have decomposed the Lorentz group $SO(3,1)$ into $SU(2)_{L,\text{Lor}} \times SU(2)_{R,\text{Lor}}$, where we label $L(R)$ for left(right) handed chirality and the 'Lor' label indicates the spinor representation.

The fields of the SM, as well as their $SU(3)_c \times SU(2)_L \times U(1)_Y$ representations are given in Table 2.1. Left handed fermions are arranged into six $SU(2)_L$ doublets, where we have denoted the three generations of quark doublets generically as $q_L = (u_L, d_L)^T$, and similarly for the three generations of lepton $l_L = (\nu_L, e_L)^T$. Right handed fermions are all singlets, again denoted generically for each generation. There are no right handed neutrinos, as they do not interact in the SM, hence there are nine fermion singlets in total. The different representations for left and right handed fermions forbids mass terms, and lead us to introduce the Higgs model, breaking

Table 2.1: The fields of the Standard Model and their $SU(3)_c \times SU(2)_L \times U(1)_Y$ representations.

Field	$SU(3)_c$	$SU(2)_L$	$U(1)_Y$
g	8	1	0
W	1	3	0
B	1	1	0
$q_L = (u_L, d_L)^T$	3	2	$+\frac{1}{6}$
u_R	3	1	$+\frac{2}{3}$
d_R	3	1	$-\frac{1}{3}$
$l_L = (\nu_L, e_L)^T$	1	2	$-\frac{1}{2}$
e_R	1	1	-1
H	1	2	$+\frac{1}{2}$

electroweak symmetry, and instilling mass into the W, Z and fermions.

2.3 Weak Decays

In this section we will outline the basis for the specific semileptonic decays that will form the backbone of this thesis.

2.3.1 Flavour changing charged currents

In Equation (2.24) we were sly. We rotated the vectors $\psi_L^i \to L_\psi^{ij}\psi_L^j$ and $\psi_R^i \to R_\psi^{ij}\psi_R^j$ so that Yukawa couplings λ were diagonal and we were able to read off the mass terms directly. In reality, the mass eigenstates are not the same as the weak eigenstates. That is to say the Yukawa couplings (in the weak eigenbasis) should have an index for each of the three fermion generations, giving terms like $\lambda_{ij}^d \overline{q_L}^i H d_R^j$. As we like to work in basis of mass eigenstates, the corollary of this is that terms coupling to the W boson in Equation (2.18) are no longer diagonal:

$$\mathcal{L}_{\text{FCCC}} = \frac{|e|}{\sqrt{2}\sin\theta_W}\Big(V_{ij}\overline{u_L}^i \not{W}^+ d_L^j + V_{ij}^*\overline{d_L}^i \not{W}^- u_L^j + \overline{\nu_L}^i \not{W}^+ e_L^i + \overline{e_L}^i \not{W}^- \nu_L^i\Big), \tag{2.27}$$

where V_{ij} is the celebrated Cabbibo-Kobayashi-Maskawa (CKM) matrix [26, 27] and its properties will be discussed below. Because neutrinos have no Yukawa mass term, we are free to rotate them as we please, and so we absorb the rotation of e_L so as to make these terms diagonal in Equation (2.27).

It will be useful to recombine $\psi = \psi_L + \psi_R$ and introduce some new notation in Equation (2.27),

$$\begin{aligned} L_\mu^{ij} &= \frac{1}{2}(\bar{\nu}^i\gamma_\mu e^j - \bar{\nu}^i\gamma_5\gamma_\mu e^j), \\ J_\mu^{ij} &= \frac{1}{2}(\bar{u}^i\gamma_\mu d^j - \bar{u}^i\gamma_5\gamma_\mu d^j) \equiv V_\mu^{ij} - A_\mu^{ij}, \end{aligned} \tag{2.28}$$

$$\mathcal{L}_{\text{FCCC}} = \frac{|e|}{\sqrt{2}\sin\theta_W}\Big(V_{ij}J_\mu^{ij}W^{+\mu} + V_{ij}^* J_\zeta^{ij\dagger}W^{-\zeta} + L_\rho^{ii}W^{+\rho} + L_\sigma^{ii\dagger}W^{-\sigma}\Big). \tag{2.29}$$

The notation here is irritating, but the meanings of the various Ls and Vs should be clear from their subscripts. J is a flavour changing charged current (FCCC), often referred to as the $V-A$ current, where the terms are vector and axial-vector respectively.

The Cabbibo-Kobayashi-Maskawa matrix

The CKM matrix [26, 27] V_{ij} which appears above is unitary by construction ($V = L_u^\dagger L_d$). The individual components of the CKM matrix appear in the amplitudes of various leptonic and semileptonic decays, some of which will be described below. These decays can be used to determine the components of V_{ij} independently, and thus confirm experimentally that V_{ij} is indeed unitary. This is considered to be a strong test of the validity of the SM. The obvious implication of a non-unitary matrix would be to imply a fourth generation of quarks that gets all its mass from the Higgs like the other three. This is ruled out [28], however, there is still room for different new heavy quarks that mix with our own. Currently (not including the work of this thesis), the global average CKM matrix element magnitudes are given by [29],

$$V_{\mathrm{CKM}} = \begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} = \begin{pmatrix} 0.97370(14) & 0.22450(80) & 0.00382(24) \\ 0.2210(40) & 0.987(11) & 0.0410(14) \\ 0.00800(30) & 0.0388(11) & 1.013(30) \end{pmatrix}, \tag{2.30}$$

and are consistent with unitarity, within uncertainty.

Being a 3×3 unitary matrix, V_{CKM} can always be specified with nine real numbers. Five of these parameters can be absored as phases into the quark fields, leaving behind four physical degrees of freedom. These can be expressed as three mixing angles and one complex phase. This phase is responsible for CP violation in the SM.

For our purposes, it is most instructive to talk about the matrix in terms of the magnitude of the transition elements as above, but other parameterisations make the mixing angles and phase more explicit. In particular, the standard parameterisation is written in terms of the mixing angles $\theta_{12}, \theta_{13}, \theta_{23}$ and the phase δ_{13}, whilst the Wolfenstein parameterisation [30] introduces parameters λ, A, ρ, and η which are related to those of the standard parameterisation. In particular [31],

$$\lambda = 0.22500^{+0.0085}_{-0.00022} \quad A = 0.8132^{+0.0119}_{-0.0060} \quad \bar{\rho} = 0.1566^{+0.0085}_{-0.0048} \quad \bar{\eta} = 0.3475^{+0.0118}_{-0.0054}, \tag{2.31}$$

where $\bar{\rho} = \rho \times (1 - \lambda^2/2 + ...)$ and similarly for $\bar{\rho}$. It is in this parameterisation which CKM unitarity is most easily visualised, as in Figure 2.1, which is a (somewhat baroque) representation

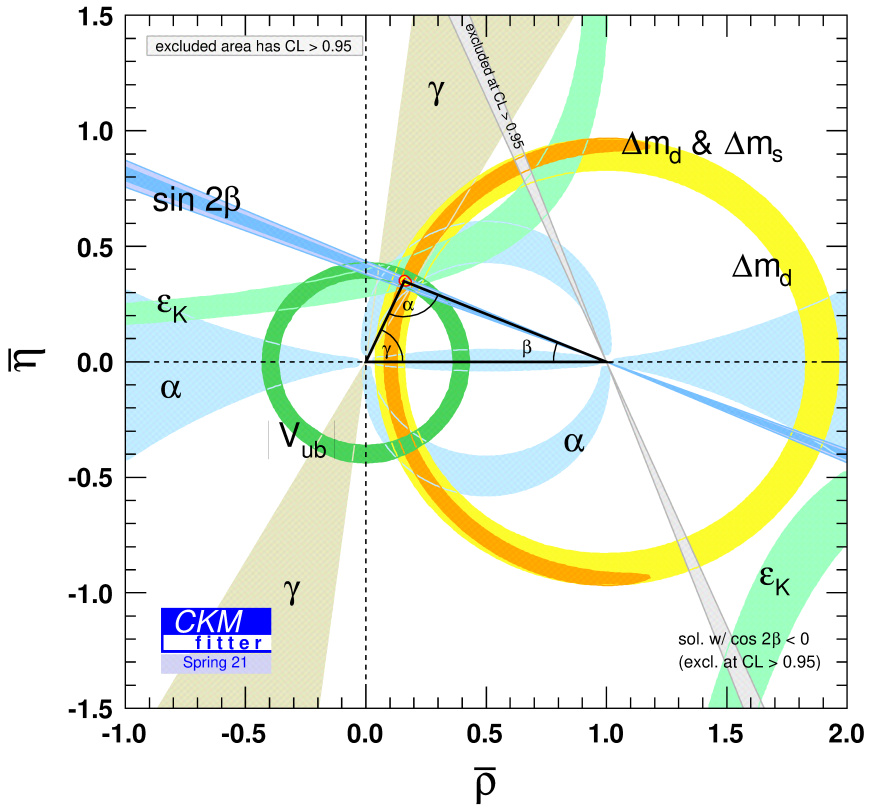

Figure 2.1: Unitarity triangle of the CKM matrix (courtesy of CKMfitter [31]).

of CKM unitarity. The unitarity triangle, constructed from angles,

$$\alpha = \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right), \quad \beta = \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right), \quad \gamma = \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right), \tag{2.32}$$

is shown, together with constraints from a variety of different processes. The fact that the lines meet in the top corner (angle α) indicates that the matrix is unitary and represents one of the great triumphs of the Standard Model.

One of the goals of the work in this thesis is to provide a more precise determination of the element $|V_{cs}|$, via the semileptonic decay $D \to K\ell\nu$. This is the subject of Chapter 5.

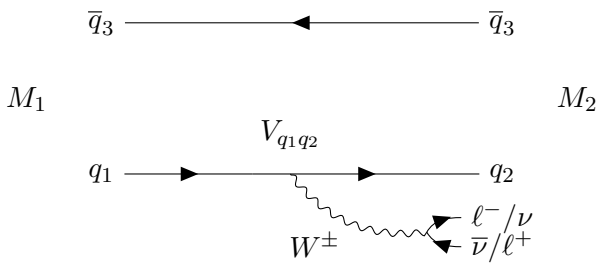

Figure 2.2: Feynman diagram for a FCCC semileptonic weak decay $M_1 \to M_2 \ell \nu$.

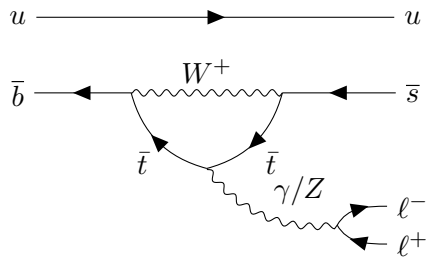

Figure 2.3: Example of a FCNC Feynman diagram for $B^+ \to K^+ \ell^+ \ell^-$.

2.3.2 Flavour changing neutral currents

The FCCC outlined above is a tree level process, and always involves the transition of an up type quark into a down type quark, or vice versa, leading to the current carrying a charge of $\pm|e|$. The associated tree level Feynman diagram is shown (in the context of a semileptonic[5] decay) in Figure 2.2.

The other form of decay which we will study in this thesis involves the transition of a up(down) type quark to another of the same type; specially we focus here on $b \to s$ decays, again in a semileptonic context. In this case Equation (2.27) shows us that there are no such flavour changing neutral currents (FCNCs) at tree level in the SM, and we must go to loop level to find examples.

An example of a FCNC loop is given for the $B^+ \to K^+ \ell^+ \ell^-$ semileptonic decay in Figure 2.3. FCNC decays are rare on account of the loop suppression, and so they constitute a good place to look for new physics. For example, a new Z type boson, the Z', which couples only to μ and τ, could cause the decay rates to vary between different leptons, violating so called Lepton Flavour

[5] I.e. a meson decaying to another meson and a pair of leptons.

Universality (LFU) [32, 33]. An alternative explanation for this, if it is definitively observed, would be to introduce a right handed neutrino, and treat the leptons in the same manner as the quarks in Equation (2.27). This would introduce neutrino masses, which we do observe, but their extremely small size means that this effect would be tiny [34, 35], and not a candidate for the level of LFU violation we already see at our present experimental uncertainty [9].

Alternatively, the existence of leptoquarks [36], coupling leptons to quarks, would completely change the diagram in Figure 2.3, and with it the measured branching fraction [37]. In order to search for evidence of these new particles, we must measure such rare FCNC decays experimentally, and also determine the SM expectations, which is what we shall do in this thesis.

Weak effective theory and Wilson coefficients

Treating the Standard Model as as effective field theory (SMEFT), is a large topic in and of itself, to which we will not even attempt to do justice (for a recent summary, see [38]). We will, however, use it to outline the purpose of effective theories, which will simplify the description of weak effective theory (WET) which is required in Chapter 7.

Aware that all good things must come to an end; that the SM, whilst remarkably powerful at low energies, cannot work up to arbitrarily high energy; we nobly seek to find out what exists beyond. Being unable to directly access the kind of energies where new physics might start to appear on shell, we make do by writing,

$$\mathcal{L}_{\text{SMEFT}} = \mathcal{L}_{\text{SM}} + \sum_{i=5}^{\infty} \sum_{j=1}^{n_j} \frac{c_j^i}{\Lambda^{i-4}} \mathcal{O}_j^i, \tag{2.33}$$

where c_j^i are a set of n_j Wilson coefficients, each multiplying an operator $\mathcal{O}_j^i$ of dimension i, and Λ is some large energy cutoff, above which we have integrated out degrees of freedom. In SMEFT, we set this cut off well above the masses of all known particles.

We have adorned the SM, previously containing only (mass) dimension four operators, with operators of dimension five and above, so called irrelevant operators, which are non-renormalisable. These operators have running couplings in scale Λ, and are suppressed at low energies E by $(E/\Lambda)^{i-4}$. They allow new physics effects from high energies to leech into the low energy SM Lagrangian, where they can be picked up by high precision experimental measurements which differ from the SM.

Coefficient	Value	Operator
$C_1(\mu_b)$	-0.294(9)	$Q_1^p = (\bar{s}_L\gamma_\mu T^a p_L)(\bar{p}_L\gamma^\mu T^a b_L)$
$C_2(\mu_b)$	1.017(1)	$Q_2^p = (\bar{s}_L\gamma_\mu p_L)(\bar{p}_L\gamma^\mu b_L)$
$C_3(\mu_b)$	-0.0059(2)	$Q_3 = (\bar{s}_L\gamma_\mu b_L)\sum_p(\bar{p}\gamma^\mu p)$
$C_4(\mu_b)$	-0.087(1)	$Q_4 = (\bar{s}_L\gamma_\mu T^a b_L)\sum_p(\bar{p}\gamma^\mu T^a p)$
$C_5(\mu_b)$	0.0004	$Q_5 = (\bar{s}_L\gamma_\mu\gamma_\nu\gamma_\rho b_L)\sum_p(\bar{p}\gamma^\mu\gamma^\nu\gamma^\rho p)$
$C_6(\mu_b)$	0.0011(1)	$Q_6 = (\bar{s}_L\gamma_\mu\gamma_\nu\gamma_\rho T^a b_L)\sum_p(\bar{p}\gamma^\mu\gamma^\nu\gamma^\rho T^a p)$
$C_7^{\text{eff},0}(\mu_b)$	-0.2957(5)	$Q_7 = \frac{e}{16\pi^2}m_b(\bar{s}_L\sigma^{\mu\nu}b_R)F_{\mu\nu}$
$C_8^{\text{eff}}(\mu_b)$	-0.1630(6)	$Q_8 = \frac{g_s}{16\pi^2}m_b(\bar{s}_L\sigma^{\mu\nu}T^a b_R)G^a_{\mu\nu}$
$C_9(\mu_b)$	4.114(14)	$Q_9 = \frac{e^2}{16\pi^2}(\bar{s}_L\gamma_\mu b_L)\sum_\ell(\bar{\ell}\gamma^\mu\ell)$
$C_{10}(\mu_b)$	-4.193(33)	$Q_{10} = \frac{e^2}{16\pi^2}(\bar{s}_L\gamma_\mu b_L)\sum_\ell(\bar{\ell}\gamma^\mu\gamma_5\ell)$

Table 2.2: Wilson coefficients used in Chapter 7, as well as their corresponding operators [39]. Central values and uncertainties are quoted at $\mu_b = 4.2\,\text{GeV}$ and taken from [40]. Note that the operators given for $C_7^{\text{eff},0}$ and C_8^{eff} are those for C_7 and C_8, with the effective coefficient containing contributions from other operators (see text). The sums over p run over all quarks except the top.

Analogous to this is weak effective theory, which takes the same principle as SMEFT, but integrates out degrees of freedom above the electroweak scale (i.e. t, W, Z, h, for cutoff M_W). These higher dimensional operators allow for both SM and non SM interactions - with their value in the SM being determined by calculating Feynman diagrams in the SM and matching the result onto the effective theory. Any difference from this SM value would indicate new physics contributions to that operator. In Section 7.4.1 below, we will require certain Wilson coefficients of WET dimension six operators [39] to construct the SM differential branching fraction for $B \to K\ell^+\ell^-$. Here we shall discuss these coefficients explicitly.

The effective Hamiltonian for $b \to s$ decays in the SM is given by [40]

$$\mathcal{H}_{\text{eff}} = \frac{4G_F}{\sqrt{2}}\left(\lambda_u^{(s)}\sum_{i=1}^{2}C_iQ_i^u + \lambda_c^{(s)}\sum_{i=1}^{2}C_iQ_i^c - \lambda_t^{(s)}\sum_{i=3}^{10}C_iQ_i - \lambda_t^{(s)}C_\nu Q_\nu + \text{h.c.}\right), \tag{2.34}$$

where $\lambda_p^{(q)} = V_{pb}V_{pq}^*$ (with CKM matrix elements V_{ab}) and we have denoted operators Q_i and Wilson coefficients C_i. We focus on $\ell^+\ell^-$ decays, so we ignore the Q_ν terms for now. The operators Q_i are given in Table 2.2 along with their SM values at scale $\mu = 4.2$ GeV [40].

Under renormalisation, QCD corrections lead to a mixing of operators, governed by the anoma-

lous dimension matrix γ,

$$\mu \frac{d}{d\mu} \vec{C} = \gamma^T \vec{C}. \tag{2.35}$$

This is required because SM matching to WET occures at the electroweak scale, and coefficients must then be run down to our scale of choice (in our case $\mu = 4.2$ GeV).

Effective coefficients C_7^{eff} and C_8^{eff} which are regularisation scheme independent are often defined as combinations of $C_{7/8}$ and the other Wilson coefficients (see [40] for details). It is these values which are given in Table 2.2. We also calculate a C_9^{eff}, which is discussed in more detail in Chapter 7.

2.3.3 Semileptonic decays

In this work, we shall concern ourselves with semileptonic decays of mesons. A generic (FCCC) semileptonic decay from meson M_1 to meson M_2 is depicted at tree level in Figure 2.2. The amplitude is given by,

$$\begin{aligned} \mathcal{A} =& \left(\frac{ie}{\sqrt{2}\sin\theta_W} \right)^2 V_{q_1 q_2} \left\langle M_2, \ell\nu \right| J_\mu^{q_1 q_2} \left(\frac{g^{\mu\nu}}{p^2 - M_W^2} \right) L_\nu^l \left| M_1 \right\rangle \\ \approx& -2\sqrt{2} G_F V_{q_1 q_2} \left\langle M_2, \ell\nu \right| J_\mu^{q_1 q_2} g^{\mu\nu} L_\nu^l \left| M_1 \right\rangle \\ \approx& -2\sqrt{2} G_F V_{q_1 q_2} \left\langle \ell\nu \right| L^{\mu,l} \left| \Omega \right\rangle \left\langle M_2 \right| J_\mu^{q_1 q_2} \left| M_1 \right\rangle, \end{aligned} \tag{2.36}$$

where θ_W is the weak mixing angle, $|\Omega\rangle$ is the vacuum, and $L^{\mu,l}$ and $J_\mu^{q_1 q_2}$ are defined in Equation (2.28). In the second line we have used the fact that p^2, the 4-momentum carried by the W, is much less than its mass. Hence we can 'integrate out' the W boson $\frac{1}{p^2 - M_W^2} \approx \frac{-1}{M_W^2}$, and make use of the definition (in our convention) of Fermi's constant $G_F = \frac{\sqrt{2}e^2}{8 M_W^2 \sin^2\theta_W}$. What we've achieved is to factorise the perturbative electroweak part from the QCD part; the hadronic matrix element $\langle M_2 | J_\mu^{q_1 q_2} | M_1 \rangle$; which we cannot calculate perturbatively. The essence of this thesis is the calculation of such matrix elements numerically, using lattice QCD, which will be detailed in the next chapter. In this work we will only consider pseudoscalar to pseudoscalar decays in which the parity invariance of QCD means that only the vector component $V_\mu^{q_1 q_2}$ of $J_\mu^{q_1 q_2}$ contributes.

The hadronic matrix element is parameterised in terms of form factors, and in this work we will always use the parameterisation in terms of scalar and vector form factors $f_0(q^2)$ and $f_+(q^2)$

where,

$$\langle M_2|\, V_\mu^{q_1 q_2}\, |M_1\rangle = \Big(p_\mu^{M_1} + p_\mu^{M_2} - \frac{M_{M_1}^2 - M_{M_2}^2}{q^2} q_\mu\Big) f_+(q^2) + \frac{M_{M_1}^2 - M_{M_2}^2}{q^2} q_\mu f_0(q^2). \tag{2.37}$$

Here we define 4-momenta $p_\mu^{M_i}$ and masses M_{M_i} for each meson, and $q \equiv (p_{M_1} - p_{M_2})$. Similar equations, involving scalar and tensor current insertions, will be given in Section 4.4.

In this thesis, the only decay which involves a FCCC is $D \to K\ell\nu$, which will be discussed in Chapter 5. For this reason, we will give the differential decay rate for the case $M_1 = D$, $M_2 = K$. Integrating $|\mathcal{A}|^2$ over final state momenta, the expression for this differential decay rate can be written,

$$\begin{aligned}\frac{d\Gamma}{dq^2}(D \to K\ell\nu) = \frac{G_F^2}{24\pi^3}(\eta_{EW}|V_{cs}|)^2(1-\epsilon)^2(1+\delta_{EM})\times \\ \left[|\vec{p}_K|^3(1+\frac{\epsilon}{2})|f_+(q^2)|^2 + |\vec{p}_K| M_D^2 \left(1 - \frac{M_K^2}{M_D^2}\right)^2 \frac{3\epsilon}{8}|f_0(q^2)|^2\right],\end{aligned} \tag{2.38}$$

where $\epsilon = m_\ell^2/q^2$ and the factors η_{EW} and δ_{EM} allow for electroweak and electromagnetic corrections which are discussed in Chapter 5. It is clear from this equation that the calculation of hadronic form factors f_0 and f_+, combined with experimental measurements of Γ, provide a method to independently determine elements of the CKM matrix, in this case $|V_{cs}|$.

For the FCNC in $B \to K\ell^+\ell^-$, which is the topic of Chapter 7, we can also parameterise the vector matrix element as in Equation (2.37). We will also require scalar and tensor matrix elements, which will be discussed in Section 4.4. The scalar matrix element will allow us to separate f_0 from f_+ in Equation (2.37), whilst the tensor will be required to parameterise new physics.

2.4 The path integral formalism

The methods we shall need to construct lattice QCD in the next chapter are based on a discretisation of the path integral formalism, which we shall briefly refresh here. What we shall focus on calculating is the time ordered product of fields, which we can relate to an n point correlation

function $C_n(x_1, ..., x_n)$.

$$C_n(x_1, ..., x_n) = \langle\Omega| T\varphi_H(x_1)...\varphi_H(x_n) |\Omega\rangle , \tag{2.39}$$

where we define scalar fields in the Heisenberg picture, φ_H, and the ground state of the interacting theory $|\Omega\rangle$. We wish to convert between this expression, and one using the vacuum of the free theory. Below we shall sketch the method for this change.

For interacting Hamiltonian $H = H_0 + H_{\text{int}}$, we have the free theory vacuum $H_0 |0\rangle = 0$ and the interacting theory ground state $|\Omega\rangle$ satisfying $\langle\Omega| H |\Omega\rangle = E_0$. $H_I = e^{iH_0(t-t_0)} H_{\text{int}} e^{-iH_0(t-t_0)}$ is the interaction Hamiltonian in the interaction picture.

Because H_I is a small perturbation, there must be some overlap of the ground state of the interacting theory with that of the free theory, such that $\langle\Omega|0\rangle \neq 0$. As such, when we insert a complete set of states into $e^{-iHT} |0\rangle$, the evolution to time T of $|0\rangle$, we find,

$$e^{-iHT} |0\rangle = \sum_n e^{-iE_nT} |n\rangle \langle n|0\rangle = e^{-iE_0T} |\Omega\rangle \langle\Omega|0\rangle + \sum_{n\neq 0} e^{-iE_nT} |n\rangle \langle n|0\rangle . \tag{2.40}$$

Now because $E_0 < E_{n\neq 0}$, if we take $T \to \infty(1 - i\epsilon)$, we can kill off all the $n \neq 0$ terms faster than the E_0 term. This physically corresponds to saying our fields are non interacting in the far past (at $-T$) and the far future (at $+T$) of our interaction.

Using this trick to convert between the interacting ground state and the free theory vacuum, we can write,

$$C_n(x_1, ..., x_n) = \langle\Omega| T\varphi_H(x_1)...\varphi_H(x_n) |\Omega\rangle = \lim_{T\to\infty(1-i\epsilon)} \frac{\langle 0| T\varphi(x_1)...\varphi(x_n) e^{-i\int_{-T}^{T} dt H_I(t)} |0\rangle}{\langle 0| e^{-i\int_{-T}^{T} dt H_I(t)} |0\rangle}. \tag{2.41}$$

A full derivation of this is lengthy and tangential to this thesis, but can be found in [11]. The important aspect which we shall use later is the origin of the $T \to \infty(1 - i\epsilon)$ limit, which we have addressed above.

We have now written C_n as an expression using the vacuum of the free theory $|0\rangle$ where φ are in the interaction picture. This can be written in terms of an integral over all possible field

configurations, the path integral,

$$C_n(x_1, ..., x_n) = \frac{1}{\mathcal{Z}} \int D\varphi\varphi(x_1)...\varphi(x_n)e^{i\int d^4x\mathcal{L}[\varphi]}, \tag{2.42}$$

where $\mathcal{Z}$ is the partition function; i.e. the path integral with no field insertions.

It is the path integral which provides many of the challenges in QCD. In QED (and QCD at high energy), the coupling is small and allows for a convergent expansion of the integral in increasing powers of the coupling. For QCD, however, the coupling is large at low energies, forbidding such a perturbative expansion (which is possible at higher energies). This is what drives us to solve the path integral numerically, and thus to consider lattice QCD.

2.4.1 Correlation functions

Once we have computed an n-point correlation function (Equation (2.41)) in our case on the lattice (see Section 4.2), we need to derive the fit form required to extract various parameters from it. In this thesis, we will focus only on two- and three-point correlation functions ($C_2(x_2, x_1)$ and $C_3(x_3, x_2, x_1)$). We make use of a Wick rotation [41] ($t \to -it$), to convert from Minkowski space to Euclidean space. The effect of this is to convert the i in the exponential of the path integral to a -1. When we come to computing the integral numerically, we shall see that this removal of the oscillation makes the integral tractable. Note that this also necessitates a change in the limit: $T \to -i\infty(1 - i\epsilon)$.

For a Euclidean two point correlation function C_2 with $t_2 > t_1$, and where we have made the

interaction picture φ_I explicit, we have,

$$\begin{aligned}
C_2(x_2,x_1) &= \lim_{T\to-\infty(\epsilon+i)} \frac{\langle 0|\,\varphi_I(x_2)\varphi_I(x_1)e^{-\int_{-T}^{T} dtH_I(t)}\,|0\rangle}{\langle 0|\,e^{-\int_{-T}^{T} dtH_I(t)}\,|0\rangle} \\
&= \frac{\langle 0|\,\varphi_I(x_2)\varphi_I(x_1)\,|0\rangle}{\langle 0|0\rangle} \\
&= \sum_n \langle 0|\,\varphi_I(x_2)\,|n\rangle\,\langle n|\,\varphi_I(x_1)\,|0\rangle \\
&= \sum_n \langle 0|\,e^{H_0(t_2-t_0)}\varphi(t_0,\vec{x}_2)e^{-H_0(t_2-t_0)}\,|n\rangle\,\langle n|\,e^{H_0(t_1-t_0)}\varphi(t_0,\vec{x}_1)e^{-H_0(t_1-t_0)}\,|0\rangle \\
&= \sum_n e^{-(E_n-E_0)(t_2-t_1)}\,\langle 0|\,\varphi(t_0,\vec{x}_2)\,|n\rangle\,\langle n|\,\varphi(t_0,\vec{x}_1)\,|0\rangle \\
&\equiv \sum_n e^{-E_n t}\,\langle 0|\,\varphi(\vec{x}_2)\,|n\rangle\,\langle n|\,\varphi(\vec{x}_1)\,|0\rangle\,,
\end{aligned} \tag{2.43}$$

where we have used the fact that our interaction picture φ_I is equal to $\varphi_I = e^{H_0(t-t_0)}\varphi(t_0,\vec{x})e^{-H_0(t-t_0)}$ for free Hamiltonian H_0. In the last line, we have redefined the energy relative to the vacuum $E_n \to E_n - E_0$ and taken $t = t_2 - t_1$. We see that the two point correlation function time dependence is simply dependent on the energy relative to the vacuum and the time between the current insertions, whilst the amplitude gives us matrix elements for the insertions. For large times t, the ground state E_1 dominates. Similarly for three-point correlation functions C_3,

$$\begin{aligned}
C_3(x_3,x_2,x_1) &= \sum_{m,n} e^{-(E_m-E_0)(t_3-t_2)}e^{-(E_n-E_0)(t_2-t_1)}\,\langle 0|\,\varphi(t_0,\vec{x}_3)\,|m\rangle\,\langle m|\,\varphi(t_0,\vec{x}_2)\,|n\rangle\,\langle n|\,\varphi(t_0,\vec{x}_1)\,|0\rangle \\
&\equiv \sum_{m,n} e^{-E_m(T-t)}e^{-E_n(t)}\,\langle 0|\,\varphi(\vec{x}_3)\,|m\rangle\,\langle m|\,\varphi(\vec{x}_2)\,|n\rangle\,\langle n|\,\varphi(\vec{x}_1)\,|0\rangle\,,
\end{aligned} \tag{2.44}$$

where we again relabel the energies relative to the vacuum and define $t = t_2 - t_1$ and $T = t_3 - t_1$. These are the forms which we shall fit our correlation functions to in Section 4.3.

Chapter 3

Lattice Quantum Chromodynamics

Here we introduce the basic principles of Lattice QCD, from discretising the path integral through to the Highly Improved Staggered Quark (HISQ) formalism that we will use for all calculations in this work.

3.1 Motivation

The need for lattice QCD arises because the strong coupling $\alpha_s = g_s^2/4\pi$ is large at low energies, leading to the confinement of quarks [42] and gluons within hadrons. Consequently, whilst high energy QCD can be treated as a perturbative expansion in the coupling, at low energies, we must look for a different approach. The approach we choose is to solve the path integral numerically, which naturally involves discretising the theory in order for it to be treated on a computer.

3.2 Basic principles

A good overview of the basic principles of Lattice QCD is given in [43], and here I will walk through the key points from that introduction.

3.2.1 Discretising the path integral

Let's begin with a toy model of a theory with real scalar fields, ϕ. Starting with the path integral from Equation (2.42), and Wick-rotating the four vectors $x \to (-ix^0, \vec{x})$ (see Section 2.4.1 above), we can write the Euclidean correlation function for ϕ in terms of the Euclidean action $S[\phi]$, the partition function $\mathcal{Z} = \int \mathcal{D}\phi e^{-S[\phi]}$, and some operator insertion of our choice, which we shall call $\Gamma[\phi]$,

$$C_n(x) = \frac{1}{\mathcal{Z}} \int \mathcal{D}\phi \Gamma[\phi] e^{-S[\phi]}. \tag{3.1}$$

The purpose of Wick rotating now becomes apparent; we have replaced the oscillating phase in the integral, which is very difficult to evaluate numerically, with an exponential decay, which is much more manageable (but still a challenge in a large number of dimensions).

Next, we discretise spacetime onto a 4D torus of points $x_j = (x_j^0, x_j^1, x_j^2, x_j^3)$, with spacing a such that $x_0 = (0,0,0,0)$, $x_1 = (a,0,0,0)$, ... and so on. j then runs up to the total volume of this lattice in units of a^4; Vol $= N_t N_x N_y N_z$, where N_μ is the total number of lattice sites in the μ direction. The purpose of this step is to provide regulators (from the finite lattice spacing and finite total lattice volume), which prevent infrared and ultraviolet divergences. A side effect of this is that we can treat the problem numerically, as we now have a finite number of degrees of freedom to simulate. In this discretised world, integrals over space become sums over lattice sites, and derivatives become finite differences. The measure $\mathcal{D}\phi$ becomes,

$$\mathcal{D}\phi = \prod_{x_j \in \text{lattice}} d\phi(x_j). \tag{3.2}$$

Now we address the computation of the path integral itself, using Monte Carlo [44, 45] integration. We generate a number, N_{cf}, of random field configurations,

$$\phi^{(\alpha)} = \{\phi^{(\alpha)}(x_{j=0}), ..., \phi^{(\alpha)}(x_{j=\text{Vol}})\} \quad \alpha = 1, 2...N_{\text{cf}}, \tag{3.3}$$

where the subscript j again denotes the position on the 4D lattice of spacing a, and the superscript labels the configuration number. If these configurations are drawn from the probability distribution $P[\phi^{(\alpha)}] \propto e^{-S[\phi^{(\alpha)}]}$, then as N_{cf} becomes large the correlation function in Equa-

tion (3.1) can be approximated as,

$$C_n[\phi] \approx \frac{1}{N_{\rm cf}} \sum_{\alpha=1}^{N_{\rm cf}} \Gamma[\phi^{(\alpha)}]. \tag{3.4}$$

For a two point correlation function on the lattice, we use an operator like $\Gamma(t)\Gamma(0)$, where

$$\Gamma(t) = \frac{1}{\sqrt{N}} \sum_{\vec{x}_j} \phi(t, \vec{x}_j), \tag{3.5}$$

with the sum enforcing zero three-momentum.

This Monte Carlo method [46, 47, 48, 49, 50, 51] is the basis for lattice calculations. We generate a large number of configurations according to the distribution specified by the action, and then take an average over these to estimate the value of the path integral. The basic method for generating configurations is by a Markov Chain, such as using the Metropolis Algorithm [51]. To update ϕ_{x_j} we proceed as follows:

- Send $\phi(x_j) \to \phi(x_j) + \zeta$, where ζ is a random number drawn from a uniform distribution in the range $(-\epsilon, \epsilon)$
- Evaluate the change in the action ΔS caused by the change in $\phi(x_j)$
- If ΔS is negative, keep the new value for $\phi(x_j)$ and proceed to $\phi(x_{j+1})$
- If ΔS is positive, accept the new value only if $e^{-\Delta S}$ is greater than a random number generated between 0 and 1. Otherwise keep the old value and proceed to $\phi(x_{j+1})$.

This algorithm (and improved versions of it [46]) produces configurations obeying the desired distribution, the key step being that some changes which increase the action are allowed at random, to ensure that the configurations don't just settle into a state which minimises the action or a local minimum.

3.2.2 The gluon action

For a full QCD simulation, the principle is the same, except we need quark and gluon fields (ψ and A_μ) in our path integral, which now looks something like

$$\int \mathcal{D}A\mathcal{D}\psi\mathcal{D}\bar{\psi}e^{-S[A,\psi,\bar{\psi}]}, \tag{3.6}$$

where the action is given by the Euclidean QCD Lagrangian, as discussed in Section 2.1.3,

$$\mathcal{L}_{\text{QCD}} = -\frac{1}{4}F^a_{\mu\nu}F^{a\mu\nu} + \sum_{f\in\{u,d,s,c,t,b\}} \bar{\psi}\left(\partial\!\!\!/ + g_s A\!\!\!/^a \frac{\lambda^a}{2} + m_f\right)\psi, \tag{3.7}$$

for coupling g_s, and gluon field strength tensor $F^a_{\mu\nu}$ where summation over repeated colour indices (Latin) and repeated 4-indices (Greek) is implied.

The simplest QCD calculations we can do are quenched, that is to say involving only gluons, and no sea quarks. On the lattice, we choose to represent gluon fields not by $A_\mu(x)$ at points on the lattice, but rather by links between lattice sites, defined by

$$U_\mu(x) = \mathcal{P}e^{-i\int_x^{x+a\hat{\mu}} g_s A\cdot dx} \approx e^{-iag_sA_\mu}, \tag{3.8}$$

where again a is the lattice spacing, $\hat{\mu}$ is a unit vector in the μ direction, and $\mathcal{P}$ denotes path ordering. This trick enables us to preserve gauge invariance on the lattice. Under an SU(3) gauge transformation, $G(x) = e^{i\alpha(x)}$,

$$\begin{aligned} U_\mu(x) &\to G(x)U_\mu(x)G^\dagger(x+a\hat{\mu}) \\ \psi(x) &\to G(x)\psi(x). \end{aligned} \tag{3.9}$$

We define $U_\mu(x)$ as a link from x to $x+a\hat{\mu}$ and $U^\dagger_\mu(x) = U_{-\mu}(x+a\hat{\mu})$ as the inverse of this, from $x+a\hat{\mu}$ to x. It is clear from the fact that $GG^\dagger = 1$ that in any connected loop of links, or a chain of links from a $\bar{\psi}$ to a ψ will always lead to cancellation of all Gs, and so, remarkably, these quantities are gauge invariant on the lattice.

The simplest such gauge invariant quantity is the plaquette [42, 52], an $a\times a$ loop,

$$P_{\mu\nu}(x) = \frac{1}{3}\text{ReTr}\big(U_\mu(x)U_\nu(x+a\hat{\mu})U^\dagger_\mu(x+a\hat{\nu})U^\dagger_\nu(x)\big), \tag{3.10}$$

where the trace is over colour space. To see why this is useful we must expand in powers of a (i.e. $-iag_sA$) in the limit where we have a slowly varying gauge field centred on x_0,

$$P_{\mu\nu} = \frac{1}{3}\mathrm{ReTr}\left[1 - i\oint g_sA\cdot dx - \frac{1}{2}\Big(\oint g_sA\cdot dx\Big)^2 ...\right], \tag{3.11}$$

where, by Green's theorem, and to first order (no implied summation over indices),

$$\begin{aligned}\oint A\cdot dx &= \int_{-\frac{a}{2}}^{\frac{a}{2}} dx_\mu dx_\nu \Big[\partial_\mu A_\nu(x_0+x) - \partial_\nu A_\mu(x_0+x)\Big] \\ &= \int_{-\frac{a}{2}}^{\frac{a}{2}} dx_\mu dx_\nu \Big[F_{\mu\nu}(x_0) + (x_\mu D_\mu + x_\nu D_\nu)F_{\mu\nu}(x_0) \\ &+ \frac{1}{2}(x_\mu^2 D_\mu^2 + x_\nu^2 D_\nu^2)F_{\mu\nu}(x_0) + ...\Big] \\ &= a^2F_{\mu\nu}(x_0) + \frac{a^4}{24}(D_\mu^2 + D_\nu^2)F_{\mu\nu}(x_0) + ... \ .\end{aligned} \tag{3.12}$$

This gives[1]

$$P_{\mu\nu} = \frac{1}{3}\mathrm{Tr}\big[1 - \frac{a^4}{2}F_{\mu\nu}^2 - \frac{g_s^2a^6}{24}F_{\mu\nu}(D_\mu^2 + D_\nu^2)F_{\mu\nu} + ...\big]\big|_{x_0}, \tag{3.13}$$

from this equation, we can define the Wilson action [42], which differs from the gluon continuum action $\mathcal{L}_{\mathrm{glue}} = -\frac{1}{4}\mathrm{Tr}F_{\mu\nu}F^{\mu\nu}$ at order a^2,

$$\begin{aligned}S_{\mathrm{W}} &= \frac{6}{g_s^2}\sum_{x,\mu>\nu}(1 - P_{\mu\nu}) \\ &= \int d^4x \sum_{\mu\nu}\frac{1}{2}\mathrm{Tr}\left[F_{\mu\nu}^2 + \frac{a^2}{12}F_{\mu\nu}(x_0)(D_\mu^2 + D_\nu^2)F_{\mu\nu}(x_0) + ...\right].\end{aligned} \tag{3.14}$$

In Equation (3.14), the strength of the coupling g_s is the means by which we set the scale of our lattice. It is usually written in terms of $\beta = \frac{6}{g_s^2}$, which will be discussed more in Section 3.4.

The Wilson action is the naive action of lattice QCD. It can be shown (e.g. [43]) that adding appropriately chosen combinations of loops of different shapes to this action can cancel this a^2 and higher order differences [53], and lead to improved actions, with even better convergence to the continuum. This process is called Symanzik improvement [54]. For example, if we play the

[1] Indices here are specified by the plane of the plaquette - there are no sums involved. The trace is over colour.

same game as above with a $2a \times a$ rectangle, $R_{\mu\nu}$ we find,

$$R_{\mu\nu} = 1 - \frac{4}{6}a^4 \mathrm{Tr}(g_s F_{\mu\nu})^2 + \mathcal{O}(a^6) \tag{3.15}$$

which can be combined with the plaquette to cancel the a^2 terms in the action, leaving it accurate up to $\mathcal{O}(a^4)$

$$S_{\mathrm{imp}} = \frac{6}{g_s^2} \sum_{x,\mu>\nu} \Big(\frac{5P_{\mu\nu}}{3} - \frac{R_{\mu\nu} + R_{\nu\mu}}{12} \Big). \tag{3.16}$$

This can be repeated to higher orders with different shaped loops.

Even after Symanzik improvement, this action corresponds to pure 'glue', with no sea quarks, and so is not a wholly accurate representation of QCD. Despite this, due to the immense cost of putting sea quarks on the lattice, for many years calculations were performed with only valence quarks, the so called quenched approximation.

Finally, we mention in passing the link operator is 'tadpole improved' [55]: $U_\mu \to U_\mu / u_0$, where $u_0 = \langle 0 | P_{\mu\nu} | 0 \rangle^{1/4}$. This procedure cancels 'tadpoles'; higher order terms in the e^{-iagA_μ} expansion, which correspond to vertices with multiple powers of agA_μ.

3.3 Staggered quarks and the HISQ formalism

Modern lattice calculations are based on the methods described above, but use more sophisticated Symanzik improved [54] actions for the gluons than the basic Wilson action described above [42]. They also include both valence and sea quarks (including only valence quarks is often described as working in the quenched approximation). In this section, we will discuss the addition of fermions to the gluon action.

Throughout this thesis we will use a staggered quark formalism called Highly Improved Staggered Quarks (HISQ) which was developed in [10] by the HPQCD collaboration, and which we will construct below.

3.3.1 The naive quark action and doubling

Quarks appear in the QCD action via the second term of $\mathcal{L}_{\mathrm{QCD}}$ in (3.7). We can write,

$$S = S_{\mathrm{glue}} + \sum_x \bar{\psi}(\not{D} + m)\psi = S_{\mathrm{glue}} + \bar{\psi} M \psi, \tag{3.17}$$

where we have defined matrix $M = (\not{D} + m)$.

As discussed in Section 3.2.2, it's possible to use various Wilson loops to reduce the discretisation errors in the gluon part of the action, but we still have to discretise the quark part of the action $S_q = \bar{\psi} M \psi$. Naively, we might consider simply transforming the covariant derivative into a discrete difference, Δ_μ, such that,

$$\Delta_\mu \psi(x) = \frac{1}{2a}\Big(U_\mu(x)\psi(x + a\hat{\mu}) - U^\dagger_\mu(x - a\hat{\mu})\psi(x - a\hat{\mu})\Big). \tag{3.18}$$

This action suffers from a 'doubling' symmetry [56, 57], as explained in [10]. Recalling that in Euclidean space, where $\{\gamma_\mu, \gamma_\nu\} = 2\delta_{\mu\nu}$,

$$\gamma^\dagger_\mu = \gamma_\mu \qquad \gamma^2_\mu = 1, \tag{3.19}$$

we can can define a complete set of 16 spinor matrices, γ_n, for 4-vectors n such that $n_\mu \in \{0, 1\}$,

$$\gamma_n \equiv \prod_{\mu=0}^{3} (\gamma_\mu)^{n_\mu}. \tag{3.20}$$

The naive action is symmetric under the transformation

$$\psi(x) \to \tilde{\psi}(x) \equiv \gamma_5 \gamma_\mu (-1)^{\frac{x_\mu}{a}} \psi(x) = \gamma_5 \gamma_\mu e^{\left(\frac{i\pi x_\mu}{a}\right)} \psi(x), \tag{3.21}$$

which means that any low momentum mode, $\psi(x)$ has an equivalent mode with momentum $p_\mu \approx \frac{\pi}{a}$, a so called doubler. We can apply this transformation independently in different directions, giving the general transformation,

$$\begin{aligned} \psi(x) &\to \mathcal{B}_\zeta(x)\psi(x) \\ \bar{\psi}(x) &\to \bar{\psi}(x)\mathcal{B}^\dagger_\zeta(x) \\ \mathcal{B}_\zeta(x) &\equiv \gamma_{\bar{\zeta}} (-1)^{\frac{\zeta \cdot x}{a}} \\ &\propto \prod_\mu (\gamma_5 \gamma_\mu)^{\zeta_\mu} e^{\frac{i x \cdot \zeta \pi}{a}}, \end{aligned} \tag{3.22}$$

where $\zeta_\mu \in \{0, 1\}$ and

$$\bar{\zeta_\mu} \equiv \sum_{\nu \neq \mu} \zeta_\nu \text{ mod2}.$$

There are sixteen possible ζs, which provide fifteen doublers in addition to the 'real' $\zeta = (0,0,0,0)$ mode. These doublers manifest themselves as sixteen equivalent flavours of quark in the lattice. We call these flavours tastes to avoid confusion with the flavours we intend to create.

These additional tastes cause a headache for the would be naive action user, however, for identical tastes, a simple factor of $\frac{1}{16}$ in the power of $\mathrm{Det}(M)$ will remove the copies. This doesn't work in reality, due to taste exchange. The simplest manifestation of this is a quark which emits a gluon of momentum $p \approx \frac{\zeta\pi}{a}$, which is absorbed by another low momentum quark. Both quarks are not driven far off shell due to the doubling symmetry, and so each changes taste to another low energy quark. Such an exchange is suppressed by a^2, and is a source of discretisation error on the lattice.

Various solutions to the doubling problem have been tried. One solution was to give the doublers mass, which causes them to decouple as $a \to 0$. This lead to the Wilson quark action [42] and from that the Clover action [58]. Here we will discuss the method employed by the HISQ action.

3.3.2 Symanzik improvement at tree level

Firstly, we can correct the a^2 errors in the finite derivative using a so called Naik term [59] $\Delta_\mu \to \Delta_\mu - \frac{a^2}{6}\Delta_\mu^3$. Next, the HISQ action addresses the discretisation errors arising from leading order taste exchange. This is achieved by introducing a form factor $f_\mu(p)$ to the gluon quark vertex $\bar{\psi}\gamma_\mu U_\mu \psi$, which vanishes for gluons of momentum $p = \frac{\zeta\pi}{a}$ for each non-zero ζ, noting that in the case where $\zeta_\mu = 1$, f_μ need not be 0 as the original interaction vanishes anyway. We arrive at,

$$f_\mu(p) = \begin{cases} 1 & p \to 0 \\ 0 & p \to \frac{\zeta\pi}{a},\ \zeta^2 \neq 0,\ \zeta_\mu = 0, \end{cases} \tag{3.23}$$

which is achieved by replacing,

$$\begin{aligned} U_\mu(x) &\to \mathcal{F}_\mu U_\mu(x) \\ \mathcal{F}_\mu &\equiv \prod_{\nu\neq\mu} \left(1 + \frac{a^2\delta_\nu^{(2)}}{4}\right)\Bigg|_{\mathrm{symm}} \\ \delta_\nu^{(2)} U_\mu(x) &\equiv \frac{1}{a^2}\big(U_\nu(x)U_\mu(x+a\hat{\nu})U_\nu^\dagger(x+a\hat{\mu}) \\ +U_\nu^\dagger(x-a\hat{\nu})U_\mu(x-a\hat{\nu})U_\nu(x-a\hat{\nu}+a\hat{\mu}) &- 2U_\mu(x)\big), \end{aligned} \tag{3.24}$$

where $\mathcal{F}$ is the smearing operator, which vanishes as $\delta_\nu^{(2)} U_\mu \to -\frac{4}{a^2} U_\mu$, when $p_\sigma \to \frac{\pi}{a}$, as required. The 'symm' here indicates that the product in $\mathcal{F}_\mu$ is symmeterised over all possible orderings of the operators [60]. This smearing process introduces additional $\mathcal{O}(a^2)$ errors which are then removed by a further replacement,

$$\begin{aligned} \mathcal{F}_\mu^{\text{ASQTAD}} &= \mathcal{F}_\mu - \sum_{\nu \neq \mu} \frac{a^2 (\delta_\nu)^2}{4} \\ \delta_\nu U_\mu(x) &\equiv \frac{1}{a}\big(U_\nu(x) U_\mu(x + a\hat{\nu}) U_\nu^\dagger(x + a\hat{\mu}) \\ -U_\nu^\dagger(x - a\hat{\nu}) U_\mu(x - a\hat{\nu}) U_\nu(x - a\hat{\nu} + a\hat{\mu}). \end{aligned} \tag{3.25}$$

This procedure results in the celebrated ASQTAD action [61],

$$\begin{aligned} \sum_x \bar{\psi}(x) \left(\sum_\mu \gamma_\mu \Big(\Delta_\mu(V) - \frac{a^2}{6} \Delta_\mu^3(U) \Big) + m_0 \right) \psi(x) \\ V_\mu \equiv \mathcal{F}_\mu^{ASQTAD} U_\mu(x). \end{aligned} \tag{3.26}$$

3.3.3 Symanzik improvement at one loop

The ASQTAD action still suffers from taste exchange errors as smearing links introduces other unsmeared links in orthogonal directions. We counter this by using repeated smearing, in the form of the HISQ action [10],

$$\begin{aligned} \sum_x \bar{\psi}(x)(\not{D}^{\text{HISQ}} + m)\psi \\ D_\mu^{\text{HISQ}} &\equiv \Delta_\mu(W) - \frac{a^2}{6}(1 + \epsilon)\Delta_\mu^3(X) \\ W_\mu(x) &\equiv \mathcal{F}_\mu^{\text{HISQ}} U_\mu(x) \\ X_\mu(x) &\equiv \mathcal{U}\mathcal{F}_\mu U_\mu(x) \\ \mathcal{F}_\mu^{\text{HISQ}} &\equiv \left(\mathcal{F}_\mu - \sum_{\nu \neq \mu} \frac{a^2 (\delta_\nu)^2}{2} \right) \mathcal{U}\mathcal{F}_\mu, \end{aligned} \tag{3.27}$$

where $\mathcal{U}$ unitarises whatever it acts on and the calculation of the Naik ϵ is discussed below. This HISQ action is what will be used throughout this work and is corrected for taste exchanges through order $\alpha_s (am)^2$.

The Naik ϵ

In lattice units, the tree level pole mass m_{tree} in the HISQ action can be written in terms of the bare mass m

$$m_{\text{tree}}(m) = m\Big(1 - \frac{3}{80}m^4 + \frac{23}{2240}m^6 + \frac{1783}{537600}m^8 - \frac{76943}{23654400}m^{10} + \mathcal{O}(m^{12})\Big). \tag{3.28}$$

The Naik ϵ [62] is defined at tree level by demanding that the tree level pole mass is equal to the tree level kinetic mass [63],

$$\epsilon_{\text{tree}}(m) = \frac{4 - \sqrt{4 + \frac{12 m_{\text{tree}}}{\sinh(m_{\text{tree}})\cosh(m_{\text{tree}})}}}{\sinh^2(m_{\text{tree}})} - 1, \tag{3.29}$$

where we have suppressed the argument of $m_{\text{tree}}(m)$. In all the calculations described in this work ϵ_{tree} is defined to be identically zero for strange and light quarks.

3.3.4 Staggered quarks

As well as the smearing, we also stagger quarks [64]. This is purely to improve the speed of simulations and staggered quarks are formally equivalent to naive quarks.

Consider transforming the naive quark field

$$\psi(x) \to \Omega(x)\chi(x) \qquad \bar{\psi}(x) \to \bar{\chi}(x)\Omega^\dagger(x), \tag{3.30}$$

where

$$\Omega(x) \equiv \prod_{\mu=0}^{3} (\gamma_\mu)^{x_\mu}. \tag{3.31}$$

There are 16 different Ωs, with the following useful properties,

$$\Omega^\dagger(x)\Omega(x) = 1, \tag{3.32}$$

$$\Omega(x) = \gamma_n \quad \text{for} \quad n_\mu = x_\mu \text{mod} 2, \tag{3.33}$$

$$\alpha_\mu(x) \equiv \Omega^\dagger(x)\gamma_\mu\Omega(x \pm \hat{\mu}) = (-1)^{x_\mu^<}, \tag{3.34}$$

where $x_\mu^< \equiv \sum_{\nu=0}^{\mu-1} x_\nu$ and similarly for $x_\mu^>$. We also define $\bar{n}_\mu = (\sum_{\nu\neq\mu} n_\nu)\text{mod}2$. Putting this

together to form the staggered quark action,

$$\bar{\psi}(x)(\gamma \cdot \Delta + m)\psi(x) = \bar{\chi}(x)(\alpha(x).\Delta + m)\chi(x), \tag{3.35}$$

where Δ is the discretised covariant derivative used in the naive quark action (Equation (3.18)). This action is diagonal in spinor space, that is to say all components of χ are the same, so we only need one of them. The same principles apply to the HISQ action. We can write,

$$\langle \chi(x)\bar{\chi}(y)\rangle = g(x,y)\mathbf{1}_{\text{spinor}}, \tag{3.36}$$

where $g(x,y)$ is the one-spinor-component staggered quark propagator. Having only one spinor component to compute reduces the cost by a factor of 4. An additional effect of the staggering process that follows from this is to reduce the total number of tastes from 16 to 4.

3.3.5 Twisted boundary conditions

Our preferred method for imparting momentum into correlation functions is using twisted boundary conditions [65, 66]. We do this by inserting a factor of $e^{i\vec{p}\cdot\vec{x}}$ into the sum over space, where $p_j = 2\pi\theta_j/(aN_j)$, with twist θ and aN_j the total extent of the lattice in the j direction. Under the transformation,

$$\psi(x) \to e^{-i\vec{p}\cdot\vec{x}}\psi(x), \tag{3.37}$$

we create the boundary conditions [66, 67, 68],

$$\psi(x + aN_j\hat{j}) = e^{2\pi i\theta_j}\psi(x). \tag{3.38}$$

In order to achieve this, we in fact multiply the gauge links by e^{iap_j},

$$\bar{\psi}(x)e^{iap_j}U_j(x)\psi(x+a\hat{j}) = \bar{\psi}(x)e^{-i\vec{p}\cdot\vec{x}}U_j(x)e^{i\vec{p}\cdot(\vec{x}+a\hat{j})}\psi(x+a\hat{j}). \tag{3.39}$$

Twist dependent finite volume effects are very small for our work, as the values taken for θ do not correspond to lattice momenta which are integer multiples of $\frac{2\pi}{N_j}$ [69].

3.3.6 Quarks on the lattice

Repeating our gluon and quark action from above,

$$S = S_{\text{glue}} + \sum_x \bar{\psi}(\not{D} + m)\psi = S_{\text{glue}} + \bar{\psi}M\psi, \tag{3.40}$$

and recalling that $M = (\not{D} + m)$, we note that computers can't deal with the anti-commuting quark fields, so we have to do this bit of the integral by hand,

$$\begin{aligned} \mathcal{Z} &= \int \mathcal{D}U\mathcal{D}\bar{\psi}\mathcal{D}\psi e^{-(S_{\text{glue}}+\bar{\psi}M\psi)} \\ &= \int \mathcal{D}U\text{Det}(M)e^{-S_{\text{glue}}} \\ &= \int \mathcal{D}U e^{-S_{\text{glue}}+\ln\text{Det}(M)}. \end{aligned} \tag{3.41}$$

This $\text{Det}(M)$ term is the most costly part of any lattice calculation (see below for details), and is why people chose to work with quenched lattices for so long. Worse still, the cost grows as the quark mass shrinks. We need to add a $\text{Det}(M)$ for each flavour of quark on the lattice, but shortcuts can be made by setting $m_u = m_d$, and including only the u, d and s quarks in the sea. Light quarks are the main contributions to the sea for low energy processes, and we do not work on fine enough lattices to include b (let alone t) quarks in the sea, though modern actions (and the action used throughout this work) include the c.

For a typical lattice calculation, we will create a meson at time t_0 and destroy it at time $t_0 + T$. The combination of γs, Γ, in the operator used to create the meson (here we use generic 'heavy' meson H) can be any local operator with the correct spin and parity. Locality means that the state we will create will not be the ground state but will have overlap with all of the states of the desired meson. We deal with this when we come to fit the data, discussed below. Like the partition function above, the correlation function for the meson containing quarks a and b now

also needs to have the quarks integrated over by hand.

$$\begin{aligned}\langle 0|\, H^\dagger(T+t_0)H(t_0)\,|0\rangle = \\ &\frac{1}{\mathcal{Z}}\int \mathcal{D}U\mathcal{D}\bar{\psi}\mathcal{D}\psi \sum_{\vec{x},\vec{x}_0} \bar{\psi}^a(T+t_0,\vec{x})\Gamma\psi^b(T+t_0,\vec{x})\bar{\psi}^b(t_0,\vec{x}_0)\Gamma\psi^a(t_0,\vec{x}_0)e^{-(S_{\text{glue}}+\bar{\psi}M\psi)} \\ &= \frac{1}{\mathcal{Z}}\int \mathcal{D}U \text{Tr}_{\text{spin,colour},\vec{x}}(M_a^{-1}\Gamma M_b^{-1}\Gamma)e^{-S_{\text{glue}}+\ln \text{Det}(M)}.\end{aligned} \tag{3.42}$$

Once we have done the costly process of generating the configurations, involving the $\text{Det}(M)$ term, the inversion of the massive matrix M is the next most expensive part of the calculation.

Calculating M^{-1} and Det(M)

The inversion of the matrix M is, as mentioned, very computationally expensive. Because M is large and sparse, it lends itself to a conjugate gradient algorithm [70], a well known method for solving systems of linear equations. Consider an equation,

$$M^\dagger M x = \xi \quad \Longrightarrow \quad x = (M^\dagger M)^{-1}\xi \tag{3.43}$$

where we seek to invert $M^\dagger M$ to ensure an Hermitian and positive definite matrix[2]. ξ here is the propagator source which will be discussed more in Section 4.2.2. Solving this equation for x is equivalent to minimising,

$$f(x) = \frac{1}{2}x^T M^\dagger M x - \xi^T x. \tag{3.44}$$

The conjugate gradient solves this equation by descending in the direction of steepest gradi-

[2]Applying an $M^\dagger$ to $(M^\dagger M)^{-1}$ recovers M^{-1}.

ent,

$$\begin{aligned} x_{n+1} &= x_n + \alpha_n d_n \\ \alpha_n &= \frac{r_n^T r_n}{d_n^T M^\dagger M d_n^T} \\ d_0 &= r_0 \\ d_{n\neq 0} &= r_n + \beta_n d_{n-1} \\ r_n &= \xi - M^\dagger M x_n \\ \beta_n &= \frac{r_n^T r_n}{r_{n-1}^T r_{n-1}}. \end{aligned} \tag{3.45}$$

After initiating a random x_0 and calculating the gradient $\nabla f(x) = M^\dagger M x_0 - \xi = -r_0$, one chooses the step direction d_n such that it is orthogonal with respect to $M^\dagger M$ ($d_n^T M^\dagger M d_{m\neq n} = 0$). The residue, r_n tells us how close we are to the solution, so we choose some tolerance in our calculation based on this, which acts as a stopping point. The cost of this inversion scales with quark mass as m^{-2}, which is why quarks are also often given artificially heavy masses to make this process cheaper, and then the results of the lattice calculation are extrapolated to the physical masses.

Calculating $\mathrm{Det}(M)$ is even more expensive, because of the non-local nature of the determinant. In modern simulations, M can easily have dimension of 10^8. The evaluation is carried out by making use of our $M^\dagger M$ inversion, with the so called Φ algorithm [71]. Heuristically, we can introduce an artificial scalar field Φ and add a term to the action,

$$\mathrm{Det}(M^\dagger M) = \int \mathcal{D}\Phi^\dagger \mathcal{D}\Phi e^{-\Phi^\dagger (M^\dagger M)^{-1}\Phi}, \tag{3.46}$$

which is also evaluated by Monte Carlo sampling. The problem with this is that it creates two factors of M for each flavour of quark, but this can be resolved by taking the root of the determinant[3].

[3] In fact we take the 8th root, as staggered quarks have four degenerate tastes, which we shall discuss below.

3.4 HISQ ensembles

All of the work in this thesis will use some combination of the eight gluon field ensembles listed in Table 3.1. In lattice QCD, the lattice spacing a is set by choosing a value for the coupling g_s, in the form of $\beta = \frac{6}{g_s^2}$. We do not directly select a spacing. Once β has been chosen, we determine a by calculating the Wilson flow parameter w_0 [72] (actually we calculate w_0/a). This can be straightforwardly determined on the lattice by taking the Wilson flow [73] and smearing and integrating until a dimensionless observable reaches a chosen value (see [72]). Once we have this defined, we must match to the continuum value in order to determine the spacing. We do this using the mass ratios of the Ω baryon, pion and kaon (see [72, 74]). In our case, we use the continuum value $w_0 = 0.1715(9)$fm [75].

All of the ensembles in Table 3.1 are generated by the MILC Collaboration with an improved gluon action and HISQ sea quarks [76, 77, 78]. The ensembles are $N_f = 2+1+1$, which indicates that there are two degenerate light quarks l in the sea, as well as strange and charm quarks. Systematic uncertainties from the missing b and t sea quarks are small when working at low energies, such as this work, though the sea b will become more important for achieving high precision calculations at the physical valence b mass in future.

Whilst the strange and charm quark masses are approximately physical on each ensemble, the degenerate light quarks take one of two values. In ensembles which we shall label with a '5', they are taken to be 1/5 the mass of the strange quark on that ensemble, whilst on ensembles described as physical ('phys'), the light quarks will take a value which is close to the physical average of u and d, which is 1/27.18(10) the strange mass [79].

The valence s and c masses we use are not exactly the same as the sea masses (see Table 3.1) as we are able to tune them to get them closer to their physical values, as is done in [80]. We further correct for any remaining mistuning in our final calculation (see Chapters 5, 6 and 7).

Table 3.1: Gluon ensembles used in this work. In column 3, β is the same one defined in Section 3.2.2 and is discussed further in the text. The Wilson flow parameter, $w_0 = 0.1715(9)$fm (also discussed in the text), is determined in [75], and is used to calculate the lattice spacing a via values for w_0/a [72] in column 4. Column 5 gives the spatial (N_x) and temporal (N_t) dimensions of each lattice in lattice units whilst columns 6-10 give the masses of the valence and sea quarks, noting that these are the same in the case of the light quark.

Set	label	β	w_0/a	$N_x^3 \times N_t$	$am_l^{\text{sea/val}}$	am_s^{sea}	am_c^{sea}	am_s^{val}	am_c^{val}
1	v. coarse phys	5.8	1.1367(5)	$32^3 \times 48$	0.00235	0.0647	0.831	0.0678	0.8605
2	coarse phys	6.0	1.4149(6)	$48^3 \times 64$	0.00184	0.0507	0.628	0.0527	0.643
3	fine phys	6.3	1.9518(7)	$64^3 \times 96$	0.00120	0.0363	0.432	0.036	0.433
4	v. coarse 5	5.8	1.1119(10)	$16^3 \times 48$	0.013	0.065	0.838	0.0705	0.888
5	coarse 5	6.0	1.3826(11)	$24^3 \times 64$	0.0102	0.0509	0.635	0.0545	0.664
6	fine 5	6.3	1.9006(20)	$32^3 \times 96$	0.0074	0.037	0.440	0.0376	0.449
7	superfine 5	6.72	2.896(6)	$48^3 \times 144$	0.0048	0.024	0.286	0.0234	0.274
8	ultrafine 5	7.0	3.892(12)	$64^3 \times 192$	0.00316	0.0158	0.188	0.0165	0.194

Chapter 4

Methods for lattice calculations of form factors

In this chapter we will shed light on the core methods which will be integral to the form factor calculations discussed in Chapters 5, 6 and 7: the calculation of correlation functions on the lattice, their subsequent fitting to extract the ground state masses and three-point amplitudes, the construction of form factors from these results, and finally the heavy-HISQ method which enables us to perform fully relativistic HISQ calculations up to the physical b mass.

4.1 A note on units

The lattice spacing a is a worthy adversary to anyone who wishes to write about the lattice without making a mess. In what follows, we shall abuse notation somewhat. The vast majority of the time, unless physical units (GeV) are expressly stated, we will be working in lattice units and suppressing a; that is to say energies E and masses M will actually be aE, aM, and so on. More importantly, times t will refer to integers $0 < t/a < N_t$, and similarly for spatial dimensions. However, to avoid endless as, t will (very occasionally) also refer to the time in physical units. As most equations are on the lattice, this distinction should be obvious and often irrelevant, and we will reinstate a for clarity when it is important. Now we've done a bit of housekeeping, let's begin.

4.2 Calculating correlation functions

Correlation functions are the bread and butter of lattice QCD calculations. Fundamentally, the simplest object that we might wish to study is the two-point correlation function, as defined in Equation (2.41) for $n = 2$. That is to say, we want to create a meson by acting on the vacuum with creation operators at some point in spacetime on our lattice, $x_0 = (t_0, \vec{x}_0)$, leave it to potter along on its own for a bit, and then destroy it with annihilation operators at some later time and place $x_1 = (t_1, \vec{x}_1)$. With different t_1 values, and averaging over all possible spatial start and end locations, we end up mapping out a function $C_2(t_0, t_1)$ which tells us something about the probability of the meson surviving to that particular time. We have shown in Equation (2.43) that this correlation function is a sum over exponential decays, with the meson ground state decaying most slowly. Naturally, we demand that $t_0 < t_1$. Using the notation of Equattion 2.41,

$$C_2(t_0, t_1) = \sum_{\vec{x}_0, \vec{x}_1} \langle \Omega | \, \varphi_H(x_0) \varphi_H(x_1) \, | \Omega \rangle \, . \tag{4.1}$$

Generalising this to an n-point function $C_n(t_0, t_1, ..., t_{n-1})$, we can add other destinations to our particle's itinerary besides the start and end point. Perhaps our particle has an interaction with another particle at time t_1 in position $\vec{x}_1$, and then another interaction at some later time and place. In this work, we will only concern ourselves with two- and three-point functions. The former, as outlined above, describes a particle minding its own business, whilst the latter allows for one interaction en route from creation to annihilation. Again, using the notation of Equattion 2.41,

$$C_3(t_0, t_1, t_2) = \sum_{\vec{x}_0, \vec{x}_1, \vec{x}_2} \langle \Omega | \, \varphi_H(x_0) \varphi_H(x_1) \varphi_H(x_2) \, | \Omega \rangle \, . \tag{4.2}$$

In general, we will use the notation $x_0 = (t_0, \vec{x}_0)$ for the origin of our particle. In the case of two-point functions, the particle will be annihilated at $x_t = (t_0 + t, \vec{x}_t)$, giving correlation function $C_2((t_0 + t) - t_0)$ which will only depend on t, the difference between the origin and the terminus. For the three-point case, we instead annihilate by convention at $x_T = (t_0 + T, \vec{x}_T)$, again using x_0 as the origin, and this time using x_t to denote the position of some interaction in the form of a current insertion. Thus we have a three-point function $C_3(t_0, t, T)$, which is only meaningful if our particle interacts after it is created and before it is destroyed, so $t_0 < t < T$[1]. For the purposes of the vast majority of the following discussion, it is simplest to just assume $t_0 = 0$. This is not true in reality, as we typically increase our statistics by using several different t_0

[1] This is not strictly true, since our lattice has periodic boundary conditions.

values on each gauge field configuration. Because our lattice has periodic boundary conditions, we can always just rotate our loop of lattice back around to $t_0 = 0$, effectively just relabelling all the points. We need to worry about t_0 for our discussion of phases below, but otherwise it is safe to assume $t_0 = 0$ unless explicitly stated.

4.2.1 Phases in staggered quarks

The following is based on the appendices on staggered quarks [64] in [10]. Here we will repeat some of the necessary formulae from Section 3.3 and derive the phases which are relevant to the spin-taste combinations used in this work.

Firstly, let's recall some useful formulae from Chapter 3. We take our Euclidean space hermitian gamma matrices, ($\gamma_\mu^\dagger = \gamma_\mu,\ \gamma_\mu^2 = 1,\ \{\gamma_\mu, \gamma_\nu\} = 2\delta_{\mu\nu}$) and define γ_n such that,

$$\gamma_n \equiv \prod_{\mu=0}^{3} (\gamma_\mu)^{n_\mu}. \tag{4.3}$$

Using $\gamma_5 = \gamma_0\gamma_1\gamma_2\gamma_3$ means $\gamma_5 = \gamma_{n=(1,1,1,1)}$.[2] Taking our naive quark $\psi(x)$ and antiquark $\bar{\psi}(x)$, we stagger them with the local field transformation,

$$\psi(x) \to \Omega(x)\chi(x) \qquad \bar{\psi}(x) \to \bar{\chi}(x)\Omega^\dagger(x), \tag{4.4}$$

where

$$\Omega(x) \equiv \prod_{\mu=0}^{3} (\gamma_\mu)^{x_\mu} = \gamma_{n=(x)\mathrm{mod}2}. \tag{4.5}$$

There are 16 different Ωs, with the following useful properties,

$$\Omega^\dagger(x)\Omega(x) = 1, \tag{4.6}$$

$$\alpha_\mu(x) \equiv \Omega^\dagger(x)\gamma_\mu\Omega(x \pm \hat{\mu}) = (-1)^{x_\mu^<}, \tag{4.7}$$

[2] Depending on your outlook, there may be an overall phase of -1 here. In reality the MILC code version 7.1.1 (http://www.physics.utah.edu/ detar/milc/milcv7.html) which we use, takes γ_4 in place of γ_0 which negates this. We'll turn a blind eye to this irrelevant overall phase in the interests of using the more familiar γ_0.

where $x_\mu^< \equiv \sum_{\nu=0}^{\mu-1} x_\nu$. The staggered quark action is diagonal in spinor space: we can write,

$$\langle \chi(x)\bar{\chi}(y)\rangle = g(x,y)\mathbf{1}_{\text{spinor}}, \tag{4.8}$$

where $g(x,y)$ is the one-spinor-component staggered quark propagator. To return to naive quarks, we can simply reinstate the Ωs and write the Feynman propagator,

$$S_F(x,y) \equiv \langle \psi(x)\bar{\psi}(y)\rangle = g(x,y)\Omega(x)\Omega^\dagger(y), \tag{4.9}$$

where $S_F(x,y) = \langle \psi(x)\bar{\psi}(y)\rangle$ is the propagator for a quark moving from y to x, and $\gamma_5 S_F^\dagger(x,y)\gamma_5 = \langle \psi(y)\bar{\psi}(x)\rangle$ the propagation of an antiquark from y to x. The additional γ_5s are a consequence of the γ_5 hermitian symmetry of the action, and act to reverse the arguments in S_F.

The basic method for our lattice calculations will be to:

- Construct the correlation functions we are interested in with naive quarks
- Convert these to combinations of staggered quark propagators and lattice site dependent phases
- Calculate the staggered quark propagators on each configuration by inverting the Dirac matrix M off source ξ
- Tie together the propagators into the relevant combinations
- Average the resulting correlation functions over the N_{cf} configurations

Spin-taste

Because of the doubling symmetry addressed in Section 3.3.1, each naive quark operator leads to 16 different operators, labelled by the 16 possible vectors s consisting only 1s and 0s. These tastes are defined exactly the same way as the gamma matrix spin structure, but we shall use ξ_s in place of γ_n to distinguish the two. A spin-taste structure might be defined $\gamma_n \otimes \xi_s$ and the most important relation we will need is,

$$\gamma_n \otimes \xi_s \psi(x) = (-1)^{\bar{s}.x}\gamma_s^\dagger \gamma_n \psi(x \oplus (n+s)), \tag{4.10}$$

where $\oplus$ adds $n+s$ to x modulo the hypercube x lies in and $\bar{s}_\mu = (\sum_{\nu\neq\mu} s_\nu)\text{mod}2$. If we label the hypercube x lies in x_B then $x_{B\mu}\text{mod}2 = 0$ and,

$$(x \oplus (n+s))_\mu \equiv x_{B\mu} + (x_\mu - x_{B\mu} + (n+s)_\mu)\text{mod}2. \tag{4.11}$$

As $x_\mu - x_{B\mu} \in \{0,1\}$, provided $(n+s)_\mu \neq 1$ (which means $s_\mu = n_\mu$) we will be adding 0 or 2 to this, which is removed by the mod, leaving us with just x_μ. The situation if $s_\mu \neq n_\mu$ is more complicated. In this case $(s+n)_\mu = 1$. If x_μ is even, such that $x_\mu - x_{B\mu} = 0$, the result is to add 1 in the μ direction. In the case of odd x_μ, $x_\mu - x_{B\mu} = 1$ and we are left with $x_{B\mu}$ which is equal to $x_\mu - 1$.

$$(x \oplus (n+s))_\mu = \begin{cases} x_\mu & (s_\mu = n_\mu) \\ x_\mu + 1 & (s_\mu \neq n_\mu, x_\mu\text{mod}2 = 0) \\ x_\mu - 1 & (s_\mu \neq n_\mu, x_\mu\text{mod}2 = 1) \end{cases} \tag{4.12}$$

It should be noted that in Equation (4.10) we have suppressed a link operator. For example, if we end up with $\psi(x+\hat{\mu})$, where $\hat{\mu}$ is a unit vector in the μ direction, we in fact mean $U_\mu(x)\psi(x+\hat{\mu})$, similarly for $U_\mu^\dagger(x-\hat{\mu})\psi(x-\hat{\mu})$. These operators will differ with the cases above so it is simpler to implicitly include them in ψ. All this means that in the case where $s = n$ and we have the same spin and taste, sandwiching such a spin-taste combination between two fermions will result in a local operator (with both fermions at the same point x). Cases where $s \neq n$ however create non-local or 'point-split' operators. In this case the fermions are separated by a number of gauge links, which require link operators to be inserted. In general point-split operators lead to noisier correlators so it is best to avoid them where possible.

The MILC code we use [81, 82] takes inputs in the form of spin taste combinations and implements them as phases which depend on lattice site. Choosing the spin-taste combinations is no great secret. Firstly, the spin part of the operator needs to marry up with the physical mesons or currents we want to perform the calculation with. Secondly, the unphysical taste elements should be self consistent so as to cancel, or the corresponding correlator will vanish, and thirdly, we seek to minimise the use of point-split operators. Below we will calculate the phases which will be relevant to us, there's nothing fantastic to see here, but it's instructive to go through the motions. We note that we will not bother to keep track of overall phases - these can easily be sorted out at the end by correcting the sign of the resulting correlation function.

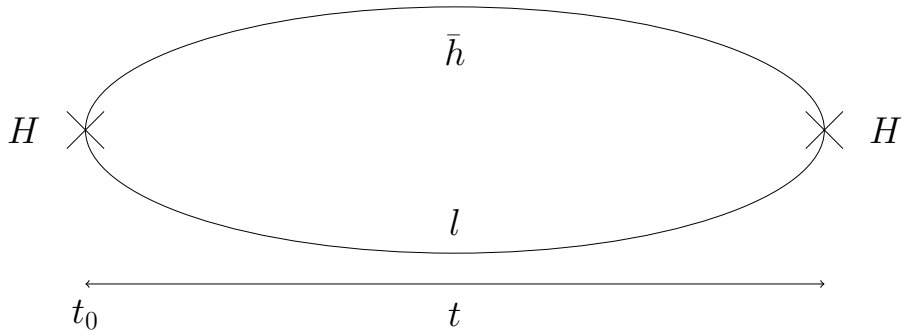

Figure 4.1: Schematic of a two-point correlation function.

Two-point correlation functions

Figure 4.1 is a schematic representation of a $H = \bar{h}l$ two-point function. We have a random colour wall source on timeslice t_0, where we initiate the heavy and light propagators, and apply the relevant spin-taste operator to one of them. We then propagate these quarks to the sink, at timeslice $t_0 + t$, where they are destroyed. Further discussion of sources and random walls will follow in Section 4.2.2.

In this work we will require three spin-taste combinations of pseudoscalar two-point function. We shall consider the case of a meson consisting a heavy anti-quark $\bar{h}$ and a light quark l, though these can easily be relabelled as any quarks of our choosing. The first case we will look at is for the meson with current $J_5^5(x_t) = \bar{\psi}_h(x_t)\gamma_5 \otimes \xi_5\psi_l(x_t)$ which annihilates a Goldstone meson $\bar{h}l$ at x_t defined above. Using Equation (4.10) and the fact that for γ_5, $n = s = \bar{s} = (1,1,1,1)$ we can write,

$$\begin{aligned}\bar{\psi}_h(x_t)\gamma_5 \otimes \xi_5\psi_l(x_t) &= \bar{\psi}_h(x_t)(-1)^{x_t^0+x_t^1+x_t^2+x_t^3}\gamma_5^\dagger\gamma_5\psi_l(x_t), \\ &= (-1)^{\tilde{x}_t}\bar{\psi}_h(x_t)\psi_l(x_t),\end{aligned} \tag{4.13}$$

where we use upstairs notation on x_t for clarity; $x_t = (x_t^0, x_t^1, x_t^2, x_t^3)$, and we abbreviate the sum of all x components $\sum_\mu x^\mu = \tilde{x}$. In order to construct a two-point correlation function we will

also need to create the same meson at x_0, using the creation operator,

$$\begin{aligned} J_5^{5\dagger}(x_0) &= (-1)^{\tilde{x}_0}(\bar{\psi}_h(x_0)\psi_l(x_0))^\dagger, \\ &= (-1)^{\tilde{x}_0}(\psi_h^\dagger(x_0)\gamma_0\psi_l(x_0))^\dagger, \\ &= (-1)^{\tilde{x}_0}\psi_l^\dagger(x_0)\gamma_0^\dagger\psi_h(x_0), \\ &= (-1)^{\tilde{x}_0}\bar{\psi}_l(x_0)\psi_h(x_0). \end{aligned} \tag{4.14}$$

Putting this together into a two-point function, we follow a simple recipe,

$$\langle J_5^5(x_t)J_5^{5\dagger}(x_0)\rangle = \langle 0|\, T\{J_5^5(x_t)J_5^\dagger(x_0)\}\,|0\rangle = (-1)^{\tilde{x}_0+\tilde{x}_t}\,\langle 0|\, T\{\bar{\psi}_h(x_t)\psi_l(x_t)\bar{\psi}_l(x_0)\psi_h(x_0)\}\,|0\rangle\,. \tag{4.15}$$

Once we have the expression for our time ordered fields, we can take the relevant Wick contractions that correspond to the Feynman diagrams for two-point function we are interested in. In our case, we simply contract the $\bar{h}$ and l fields with themselves,

$$\begin{aligned} \langle J_5^5(x_t)J_5^{5\dagger}(x_0)\rangle &= (-1)^{\tilde{x}_0+\tilde{x}_t}\,\langle 0|\,\overbrace{\bar{\psi}_h(x_t)\psi_l(x_t)\bar{\psi}_l(x_0)\psi_h(x_0)}\,|0\rangle\,, \\ &= (-1)^{\tilde{x}_0+\tilde{x}_t}\,\langle 0|\,\overbrace{\psi_h(x_0)\bar{\psi}_h(x_t)}\overbrace{\psi_l(x_t)\bar{\psi}_l(x_0)}\,|0\rangle\,, \end{aligned} \tag{4.16}$$

where in the second line we ignore the overall phases resulting from the exchange of fermions[3]. These contractions are exactly the propagators defined in Equation (4.9) above for an anti-quark $\bar{h}$ moving from x_0 to x_t and a quark l propagating from x_0 to x_t,

$$\begin{aligned} \langle J_5^5(x_t)J_5^{5\dagger}(x_0)\rangle &= (-1)^{\tilde{x}_0+\tilde{x}_t}\mathrm{Tr}\big(\gamma_5 S_F^{h\dagger}(x_t,x_0)\gamma_5 S_F^l(x_t,x_0)\big), \\ &= (-1)^{\tilde{x}_0+\tilde{x}_t}\mathrm{Tr}\big(\gamma_5\Omega(x_0)\Omega^\dagger(x_t)\mathrm{tr}(g^{h\dagger}(x_t,x_0))\gamma_5\mathrm{tr}(g^l(x_t,x_0))\Omega(x_t)\Omega^\dagger(x_0)\big), \\ &= (-1)^{\tilde{x}_0+\tilde{x}_t}\mathrm{Tr}\big(\gamma_5\Omega(x_0)\Omega^\dagger(x_t)\gamma_5\Omega(x_t)\Omega^\dagger(x_0)\big)\mathrm{tr}(g^{h\dagger}(x_t,x_0)g^l(x_t,x_0)), \end{aligned} \tag{4.17}$$

where we have used Equation (4.9) in the first step and tr is the trace over colour. The trace over spinor components (Tr) arises because we have suppressed spinor indices thus far. Reinstating them we would have something like $\gamma_5^{\alpha\beta}S_F^{\dagger\beta\rho}\gamma_5^{\rho\delta}S_F^{\delta\alpha}$, giving the trace. Recalling the definition of Ω in Equation (4.5) and the fact that $\gamma_\mu\gamma_5 = -\gamma_5\gamma_\mu$, we can see that $\gamma_5\Omega(x) = (-1)^{\tilde{x}}\Omega(x)\gamma_5$ and

[3] We go through this first case in painstaking detail, but will skip steps in later examples.

similarly for $\gamma_5\Omega^\dagger(x)$, such that

$$\begin{aligned}\langle J_5^5(x_t)J_5^{5\dagger}(x_0)\rangle &= (-1)^{\tilde{x}_0+\tilde{x}_t}(-1)^{\tilde{x}_0+\tilde{x}_t}\mathrm{Tr}\big(\gamma_5\gamma_5\Omega(x_0)\Omega^\dagger(x_t)\Omega(x_t)\Omega^\dagger(x_0)\big)\mathrm{tr}(g^{h\dagger}(x_t,x_0)g^l(x_t,x_0)),\\ &= 4\mathrm{tr}(g^{h\dagger}(x_t,x_0)g^l(x_t,x_0)),\end{aligned} \tag{4.18}$$

where the 4 comes from the spinor trace over the identity. So the $\gamma_5 \otimes \xi_5$ Goldstone spin-taste combination requires no phase to be added to the staggered quark propagators.

We will now work though the other two spin-taste combinations we require in this work. The first is the local $J_{5t}^{5t}(x_t) = \bar{\psi}_h(x_t)\gamma_5\gamma_0 \otimes \xi_5\xi_0\psi_l(x_t)$, where we will require the additional result that $\gamma_m \otimes \xi_r\gamma_n \otimes \xi_s = \gamma_m\gamma_n \otimes \xi_r\xi_s$. For γ_0, $s = n = (1,0,0,0)$ and $\bar{s} = \bar{n} = (0,1,1,1)$[4]

$$\begin{aligned}J_{5t}^{5t}(x_t) &= \bar{\psi}_h(x_t)\gamma_5\gamma_0 \otimes \xi_5\xi_0\psi_l(x_t),\\ &= \bar{\psi}_h(x_t)\gamma_5 \otimes \xi_5\gamma_0 \otimes \xi_0\psi_l(x_t),\\ &= \bar{\psi}_h(x_t)(-1)^{\tilde{x}_t}(-1)^{\tilde{x}_t - x_t^0}\psi_l(x_t),\\ &= (-1)^{x_t^0}\bar{\psi}_h(x_t)\psi_l(x_t),\\ \implies J_{5t}^{5t\dagger}(x_0) &= (-1)^{x_0^0}\bar{\psi}_l(x_0)\psi_h(x_0).\end{aligned} \tag{4.19}$$

$$\begin{aligned}\langle J_{5t}^{5t}(x_t)J_{5t}^{5t\dagger}(x_0)\rangle &= (-1)^{x_0^0+x_t^0}\langle 0|\,\overbrace{\psi_h(x_0)\bar{\psi}_h}(x_t)\overbrace{\psi_l(x_t)\bar{\psi}_l}(x_0)\,|0\rangle\,,\\ &= (-1)^{x_0^0+x_t^0}(-1)^{\tilde{x}_0+\tilde{x}_t}4\mathrm{tr}(g^{h\dagger}(x_t,x_0)g^l(x_t,x_0)),\\ &= (-1)^{\tilde{x}_0+\tilde{x}_t+x_0^0+x_t^0}4\mathrm{tr}(g^{h\dagger}(x_t,x_0)g^l(x_t,x_0)).\end{aligned} \tag{4.20}$$

So the correlator from this local spin taste operator has a phase which is dependent on the spatial components of x_0 and x_t. Finally, we also require the point split operator $J_5^{5x}(x_t) = \bar{\psi}_h(x_t)\gamma_5 \otimes \xi_5\xi_1\psi_l(x_t)$. In this case we have spin matrix $\mathbb{1}$ for which $n = \bar{n} = (0,0,0,0)$ and ξ_1, for which $s = (0,1,0,0)$ and $\bar{s} = (1,0,1,1)$, which causes the point splitting. We require the additional relation $\gamma_n \otimes \xi_r\xi_s = (-1)^{\bar{r}\cdot s}\gamma_n \otimes \xi_s\xi_r$. In this case $r = \bar{r} = (1,1,1,1)$, giving an overall

[4]We can alternatively use $\gamma_5\gamma_0 = -\gamma_{(0,1,1,1)}$ directly, giving $\bar{s} = \bar{n} = (1,0,0,0)$ and thus the same result.

phase of -1 from this swap. We're not interested in overall phases, so we shall drop this[5].

$$\begin{aligned} J_5^{5x}(x_t) &= \bar{\psi}_h(x_t)\gamma_5 \otimes \xi_1\xi_5\psi_l(x_t), \\ &= \bar{\psi}_h(x_t)\mathbb{1} \otimes \xi_1\gamma_5 \otimes \xi_5\psi_l(x_t), \\ &= (-1)^{\tilde{x}_t}(-1)^{\tilde{x}_t - x_t^1}\bar{\psi}_h(x_t)\gamma_1^\dagger\psi_l(x_t \oplus (0,1,0,0)), \\ &= (-1)^{x_t^1}\bar{\psi}_h(x_t)\gamma_1^\dagger\psi_l(x_t \oplus \hat{1}), \\ \implies J_5^{5x\dagger}(x_0) &= (-1)^{x_0^1}\bar{\psi}_l(x_0 \oplus \hat{1})\gamma_1\psi_h(x_0), \end{aligned} \tag{4.21}$$

where the $U_1(x_t)$ link operators are implicit in the x_t^1 direction, which is required to join the two fermions, and $\hat{1}$ is a unit vector in the same direction. Note that U acts on colour space and so does not act on γ. Also note that there should be another overall minus in the hermitian conjugate from the swapping of γ_0 and γ_1, which we again ignore.

$$\begin{aligned} \langle J_5^{5x}(x_t)J_5^{5x\dagger}(x_0)\rangle &= (-1)^{x_0^1+x_t^1}\langle 0|\overbrace{\psi_h(x_0)\bar{\psi}_h}(x_t)\gamma_1^\dagger\overbrace{\psi_l(x_t\oplus\hat{1})\bar{\psi}_l}(x_0\oplus\hat{1})\gamma_1|0\rangle, \\ &= (-1)^{x_0^1+x_t^1}\mathrm{Tr}\big(\gamma_5\Omega(x_0)\Omega^\dagger(x_t)\gamma_5\gamma_1^\dagger\Omega(x_t\oplus\hat{1})\Omega^\dagger(x_0\oplus\hat{1})\gamma_1\big), \\ &\times \mathrm{tr}(g^{h\dagger}(x_t,x_0)g^l(x_t\oplus\hat{1},x_0\oplus\hat{1})), \\ &= (-1)^{x_0^1+x_t^1}(-1)^{\tilde{x}_0+\tilde{x}_t}(-1)^{x_t^0}(-1)^{x_0^0}\mathrm{Tr}\big(\Omega(x_0)\Omega^\dagger(x_t)\Omega(x_t)\Omega^\dagger(x_0)\big), \\ &\times \mathrm{tr}(g^{h\dagger}(x_t,x_0)g^l(x_t\oplus\hat{1},x_0\oplus\hat{1})), \\ &= (-1)^{x_0^2+x_t^2+x_0^3+x_t^3}4\mathrm{tr}(g^{h\dagger}(x_t,x_0)g^l(x_t\oplus\hat{1},x_0\oplus\hat{1})), \end{aligned} \tag{4.22}$$

where we have used the fact that from the definition of $\Omega(x)$, $\Omega(x\oplus\hat{1}) = (-1)^{x^0}\gamma_1\Omega(x)$ and similarly $\Omega^\dagger(x\oplus\hat{1}) = (-1)^{x^0}\Omega^\dagger(x)\gamma_1^\dagger$ (see Equation (4.7)). We see that the phases follow a similar pattern to the previous case, with components of x that correspond to the γs which are absent from the current. The point splitting is clear (recalling that the gauge links are suppressed).

Three-point correlation functions

Figure 4.2 is a schematic representation of a $K \to H$ three-point function with a current insertion J. We have a random colour wall source on timeslice t_0, where we initiate the strange and light propagators, and apply the relevant spin-taste operator to one of them (see 4.2.2). We then propagate the strange quark to the current, at timeslice $t_0 + t$. The light quark is propagated

[5] In any case, we could easily have defined these the other way around in the current.

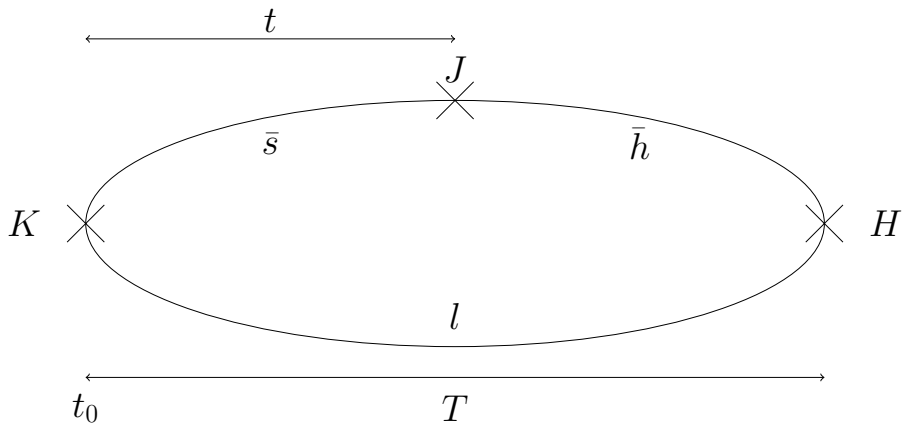

Figure 4.2: Schematic of a three-point correlation function.

to timeslice $t_0 + T$, where it has a sink phase added and acts as a source for the heavy quark h which is propagated back to meet the current at $t_0 + t$. This set up is the reverse of what happens in reality, but is set up this way because it involves calculating the smallest number of expensive strange propagators. The physics remains unchanged under this time reversal. Again, further discussion of sources will follow in Section 4.2.2.

In this work we will use three different current insertions, S, V and T which will be combined with the appropriate spin-taste combination of meson from the three outlined above. We shall label the quarks as in Figure 4.2, that is an $\bar{s}$ and $\bar{h}$ each paired with an l. As before, it is simple to amend these labels to cover the specific cases in the chapters below. Again we will work through the first case carefully and the latter two more quickly.

The scalar current insertion $S(x_t) = \bar{\psi}_s(x_t)\psi_h(x_t)$ annihilates an $\bar{s}$ and creates an $\bar{h}$ at x_t. We'll

label the meson operators (formerly J) as H and K to distinguish them.

$$\begin{aligned}
\langle H_5^5(x_T)S(x_t)K_5^{5\dagger}(x_0)\rangle &= (-1)^{\tilde{x}_T+\tilde{x}_0}\,\langle 0|\,\overbrace{\bar{\psi}_h(x_T)\psi_l(x_T)\bar{\psi}_s(x_t)\psi_h(x_t)\bar{\psi}_l(x_0)\psi_s(x_0)}\,|0\rangle\,,\\
&= (-1)^{\tilde{x}_T+\tilde{x}_0}\,\langle 0|\,\overbrace{\psi_h(x_t)\bar{\psi}_h(x_T)}\overbrace{\psi_l(x_T)\bar{\psi}_l(x_0)}\overbrace{\psi_s(x_0)\bar{\psi}_s(x_t)}\,|0\rangle\,,\\
&= (-1)^{\tilde{x}_T+\tilde{x}_0}\mathrm{Tr}\big(\gamma_5 S_F^{h\dagger}(x_T,x_t)\gamma_5 S_F^{l}(x_T,x_0)\gamma_5 S_F^{s\dagger}(x_t,x_0)\gamma_5\big),\\
&= (-1)^{\tilde{x}_T+\tilde{x}_0}\mathrm{Tr}\big(\gamma_5\Omega(x_t)\Omega^\dagger(x_T)\gamma_5\Omega(x_T)\Omega^\dagger(x_0)\gamma_5\Omega(x_0)\Omega^\dagger(x_t)\gamma_5\big),\\
&\times \mathrm{tr}(g^{h\dagger}(x_T,x_t)g^{l}(x_T,x_0)g^{s\dagger}(x_t,x_0)),\\
&= (-1)^{\tilde{x}_T+\tilde{x}_0}(-1)^{\tilde{x}_T+\tilde{x}_0}\mathrm{Tr}\big(\Omega(x_t)\Omega^\dagger(x_T)\Omega(x_T)\Omega^\dagger(x_0)\Omega(x_0)\Omega^\dagger(x_t)\big),\\
&\times \mathrm{tr}(g^{h\dagger}(x_T,x_t)g^{l}(x_T,x_0)g^{s\dagger}(x_t,x_0)),\\
&= 4\mathrm{tr}(g^{h\dagger}(x_T,x_t)g^{l}(x_T,x_0)g^{s\dagger}(x_t,x_0)),
\end{aligned} \tag{4.23}$$

where we have used the fact that we can move spinors through one another by picking up an overall phase which we ignore. We find no position dependent phase in the case of the scalar current insertion.

The temporal vector current insertion $V(x_t) = \bar{\psi}_s(x_t)\gamma_0 \otimes \xi_0\psi_h(x_t)$ proceeds similarly (noting $s = n = (1,0,0,0)$, $\bar{s} = \bar{n} = (0,1,1,1)$),

$$\begin{aligned}
V(x_t) &= \bar{\psi}_s(x_t)\gamma_0 \otimes \xi_0\psi_h(x_t),\\
&= (-1)^{\tilde{x}_t - x_t^0}\bar{\psi}_s(x_t)\psi_h(x_t),
\end{aligned} \tag{4.24}$$

$$\begin{aligned}
\langle H_{5t}^{5t}(x_T)V(x_t)K_5^{5\dagger}(x_0)\rangle &= (-1)^{x_T^0}(-1)^{\tilde{x}_t-x_t^0}(-1)^{\tilde{x}_0}\,\langle 0|\,\overbrace{\psi_h(x_t)\bar{\psi}_h}(x_T)\overbrace{\psi_l(x_T)\bar{\psi}_l}(x_0)\overbrace{\psi_s(x_0)\bar{\psi}_s}(x_t)\,|0\rangle\,,\\
&= (-1)^{x_T^0}(-1)^{\tilde{x}_t-x_t^0}(-1)^{\tilde{x}_0}(-1)^{\tilde{x}_T+\tilde{x}_0}4\mathrm{tr}(g^{h\dagger}(x_T,x_t)g^{l}(x_T,x_0)g^{s\dagger}(x_t,x_0)),\\
&= (-1)^{\tilde{x}_t-x_t^0+\tilde{x}_T-x_T^0}4\mathrm{tr}(g^{h\dagger}(x_T,x_t)g^{l}(x_T,x_0)g^{s\dagger}(x_t,x_0)).
\end{aligned} \tag{4.25}$$

Finally, the tensor current insertion $T(x_t) = \bar{\psi}_s(x_t)\gamma_0\gamma_1 \otimes \xi_0\xi_1\psi_h(x_t)$ which uses the point split K from above,

$$\begin{aligned}
T(x_t) &= \bar{\psi}_s(x_t)\gamma_0\gamma_1 \otimes \xi_0\xi_1\psi_h(x_t),\\
&= \bar{\psi}_s(x_t)\gamma_0 \otimes \xi_0\gamma_1 \otimes \xi_1\psi_h(x_t),\\
&= (-1)^{x_t^0+x_t^1}\bar{\psi}_s(x_t)\psi_h(x_t),
\end{aligned} \tag{4.26}$$

$$
\begin{aligned}
\langle H_{5t}^{5t}(x_T)T(x_t)K_5^{5x\dagger}(x_0)\rangle &= (-1)^{x_T^0}(-1)^{x_t^0+x_t^1}(-1)^{x_0^1}\,\langle 0|\,\overbrace{\bar\psi_h(x_T)\psi_l(x_T)\bar\psi_s(x_t)\psi_h(x_t)\bar\psi_l(x_0\oplus\hat 1)\gamma_1\psi_s(x_0)}\,|0\rangle\,,\\
&= (-1)^{x_T^0}(-1)^{x_t^0+x_t^1}(-1)^{x_0^1}\,\langle 0|\,\overbrace{\psi_h(x_t)\bar\psi_h(x_T)}\overbrace{\psi_l(x_T)\bar\psi_l(x_0\oplus\hat 1)\gamma_1}\overbrace{\psi_s(x_0)\bar\psi_s(x_t)}\,|0\rangle\,,\\
&= (-1)^{x_t^0+x_t^1+x_0^1+x_T^0}\mathrm{Tr}\big(\gamma_5\Omega(x_t)\Omega^\dagger(x_T)\gamma_5\Omega(x_T)\Omega^\dagger(x_0\oplus\hat 1)\gamma_1\gamma_5\Omega(x_0)\Omega^\dagger(x_t)\gamma_5\big),\\
&\times\,\mathrm{tr}(g^{h\dagger}(x_T,x_t)g^l(x_T,x_0\oplus\hat 1)g^{s\dagger}(x_t,x_0)),\\
&= (-1)^{x_t^0+x_t^1+x_0^1+x_T^0+x_0^0}\mathrm{Tr}\big(\Omega(x_t)\Omega^\dagger(x_T)\gamma_5\Omega(x_T)\Omega^\dagger(x_0)\gamma_5\Omega(x_0)\Omega^\dagger(x_t)\big),\\
&\times\,\mathrm{tr}(g^{h\dagger}(x_T,x_t)g^l(x_T,x_0\oplus\hat 1)g^{s\dagger}(x_t,x_0)),\\
&= (-1)^{x_t^0+x_t^1+x_0^1+x_T^0+x_0^0+\tilde x_0+\tilde x_T}\mathrm{Tr}\big(\Omega(x_t)\Omega^\dagger(x_T)\Omega(x_T)\Omega^\dagger(x_0)\Omega(x_0)\Omega^\dagger(x_t)\big),\\
&\times\,\mathrm{tr}(g^{h\dagger}(x_T,x_t)g^l(x_T,x_0\oplus\hat 1)g^{s\dagger}(x_t,x_0)),\\
&= (-1)^{x_t^0+x_t^1+x_0^2+x_0^3+x_T^0+\tilde x_T}4\mathrm{tr}(g^{h\dagger}(x_T,x_t)g^l(x_T,x_0\oplus\hat 1)g^{s\dagger}(x_t,x_0)).
\end{aligned}
\tag{4.27}
$$

In the second line, we have to be careful about moving spinors because of the γ_1. Dropping some labels and reinstating spinor indices we have $\bar\psi_h^\alpha\psi_l^\alpha\bar\psi_s^\beta\psi_h^\beta\bar\psi_l^\delta\gamma_1^{\delta\sigma}\psi_s^\sigma$, which can be rearranged as $\psi_h^\beta\bar\psi_h^\alpha\psi_l^\alpha\bar\psi_l^\delta\gamma_1^{\delta\sigma}\psi_s^\sigma\bar\psi_s^\beta$, leading to the second line.

4.2.2 Sources and random walls

We will now elaborate somewhat on the construction of the two- and three-point functions as outlined at the beginning of the subsections above. From the procedures above, we end up with an expression for the two-point correlation functions of the generic form

$$C_2(x_t,x_0) = (-1)^{\alpha(x_0,x_t)}\mathrm{tr}(g^{h\dagger}(x_t,x_0)g^l(x_t,x_0)), \tag{4.28}$$

where $(-1)^{\alpha(x_0,x_t)}$ is some phase peculiar to the spin-taste combination we chose. We are usually looking for a meson in a momentum eigenstate with 3-momentum $\vec p$, so we Fourier transform, putting the momentum on one quark (in this case $\bar h$) using the twisted boundary conditions described in Section 3.3.5 and recalling $x_0^0 = t_0$, $x_t^0 = t_0 + t$,

$$\tilde C_2^{\vec p}(t_0,t_0+t) = \frac{1}{N_x^3}\sum_{\vec x_t,\vec x_0} e^{-i\vec p\cdot(\vec x_0-\vec x_t)}C_2(x_t,x_0). \tag{4.29}$$

In order to calculate this, we need to create staggered propagators $g(x_t,x_0)$ for each x_0 and x_t choice on our lattice, leading to a cost which goes like $2\mathrm{vol}^2$. The cost of the calculation can be reduced using random wall sources, at the expense of introducing random noise. We first define

new propagators,

$$\begin{aligned} P^{l,\vec{p}}(x_t,t_0) &\equiv \frac{1}{\sqrt{N_x^3}}\sum_{\vec{x}_0} e^{i\vec{p}\cdot(\vec{x}_0-\vec{x}_t)}\xi(\vec{x}_0)g^l(x_t,x_0), \\ P^{h,\vec{p}\dagger}(x_t,t_0) &\equiv \frac{1}{\sqrt{N_x^3}}\sum_{\vec{x}_0} e^{-i\vec{p}\cdot(\vec{x}_0-\vec{x}_t)}(-1)^{\alpha(x_0,0)}g^{h\dagger}(x_t,x_0)\xi^\dagger(\vec{x}_0), \end{aligned} \tag{4.30}$$

where we note that the source phase, like the momentum, is only applied to one of the propagators. Here $\xi(\vec{x_0})$ [83, 84] is a random field of complex colour vectors across the spatial components of the source timeslice $x_0^0 = t_0$, chosen differently for each gauge field configuration. It has the property

$$\langle f(\vec{x},\vec{x}')\xi^\dagger(\vec{x}')\xi(\vec{x})\rangle = \delta_{\vec{x}'\vec{x}}\langle f(\vec{x},\vec{x}')\rangle. \tag{4.31}$$

From this we can see that the construction,

$$\begin{aligned} \zeta &= \sum_{\vec{x}_t}(-1)^{\alpha(0,x_t)}\mathrm{tr}(P^{h,\vec{p}\dagger}(x_t,t_0)P^{l,\vec{0}}(x_t,t_0)), \\ &= \frac{1}{N_x^3}\sum_{\vec{x}_t}(-1)^{\alpha(0,x_t)}\mathrm{tr}\Big(\sum_{\vec{x}_0'} e^{-i\vec{p}\cdot(\vec{x}_0'-\vec{x}_t)}(-1)^{\alpha(x_0',0)}g^{h\dagger}(x_t,x_0')\xi^\dagger(\vec{x}_0')\sum_{\vec{x}_0}\xi(\vec{x}_0)g^l(x_t,x_0)\Big), \\ &= \frac{1}{N_x^3}\sum_{\vec{x}_t}(-1)^{\alpha(0,x_t)}\sum_{\vec{x}_0',\vec{x}_0} e^{-i\vec{p}\cdot(\vec{x}_0'-\vec{x}_t)}(-1)^{\alpha(x_0',0)}\mathrm{tr}(g^{h\dagger}(x_t,x_0')\xi^\dagger(\vec{x}_0')\xi(\vec{x}_0)g^l(x_t,x_0)), \\ &= \frac{1}{N_x^3}\sum_{\vec{x}_t,\vec{x}_0}(-1)^{\alpha(x_0,x_t)}e^{-i\vec{p}\cdot(\vec{x}_0-\vec{x}_t)}\mathrm{tr}(g^{h\dagger}(x_t,x_0)g^l(x_t,x_0)), \\ &= \tilde{C}_2^{\vec{p}}(t_0,t_0+t), \end{aligned} \tag{4.32}$$

where we give momentum (where necessary) to the heavy quark and note that $(-1)^{\alpha(x,y)} = (-1)^{\alpha(x,0)}(-1)^{\alpha(0,y)}$. Note that we define the dummy variable $x_0' = (x_0^0, \vec{x}_0')$, i.e. the first component is not a dummy variable as it doesn't appear in the sum. This ζ is exactly the process we described above. We have generated quark and antiquark propagators on random colour walls, applied a source spin-taste phase and momentum to one of them ($\bar{h}$ in this case) and then destroyed them at the sink, applying the sink phase here. Averaging over the random wall effectively ties the propagators together at the source.

This approach is considerably cheaper than calculating the all-to-all propagator $g(x,y)$ in Equation (4.28)). To compute such an all-to-all propagator would require us to solve a system

like

$$\begin{aligned} g(x^0,\vec{x},y^0,\vec{y}) &= \sum_{z^0,\vec{z}} D^{-1}(x^0,\vec{x},z^0,\vec{z})\delta^3(\vec{y}-\vec{z})\delta(y^0-z^0) \\ &= \sum_{\vec{z}} D^{-1}(x^0,\vec{x},y^0,\vec{z})\delta^3(\vec{y}-\vec{z}) \\ &\Longrightarrow \\ g_{ac}(x^0,y^0) &= D^{-1}_{ab}(x^0,y^0)\delta_{bc}, \end{aligned} \tag{4.33}$$

where in the last line we have given each spatial point an index, i.e. $a = 0$ is $(0,0,0)$, $a = 1$ is $(0,0,1)$ and so on. We can see that for each of the N_x^3 choices of index c, we have to explicitly invert the matrix D (using the inversion process of Section 3.3.6) off the source vector where the only non-zero element is the 1 in position c. However, by introducing the random wall, we only need compute the propagator,

$$\begin{aligned} P(x^0,\vec{x},y^0) &= \sum_{\vec{y}} \Big(\sum_{z^0,\vec{z}} D^{-1}(x^0,\vec{x},z^0,\vec{z})\delta^3(\vec{y}-\vec{z})\delta(y^0-z^0) \Big) \xi(\vec{y}) \\ &= \sum_{\vec{z}} D^{-1}(x^0,\vec{x},y^0,\vec{z})\xi(\vec{z}) \\ &\Longrightarrow \\ P_a(x^0,y^0) &= D^{-1}_{ab}(x^0,y^0)\xi_b, \end{aligned} \tag{4.34}$$

where the random wall vector ξ_b means that we effectively combine all of the N_x^3 inversions in Equation (4.33) into one single inversion, at the cost of introducing random noise.

In the case of three-point functions we proceed similarly to the two-points, with the $\bar{s}$ propagator being given a source phase and momentum, and the l not. We use the l propagator as a source to generate the $\bar{h}$ one, a so called extended source.

$$C_3(x_T,x_t,x_0) = (-1)^{\alpha(x_0,x_t,x_T)}\mathrm{tr}(g^{h\dagger}(x_t,x_T)g^{l}(x_T,x_0)g^{s\dagger}(x_t,x_0)), \tag{4.35}$$

$$\tilde{C}_3^{\vec{p}}(t_0,t_0+t,t_0+T) = \frac{1}{N_x^3}\sum_{\vec{x}_t,\vec{x}_0,\vec{x}_T} e^{-i\vec{p}\cdot(\vec{x}_0-\vec{x}_t)} C_3(x_T,x_t,x_0), \tag{4.36}$$

where we apply momentum only to the $\bar{s}$ propagator, which means x_T doesn't feature in the

Fourier transform.

$$\begin{aligned} P^{l,\vec{p}}(x_T,t_0) &\equiv \frac{1}{\sqrt{N_x^3}}\sum_{\vec{x}_0} e^{i\vec{p}\cdot(\vec{x}_0-\vec{x}_T)}\xi(\vec{x}_0)g^l(x_T,x_0) \\ P^{s,\vec{p}\dagger}(x_t,t_0) &\equiv \frac{1}{\sqrt{N_x^3}}\sum_{\vec{x}_0} e^{-i\vec{p}\cdot(\vec{x}_0-\vec{x}_t)}(-1)^{\alpha(x_0,0,0)}g^{s\dagger}(x_t,x_0)\xi^\dagger(\vec{x}_0) \\ P_{\text{ext}}^{h,\vec{p}_h,\vec{p}_l\dagger}(x_t,t_0,t_0+T) &\equiv \sum_{\vec{x}_T} e^{-i\vec{p}_h\cdot(\vec{x}_t-\vec{x}_T)}(-1)^{\alpha(0,0,x_T)}g^{h\dagger}(x_t,x_T)P^{l,\vec{p}_l}(x_T,t_0). \end{aligned} \tag{4.37}$$

$$\begin{aligned} \zeta =& \sum_{\vec{x}_t}(-1)^{\alpha(0,x_t,0)}\text{tr}(P_{\text{ext}}^{h,\vec{0},\vec{0}\dagger}(x_t,t_0,t_0+T)P^{s,\vec{p}\dagger}(x_t,t_0)) \\ =& \frac{1}{N_x^3}\sum_{\vec{x}_t}(-1)^{\alpha(0,x_t,0)}\text{tr}\Big(\sum_{\vec{x}_T}(-1)^{\alpha(0,0,x_T)}g^{h\dagger}(x_t,x_T)\times \\ & \sum_{\vec{x}_0',\vec{x}_0}\xi(\vec{x}_0')g^l(x_T,x_0')e^{-i\vec{p}\cdot(\vec{x}_0-\vec{x}_t)}(-1)^{\alpha(x_0,0,0)}g^{s\dagger}(x_t,x_0)\xi^\dagger(\vec{x}_0)\Big) \\ =& \frac{1}{N_x^3}\sum_{\vec{x}_t,\vec{x}_0,\vec{x}_T}(-1)^{\alpha(x_0,x_t,x_T)}e^{-i\vec{p}\cdot(\vec{x}_0-\vec{x}_t)}\text{tr}\Big(g^{h\dagger}(x_t,x_T)g^l(x_T,x_0)g^{s\dagger}(x_t,x_0)\Big) \\ =& \tilde{C}_3^{\vec{p}}(t_0,t_0+t,t_0+T). \end{aligned} \tag{4.38}$$

Herein lies the reason that we set up the three-point calculation backwards, as in Figure 4.2. Assuming we have already generated propagators for the $\bar{h}l$ and $\bar{s}l$ two-points, we can reuse these in the three-point case. However, one set of propagators ($\bar{h}$ or $\bar{s}$) must be regenerated off the extended source l. In this case, the heavier $\bar{h}$ propagators are much cheaper to generate than the lighter $\bar{s}$, so it makes sense to set up the calculation in this way.

4.3 Fitting correlation functions

4.3.1 Fit forms

Once we have our first principles QCD determinations of two- and three-point correlation functions, we can fit them to known forms (Equations (2.43) and (2.44)) in order to extract physics parameters like masses and amplitudes. The general form of a two-point function for a meson M of ground state energy E_0^M is an exponential decay, with some amplitude A_0^M which is related to the matrix element for the decay. However, the operators we construct on the lattice overlap not just with the ground state, but with an infinite tower of excited states of energy E_i^M $\forall i > 0$ of the

meson. We must also account for the periodic boundary conditions we impose, which means we also get mesons effectively travelling backwards from the end of the lattice N_t. Finally, because of doubling symmetry, staggered quark operators for pseudoscalar mesons couple not only to 0^- states as intended, but also to 0^+ mesons, with a phase which oscillates in t (see appendix of [10]). Some mesons are exempt from this behaviour, notably those with the same quark and anti-quark mass and twist (i.e. which use the same propagator). In our case this applies to the η_s with zero twist in Chapter 6. Putting all this together into a fit function,

$$C_2^M(t) = \sum_{i=0}^{N_{\text{exp}}} \left(|A_i^{M,\text{n}}|^2 (e^{-E_i^{M,\text{n}} t} + e^{-E_i^{M,\text{n}}(N_t - t)}) - (-1)^t |A_i^{M,\text{o}}|^2 (e^{-E_i^{M,\text{o}} t} + e^{-E_i^{M,\text{o}}(N_t - t)}) \right), \tag{4.39}$$

where we label (non)oscillating states (n)o and we truncate the sum to N_{exp} terms, because computers find it difficult to count to infinity. This truncation is made possible by discarding some number of data points (for $t < t_{\text{min}}$) from the start of each correlator. As very heavy states decay away rapidly in t this enables us to fit to a finite number of exponentials.

Fitting three-point functions follows directly from considering two such two-point functions, decaying from $t_0 = 0$ and T respectively and meeting at t, where we insert a current J (Equation (2.44)). The amplitude from the overlap between the different (non)oscillating mesons is given by parameter J, where were will use $J = S(V)[T]$ to denote scalar (vector) [tensor] current insertions. In this case, we do not need to include 'reflections' from the ends of the lattice, as generally $T << N_t$ and the interactions between such terms are negligible[6]. For a decay of meson M_1 to meson M_2

$$\begin{aligned} C_3^{M_1,M_2}(t,T) = \sum_{i,j=0}^{N_{\text{exp}}} \big(& A_i^{M_1,\text{n}} J_{ij}^{\text{nn}} A_j^{M_2,\text{n}} e^{-E_i^{M_1,\text{n}} t} e^{-E_j^{M_2,\text{n}}(T-t)} \\ & - (-1)^{(T-t)} A_i^{M_1,\text{n}} J_{ij}^{\text{no}} A_j^{M_2,\text{o}} e^{-E_i^{M_1,\text{n}} t} e^{-E_j^{M_2,\text{o}}(T-t)} \\ & - (-1)^t A_i^{M_1,\text{o}} J_{ij}^{\text{on}} A_j^{M_2,\text{n}} e^{-E_i^{M_1,\text{o}} t} e^{-E_j^{M_2,\text{n}}(T-t)} \\ & + (-1)^T A_i^{M_1,\text{o}} J_{ij}^{\text{oo}} A_j^{M_2,\text{o}} e^{-E_i^{M_1,\text{o}} t} e^{-E_j^{M_2,\text{o}}(T-t)} \big). \end{aligned} \tag{4.40}$$

[6] In an extremely high statistics calculation with precision far beyond what is achievable today, such effects may need to be included. However, at that stage, extending N_t to further suppress these effects would probably be a simpler option anyway.

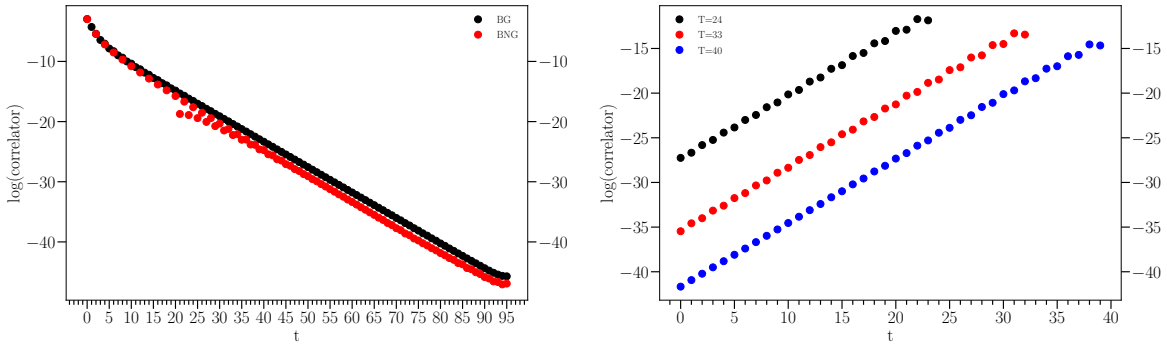

Figure 4.3: Example of two- and three-point correlator data on the ultrafine 5 (set 8) ensemble. In both cases error bars are too small to be visible. The left hand plot is the log of the (folded) two-point data for a heavy-light H meson with $am_h = 0.194$ (actually a D meson). 'BG' is a Goldstone pseudoscalar with spin-taste $\gamma_5 \otimes \xi_5$ and 'BNG' a non-Goldstone pseudoscalar $(\gamma_5\gamma_0 \otimes \xi_5\xi_0)$. Oscillations are much larger in the latter case, causing some negative values which are ignored by the log. On the right is the log of a three-point function for a $H \to K$ decay with a scalar current insertion, for $am_h = 0.6$ and twist $\theta = 1.529$. Three different T values are shown.

Both C_2 and C_3 are included as functions in the fitting package we use [85, 86]. It's worth noting that, because of the periodic boundary conditions, $C_2(t)$ is symmetric about $N_t/2$. The fitter takes advantage of this by folding the data before fitting. Figure 4.3 shows an example of the data for a typical two- and three-point function. We have taken the log and it is clear that the behaviour is very well described by an exponential fit form (the two-point data has been folded, and the three-point is plotted only up T). Note that in the two point function at small t values, oscillations cause negative values of the correlator on alternating timeslices, which are omitted from the log plot. Note also that three-point function, which is for a $H \to K$ decay with a scalar current insertion, is in fact arranged $K \to H$, as described above. This is visible in the plot because the heavier H meson decays faster and oscillates more than the lighter K. Hence the three-point is larger and has more oscillations towards $t = T$.

4.3.2 Bayesian fitting

Our correlator fits are carried out using the corrfitter package [86], which bolts onto the lsqfit package [85] to enable fitting of correlators as discussed above. Both packages make extensive use of gvar [87] which automatically keeps track of (Gaussian) uncertainties and correlations between variables throughout the fitting process. The packages use a Bayesian fitting procedure [88]

outlined in [89].

Because Equations (4.39) and (4.40) can have a very large number of unknown fit parameters A_i and E_i, it's easy to end up in a situation with more parameters than data, leading to infinite uncertainties on parameters. Bayesian fitting requires an a priori estimate of each of the fit parameters, the prior. Put simply, the prior ($\mathcal{P}$) is a conservative estimate of what the parameter might be, containing information which is not obtained from the data itself. If the fit does not have enough data with which to inform a certain parameter, then this parameter's posterior will simply be the prior it was given initially, preventing posteriors with infinite uncertainty from messing up the fit. This is incorporated into standard χ^2 least squares regression using a modification of χ^2 which includes the prior. For the example of $C_2(t)$ above, and dropping superscripts on A_i and E_i for simplicity, the χ^2 we minimise is [89]

$$\begin{aligned}\chi^2(A_n, E_n) = &\sum_{t,t'} \Delta C_2(t) \sigma^{-2}_{t,t'} \Delta C_2(t') \\ &+ \sum_i \frac{(A_i - \mathcal{P}[A_i])^2}{\sigma^2_{\mathcal{P}[A_i]}} + \sum_i \frac{(E_i - \mathcal{P}[E_i])^2}{\sigma^2_{\mathcal{P}[E_i]}},\end{aligned} \tag{4.41}$$

with

$$\Delta C_2 = \overline{C_2^{\text{data}}(t)} - C_2^{\text{fit}}(t; A_i, E_i), \tag{4.42}$$

where $\overline{C_2^{\text{data}}(t)}$ is the mean of our Monte Carlo values, and $C_2^{\text{fit}}(t; A_i, E_i)$ is the fit function that we are trying to minimise χ^2 with respect to.

$$\sigma^2_{t,t'} = \overline{C_2^{\text{data}}(t) C_2^{\text{data}}(t')} - \overline{C_2^{\text{data}}(t)}\ \overline{C_2^{\text{data}}(t')} \tag{4.43}$$

is the covarience matrix which is determined in the Monte Carlo averaging process over the number of configurations. The second and third terms in Equation (4.41) add χ^2 terms which are dependent on the prior means ($\mathcal{P}[A_i]$, $\mathcal{P}[E_i]$) and their respective standard deviations (σ), and so penalise fit parameters which did not reside within their priors. Typically, we quote χ^2/d.o.f. values, where the d.o.f.s (degrees of freedom) are the number of data points, minus the number of fit parameters. In Bayesian fitting, priors are counted like data points, so these cancel out the fit parameters, meaning the number of d.o.f.s is just the number of data points. A good fit should have χ^2/d.o.f. ≈ 1, but, as will be explained below, this is modified by priors and SVD cuts.

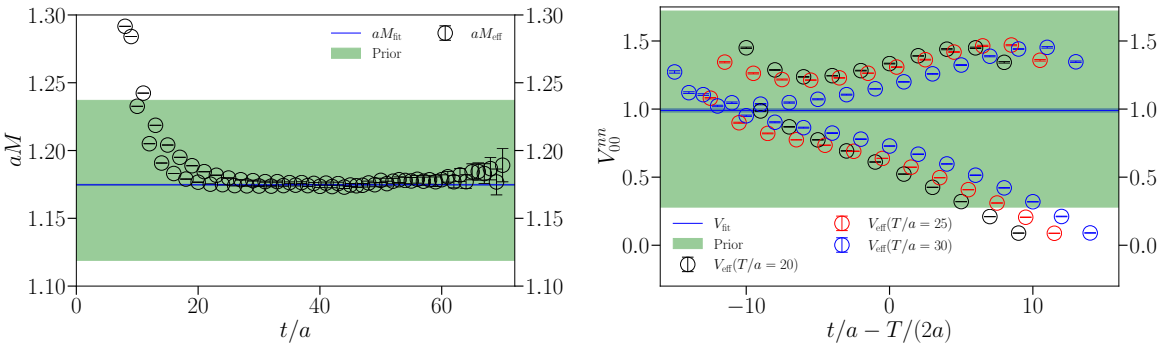

Figure 4.4: Example [2] of estimating priors for the ground state mass, and V_{00}^{nn} using Equations 4.44 and 4.46. Priors in green are based on the data points, whilst the fit result (including error band) is in blue.

4.3.3 Prior selection

Priors should be selected to include as much information as possible to help with the fit. One way to get estimates for priors is to use the effective mass, and effective amplitude, which can be obtained from Equation (4.39) in the limit of large t,

$$M_{\text{eff}} = \frac{1}{2}\cosh^{-1}\left(\frac{C_2(t-2)+C_2(t+2)}{2C_2(t)}\right), \tag{4.44}$$

$$A_{\text{eff}} = \sqrt{\frac{C_2(t)}{e^{-M_{\text{eff}}t}+e^{-M_{\text{eff}}(N_t-t)}}}. \tag{4.45}$$

These functions can be plotted and the plateau identified, which can then be used to inform the central value of the ground state energy and amplitude E_0 and A_0. A similar principle can be used to obtain an estimate for J_{00}^{nn}[7],

$$J_{\text{eff}}(t,T) = \frac{C_3^{M_1,M_2}(t,T)A_{\text{eff}}^{M_1}A_{\text{eff}}^{M_2}}{C_2^{M_1}(t)C_2^{M_2}(T-t)}. \tag{4.46}$$

An example of the effectiveness of these methods for estimating priors from [2] is shown in Figure 4.4. Here the ground state mass and V_{00}^{nn} are shown. In each case, the data points depict the result of Equations (4.44) and (4.46) respectively, with the green band the prior based off

[7]Note that it would also be legitimate to use the effective mass instead of the C_2 data. In practice we often find the results from this method to be less stable.

this (including a generous uncertainty), and the blue band (which includes uncertainty) the final fit result using this prior.

For non-ground state priors we often don't have much information beyond the fact that it ought to have a 'sensible' value, usually taking a number based off an effective calculation for the ground state.

Additional to the actual values of priors, we usually include further information. In all the work described in this thesis, we enforce $A_i > 0$ by using a log prior; that is we actually provide a prior for $\log(A_i)$. We know A_i is positive because the source and sink use the same operator in the cases we discuss. Similarly, for the energy of excited states, we actually provide a prior for $\log(E_{i+1} - E_i)$, which ensures that they are ordered with a positive energy gap, as would would expect.

Finally, in Chapters 5 and 7 ([1, 3]) we use the relativistic dispersion relation to relate the priors of the mesons at non-zero momentum, $\vec{p}$ to those at zero momentum. Allowing for discretisation effects which appear at order $|a\vec{p}|^2$ we can write (reinstating a for clarity),

$$\mathcal{P}[aE_0^{\vec{p}}] = \sqrt{(\mathcal{P}[aE_0^{\vec{0}}])^2 + |a\vec{p}|^2}\Big(1 + \tilde{C}\Big(\frac{|a\vec{p}|}{\pi}\Big)^2\Big), \tag{4.47}$$

$$\mathcal{P}[A_0^{\vec{p}}] = \frac{\mathcal{P}[A_0^{\vec{0}}]}{(1 + (|\vec{p}|/\mathcal{P}[E_0^{\vec{0}}])^2)^{1/4}}\Big(1 + \tilde{D}\Big(\frac{|a\vec{p}|}{\pi}\Big)^2\Big), \tag{4.48}$$

where $\tilde{C}$ and $\tilde{D}$ are order unity coefficients included in the fit to parameterise the discretisation effects. We include this extra information in the prior.

4.3.4 SVD cuts

This discussion is largely based on appendix D of [90] from which we adopt our notation. Owing to the large cost of achieving high statistics on most HISQ ensembles, we often find ourselves working in a regime where we have more data points N_G than samples N_s. For example, take the relatively small $D \to K$ calculation [1] (details in Chapter 5), where we only have one heavy mass, and no tensor form factor. In this case, on the coarse 5 (set 5) ensemble, where we have highest statistics, we have seven two-point correlators of length ≈ 30 (our data is folded and t_{min} points removed) and 40 (four T values times five twists times two current insertions) three-point correlators of length ≈ 15. This gives $N_G \approx 800$, which is close to our $N_s = 1053$ (total number

of configurations)[8].

When we defined χ^2 in Equation (4.41) we glossed over the exact role that the covariance matrix has to play. Consider a sample of N_s values of an (N_G dimensional) vector of data $\mathbf{G}^{(s)}$, with mean over samples $\overline{\mathbf{G}}$ and a matching fit function $\mathbf{G}(\mathbf{p})$, which takes parameters $\mathbf{p}$ as arguments and contains in its N_G elements a corresponding fit value for each data point in $\mathbf{G}^{(s)}$. The covariance matrix is given by,

$$\mathbf{M}_{\text{cov}} \approx \frac{1}{N_s(N_s-1)} \sum_s (\mathbf{G}^{(s)} - \overline{\mathbf{G}})(\mathbf{G}^{(s)} - \overline{\mathbf{G}})^T, \tag{4.49}$$

which can also be defined $\mathbf{M}_{\text{cov}} \equiv \mathbf{D}\mathbf{M}_{\text{corr}}\mathbf{D}$ in terms of the correlation matrix and $D_{ij} = \delta_{ij}\sigma_{G_i}$, the diagonal matrix of standard deviations.

Ignoring the part of χ^2 pertaining to the prior, we can now write,

$$\chi^2(\mathbf{p}) = \sum_{n=1}^{N_G} \frac{((\overline{\mathbf{G}} - \mathbf{G}(\mathbf{p}))^T \mathbf{D}^{-1}\mathbf{v}_n)^2}{\lambda_n}, \tag{4.50}$$

where λ_n and $\mathbf{v}_n$ are the eigenvalues and eigenvectors of $\mathbf{M}_{\text{corr}}$ respectively.

We can see from Equation (4.49) that if $N_s < N_G$ then there will be $N_G - N_s$ eigenvalues which are necessarily zero[9]. These cause very obvious problems for Equation (4.50). What is not so obvious is that when N_s is not sufficiently many times larger than N_G (10-100 times), the smallest eigenvalues are underestimated by Equation (4.49) [91], which again causes problems for our fit, making these parameters more heavily weighted than they ought to be. The eigenvalues for which this approximation affects χ^2 are those where [90],

$$\frac{\lambda_n^{\text{approx}}}{\lambda_n^{\text{exact}}} < 1 - \sqrt{\frac{2}{N_G}}, \tag{4.51}$$

since on average, individual terms in χ^2 should contribute approximately $1 \pm \sqrt{2/N_G}$ to the total.

[8] Technically folding the data for the two-points could be interpreted as doubling their statistics, but the point stands. It's also worth noting that, for our data, the latter half of the correlator is so highly correlated to the first half that folding it produces only a very small reduction in uncertainty versus just throwing away the second half of the data; it is in no way equivalent to a doubling of statistics.

[9] Each element of the sum is an $N_G \times N_G$ matrix with all rows linearly dependent. If we add N_s such matrices, we make N_s linearly independent rows, leaving $N_G - N_s$ eigenvalues 0.

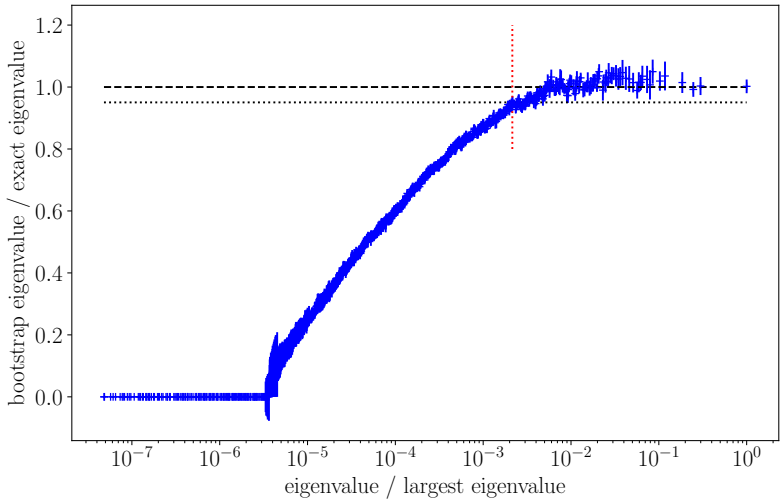

Figure 4.5: SVD diagnosis plot for the $D \to K$ coarse 5 (set 5) fit described in the text. The horizontal black dotted line indicates $1 - \sqrt{\frac{2}{N_G}}$, whilst the vertical red dotted lines gives the $\kappa = \lambda_n / \lambda_n^{\max}$ value at which this line is crossed.

The solution to this problem is two-fold. Firstly, we compare our approximated eigenvalues from Equation (4.49) to the 'exact' corresponding eigenvalues. Since we don't usually know the exact eigenvalues (otherwise this discussion would all be a bit pointless), we can generate them using bootstrapped copies of the data, and compare $\lambda_n^{\text{bstrap}}$ to $\lambda_n^{\text{approx}}$ to generate a plot like that in Figure 4.5 which is a good approximation for $\frac{\lambda_n^{\text{approx}}}{\lambda_n^{\text{exact}}}$ (see [90]). This plot, with eigenvalues on the x axis scaled by the largest value $\lambda_n^{\max}$ shows the value of $\kappa = \lambda_n / \lambda_n^{\max}$ [91, 92] at which the approximate eigenvalue drops below $1 - \sqrt{\frac{2}{N_G}}$ of its true value (horizontal dotted line). This point (vertical dotted line) is where we make our Singular Value Decomposition (SVD) cut. The second part of the process is simply to replace all eigenvalues below this point with this value, that is to say $\kappa \lambda_n^{\max}$. All of this process is handled by the corrfitter and lsqfit packages [85, 86] and more details and examples can be found in the documentation.

Some people view SVD cuts as being a little bit heathen, but we would like to make clear that the SVD cut is essential to fitting data faithfully, and is a conservative move. Increasing the SVD cut serves to increase the errors on the final fit results to reflect the weakness in our initial eigenvalue approximation.

4.3.5 Goodness of fit

As we've discussed above, our Bayesian fitting procedure seeks to minimise χ^2 (Equation 4.41), taking into account our prior expectations for fit parameters. Following this method, we expect a good fit to be one for which χ^2 per degree of freedom (d.o.f.) is approximately unity. This means that the average deviation squared of the data from the best fit line is equal to the variance of the data, for each fit parameter.

The parameter Q (sometimes called the p-value) tells us the probability that χ^2 could have been larger by chance if the best fit model is correct. That is to say $Q = \Pr(\chi^2 \leq X|H_0)$, the probability that we could have drawn a lager χ^2 from a distribution X, if the null hypothesis H_0 is true. Fits are deemed acceptable if $Q \gtrsim 0.05$. However, we can see that this could be manipulated with certain prior choices. For example, consider choosing a prior for a parameter with a central value equal to the (known from a previous fit) fit output[10] and/or an unjustifiably large uncertainty. This will cause the corresponding term in Equation (4.41) to give a negligible contribution to χ^2. Similarly, an unjustifiably precise prior (with smaller uncertainty than the data is able to resolve) will override information from the data and fix the fit parameter at that value. These examples might be considered overfitting, as would be providing a model with too many parameters, on which the priors are too loose.

To test for overfitting we add a second goodness of fit test, the Gaussian Bayes Factor (GBF), typically quoted as log(GBF). This number is an approximation to the Bayes factor for data D given models M_1 and M_2

$$\text{Bayes Factor} = \frac{\Pr(D|M_1)}{Pr(D|M_2)} = \frac{\frac{\Pr(M_1|D)\Pr(D)}{\Pr(M_1)}}{\frac{\Pr(M_2|D)\Pr(D)}{\Pr(M_2)}} = \frac{\Pr(M_1|D)\Pr(M_2)}{\Pr(M_2|D)\Pr(M_1)}, \tag{4.52}$$

in the limit where our distributions are Gaussian. log(GBF) is calculated automatically by lsqfit [85] (along with χ^2 and Q) and is the probability of obtaining the data by randomly sampling the model. This isn't very useful in isolation, but when comparing two fits to the same data with different models or priors, the log(GBF) tells us which is most likely to be the true model the data came from. We consider a change in log(GBF) of greater than 3 to be significant. log(GBF) penalises overfitting, and so can be used in conjunction with χ^2 to test a good fit. We choose fits for which χ^2 is minimised (with $Q \gtrsim 0.05$), and log(GBF) is maximised.

[10] This is obviously immoral, as this is not a priori knowledge.

Another legitimate use of log(GBF) is to tune priors for which we have no intuition. Whilst fine tuning of individual priors should be avoided, the method of empirical Bayes can be used to justify broad brush adjustments to groups of priors which boost the Bayes factor. We make limited use of this in the following Chapters, but we do use an empirical Bayes study of all prior widths to confirm that our priors are conservative (that is to say the GBF favours narrower priors).

SVD and prior noise

Unfortunately, the χ^2 values calculated for fits with broad priors and svd cuts are not always kosher. We discussed a situation with a broad prior above which could artificially reduce χ^2. Similarly, SVD cuts reduce χ^2 because random fluctuations in the uncertainty added by the SVD cut are not taken into account. For more details of this see appendix D of [90] or the lsqfit documentation [85]. lsqfit includes an option to nullify these effects by including some noise in the priors and sampling the SVD cut error distribution. Standard practice in the fits described in the following Chapters is to find a good fit with noise off, and then to check that χ^2 is still acceptable and the fit results have not moved significantly with the SVD and prior noise on. The random nature of the noise means that χ^2 values and fit results are not identically reproducible for a fit with noise, so whilst we check that these are acceptable, the fit results quoted will always be for a fit with noise off. As a final check that the fit is good, we can plot the fit function and the data, another facility which is built into lsqfit.

4.4 Form factors

In this work we focus on pseudoscalar to pseudoscalar electroweak decays, where the current for the $q_1 \to q_2$ quark transition can be written $V^\mu - A^\mu$ in terms of vector $V^\mu = \bar{\psi}_{q_2}\gamma^\mu\psi_{q_1}$ and axial vector $A^\mu = \bar{\psi}_{q_2}\gamma^\mu\gamma^5\psi_{q_1}$ components. Parity invariance means that of these two parts, only the vector component contributes to decays. We also consider a scalar current $S = \bar{\psi}_{q_2}\psi_{q_1}$, which in the HISQ formalism is also required to non-perturbatively renormalise the vector matrix element, and a tensor current insertion, $T^{\mu\nu} = \bar{\psi}_{q_2}\sigma^{\mu\nu}\psi_{q_1}$ where $\sigma^{\mu\nu} = \frac{i}{2}[\gamma^\mu, \gamma^\nu]$, which can be relevant for beyond the SM scenarios.

We can construct scalar, vector and tensor form factors f_0, f_+ and f_T using the matrix elements calculated with these three current insertions. These matrix elements and form factors (with different parent and daughter mesons) will be common to all three of the calculations that follow,

so we will describe them here for a general case which covers all three. Consider the $H \to L$ decay, where $H = \bar{h}\iota$ denotes the 'heavy' pseudoscalar parent meson, comprised a parent heavy quark[11] h and spectator quark ι and $L = \bar{s}\iota$ a 'light' pseudoscalar daughter meson comprising a daughter strange quark s and the same spectator ι (note that $\iota \in \{l, s\}$ covers all the work in this thesis). The spin-taste combinations of H and L will vary between current insertions, and this has been covered in Section 4.2. We will denote the Goldstone cases H and L and the non-Goldstone $\hat{H}$ and $\hat{L}$ respectively and restrict ourselves to component T^{10}, though we write more generally about T^{i0} where $i = 1, 2, 3$. To construct matrix elements from the three-point amplitudes J_{00}^{nn} we obtained from our correlator fits, we write

$$\langle \overset{(\wedge)}{L} | J | \overset{(\wedge)}{H} \rangle = 2Z_{\mathrm{disc}} \sqrt{M_H E_L} J_{00}^{\mathrm{nn}}, \tag{4.53}$$

for $J = S, V, T$, where we always use the Goldstone parent mass and daughter energy. These differ from the non-Goldstone cases only by discretisation effects, which we discuss and account for below. Here Z_{disc} is a normalisation which corrects for discretisation effects at $\mathcal{O}(m_h)^4$ and is defined [63],

$$Z_{\mathrm{disc}} = \sqrt{\cosh(m_{\mathrm{tree}}) \Big(1 - \frac{1 + \epsilon_{\mathrm{tree}}}{2} \sinh^2(m_{\mathrm{tree}})\Big)}, \tag{4.54}$$

where m_{tree} and ϵ_{tree} are defined in Equations (3.28) and (3.29) respectively. We suppress m arguments, but we only need to include $m_{\mathrm{tree}}(m_h)$, and $\epsilon_{\mathrm{tree}}(m_h)$, as we neglect strange and light quarks by convention (their contributions are vanishingly small anyway). In principle we could simply include another quark by multiplying by Z_{disc} evaluated at that quark mass, as is necessary for the $B_s \to D_s^*$ decay in [93].

From these matrix elements we can construct the scalar, vector and tensor form factors,

$$Z_V \langle L | V^\mu | \hat{H} \rangle = f_+(q^2) \Big(p_H^\mu + p_L^\mu - \frac{M_H^2 - M_L^2}{q^2} q^\mu \Big) + f_0(q^2) \frac{M_H^2 - M_L^2}{q^2} q^\mu, \tag{4.55}$$

$$\langle L | S | H \rangle = \frac{M_H^2 - M_L^2}{m_h - m_s} f_0(q^2), \tag{4.56}$$

$$Z_T \langle \hat{L} | T^{i0} | \hat{H} \rangle = \frac{2 i M_H p_L^i}{M_H + M_L} f_T(q^2). \tag{4.57}$$

[11] Which of these is the anti-quark is of no relevance as our decays do not include QED and our matrix elements are CP invariant.

Here, everything is defined on the lattice. Mesons have masses M_H and M_L and quarks m_h and m_s, $q^2 \equiv (p_H - p_L)^2$ is the squared four-momentum transfer and the multiplicative i in Equation (4.57) is understood to be $\sqrt{-1}$, not the same as the i index. We also note that Equation (4.55) demands that $f_+(0) = f_0(0)$, a constraint which we can enforce in our fits below, and which helps to improve the (typically noisier) f_+ calculation, by correlating it to the (typically cleaner) f_0 one.

4.4.1 Z_V and Z_T

Equation (4.56) is an application of the partially conserved vector current (PCVC) relation[12],

$$q^\mu \langle L| V_\mu |H\rangle = (m_h - m_s) \langle L| S |H\rangle , \tag{4.58}$$

a continuum expression which also holds on the lattice for the HISQ formalism [10], up to discretisation effects. In this case, the combination of valence masses and matrix element $(m_h - m_s) \langle L| S |H\rangle$ is absolutely normalised [94]. Using this and considering the construction of Equation (4.56) it is clear that no renormalisation is required in order to obtain f_0. For the f_+ expression, however, this is not the case. Whilst it is possible to construct a vector current which is conserved in the HISQ formalism, this contains both one- and three-link operators (see appendix A in [95]), which makes it very unattractive for use in lattice calculations. Firstly because it is very complicated [95], and secondly because point-split current operators can lead to noisier correlators [96], although this is not universally true [97]. Instead, we proceed with the simple local vector current insertion defined above, and use Equation (4.58) at zero recoil (where both the parent and daughter meson are at rest and all four-momentum difference is carried away by the current) to define on the lattice,

$$Z_V = \left. \frac{(m_h - m_s) \langle L| S |H\rangle}{(M_H - M_L) \langle L| V^0 |\hat{H}\rangle} \right|_{\vec{p}_L = \vec{0}} , \tag{4.59}$$

where Z_V can then be applied as a normalisation factor to the vector matrix element, as in Equation (4.55). Indeed, if we take $\vec{p}_L = \vec{0}$ in Equation (4.55) we recover Equation (4.59). This method allows fully non-perturbative normalisation of the vector matrix element, eliminating systematic errors resulting from this matching in earlier methods. Z_V is a single normalisation constant (for each gluon ensemble and m_h). This means that we can use the $\vec{p}_L = \vec{0}$ case in

[12] We drop the $\hat{H}$ notation for this continuum expression as taste is unphysical.

Equation (4.59), which gives us the most precise Z_V (because adding momentum to propagators increases noise) for cases with $\vec{p}_L \neq \vec{0}$ too. These cases differ only by discretisation effects which we will deal with below. Tests in [98] confirm this. This is also helpful because it means we can compute just the temporal component of V^μ, which we need for f_+ anyway, and avoid computing spatial components, as is the case in Chapters 5 and 6.

The normalisation of the tensor matrix element Z_T does not have such a simple calculation as Z_V and until recently was only available perturbatively, with significant errors. However, a recent HPQCD lattice calculation [99] provides Z_T in the RI-SMOM scheme and matched to $\overline{\text{MS}}$, correcting for contamination from condensates. Table VIII of [99] gives Z_T values on all of the ensembles we will use in this work, and their correlation matrix is in Table IX. It's worth noting that Z_T is the only part of the form factor expressions above which carries any dependence on the renormalisation scheme scale. The running of the normalisation to a desired scale is simple enough to implement and we will address this when relevant.

4.4.2 The modified z expansion

For each of our different ensembles we can use Equations (4.55), (4.56) and (4.57) to determine values of f_+, f_0 and f_T on the lattice at different q^2 values. We want to fit this data to the full physical q^2 range, from $q^2 = 0$ to $q^2_{\text{max}} = (M_H - M_L)^2$. We also need to tie together our different gauge field ensembles in order to fit the a dependence and extrapolate to the continuum $a = 0$. The data also contains lattice artefacts; discretisation effects that depend on the lattice spacing, as well as masses in the simulation which are not tuned to their continuum physical values.

We deal with all of these things simultaneously in one fit, using a modified z expansion, which was first introduced in [94]. For the sake of example, let's focus on the specific $D \to K$ decay. In the language of z expansions [100, 101, 102], it is very common to refer to q^2 as the Mandelstam variable t. We avoid this notation, for obvious reasons, but it is the origin of the t_0 and $t_\pm$ notation that follows. t_0 here should be distinguishable from that discussed in Section 4.2 from context. The physical range of q^2 is $0 \leq q^2 \leq q^2_{\text{max}} = (M_D - M_K)^2$ and in the complex ($t = q^2$) plane we see a branch cut from $t_+ = (M_D + M_K)^2$ upwards. This corresponds to the threshold for DK production. In general this point reflects the lowest mass combination with the correct quantum numbers that we can construct from our current with a pair of quarks out of the vacuum, so $B \to K$ and $B_s \to \eta_s$ decays both use $t_+ = (M_B + M_K)$ (actually $t_+ = (M_H + M_K)$ in the heavy-HISQ method (see Section 4.5)). We map q^2 to a function z which lies within the

unit circle on the complex z plane

$$z = \frac{\sqrt{t_+ - q^2} - \sqrt{t_+ - t_0}}{\sqrt{t_+ - q^2} + \sqrt{t_+ - t_0}}, \tag{4.60}$$

where t_0 is the point in q^2 space which maps to $z = 0$. Typically we take $t_0 = 0$, but other cases and the impact of this choice are discussed below.

Using the Bourreley-Caprini-Lellouch (BCL) parameterisation [102], which has now become the standard for the heavy to light decays we are investigating, we expand our form factors as polynomials in z,

$$f_0(q^2) = \frac{1}{P(q^2, M)} \sum_{n=0}^{N-1} a_n^0 z(q^2)^n, \tag{4.61}$$

$$f_+(q^2) = \frac{1}{P(q^2, M)} \sum_{n=0}^{N-1} a_n^+ \Big(z(q^2)^n - \frac{n}{N}(-1)^{n-N} z(q^2)^N \Big), \tag{4.62}$$

$$f_T(q^2) = \frac{1}{P(q^2, M)} \sum_{n=0}^{N-1} a_n^T \Big(z(q^2)^n - \frac{n}{N}(-1)^{n-N} z(q^2)^N \Big), \tag{4.63}$$

where we typically find $N = 3$ is more than sufficient to capture the z dependence. $a_n^{0,+,T}$ are coefficients which contain fit parameters to account for discretisation effects and quark mistunings and vary from decay to decay (see Chapters 5, 6 and 7 for details in each case), and $P(q^2, M) = 1 - q^2/M^2$ is a pole function. Removing this pole ensures that the remaining function is analytic and expressible as a polynomial in z. The pole masses in $P(q^2, M)$ vary from decay to decay and will be detailed in specific cases below. For our $D \to K$ example above, D_{s0}^* is the pole mass in the f_0 expression and D_s^* in the f_+ and f_T cases. In some cases, extra factors will be added to these expressions to allow for chiral perturbation theory to tune the light quark mass (see Section 4.4.2). These functions are described in the chapters where they are applied.

We once again perform the fits with the lsqfit package [85] which keeps track of uncertainties and correlations using gvar [87]. This time we are not dealing with many samples that need a covariance matrix estimating, and instead with fitting a modest number of data points which already have uncertainties and a correlation matrix. For this reason we don't need to conduct the same precise SVD analysis as with the correlator fits above. We do, however conduct tests by modifying the SVD cut to check this has no effect on the fit.

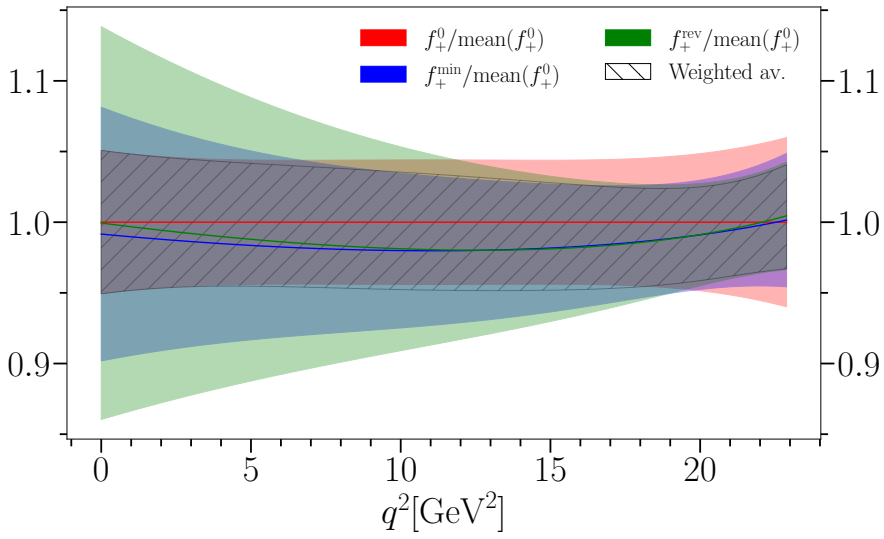

Figure 4.6: f_+ for the $t_0 = 0$ (f_+^0), 'minimum' (f_+^{min}) and 'reverse' (f_+^{rev}) choice (see text), as well as the weighted average of the three. In each case, the result is divided by the mean on the $t_0 = 0$ case.

Choosing t_0

A common choice for t_0 is $t_0 = 0$. This maps $q^2 = 0 \to z = 0$ and makes enforcing the $f_+(0) = f_0(0)$ constraint trivial. The maximal value of $|z|$ is at q^2_{max}. Other common choices are the 'reverse' choice, where[13] $t_0 = t_- = (M_H - M_L)^2$ which reverses $|z|$ with respect to q, so $z = 0$ at q^2_{max} and z is maximal at $q^2 = 0$, and the 'minimal' prescription $t_0 = t_+(1 - \sqrt{1 - t_-/t_+})$ which minimises the maximum size of $|z|$ for a given q^2 range. In general, our form factors should be insensitive, within error, to the choice of t_0, however, we do find that changes can affect the uncertainty distribution in q^2 space. Figure 4.6 shows an example of the effect of different t_0 choices on f_+ for the $B \to K$ decay (discussed in Chapter 7). The three bands show f_+ for the $t_0 = 0$ (f_+^0), 'minimum' (f_+^{min}) and 'reverse' (f_+^{rev}) choice described above, and the hatched band is the correlated, weighted average. In each case, the result is divided by the mean on the $t_0 = 0$ case. We can see that, whilst the form factors agree within error, there is some variation in the way the uncertainty is distributed. This effect is most profound in f_+.

We examine this error distribution further in Figure 4.7, which shows the relative error of the

[13] Note that whilst in some cases t_+ and t_- use the same masses, and can be written $t_\pm = (M_1 \pm M_2)^2$, this is not true in general.

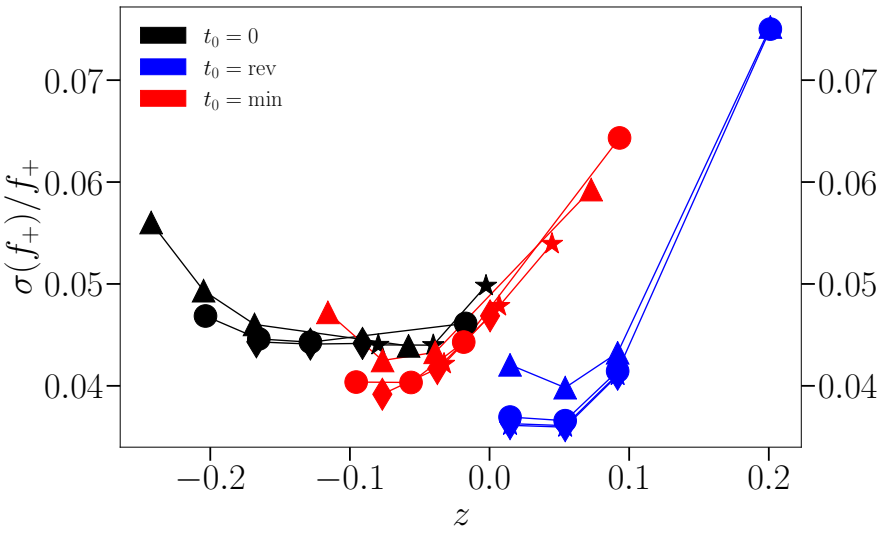

Figure 4.7: The relative error on f_+ for the $t_0 = 0$ (f_+^0), 'minimum' (f_+^{min}) and 'reverse' (f_+^{rev}) choice (see text). In each case, the data points indicate the z values for the set 8 data points in [3], with lines joining points of the same mass, which share a symbol.

continuum vector form factor with the pole removed. The points plotted are at the z values of the data points from set 8 in [3], with different m_h values having different symbols (star, diamond, circle, triangle in order of increasing mass), and different z values at the same mass joined by a solid line. Values which have $q^2 < 0$ have been excluded. The exercise is repeated for the three t_0 choices above, and we can see that t_0 has some effect on the uncertainty. In particular, the general trend seems to suggest that points which are mapped to smaller values of z have a lower uncertainty, which supports our observations in Figure 4.6, where $t_0 = 0$ leads to the lowest uncertainty at $q^2 = 0$ and the reverse case reduces error at q^2_{max}. It is worth noting that Figure 4.7 represents the clearest example of this effect, which is much less apparent on coarser ensembles, possibly as a result of the z range being greatly reduced. It's still not very clear to us what is driving this, but it potentially warrants future investigation. In the case we have given, this seems to allow us to make a choice about where we want to minimise our uncertainty. In general this is at $q^2 = 0$, which also also marries up with our standard (and by far the most convenient) choice $t_0 = 0$. Alternatively, doing several fits with different t_0 choices and then taking a weighted average of the (correlated) results, is a time consuming, though effective method to minimise errors across the q^2 range as in Figure 4.6.

Enforcing $f_+(0) = f_0(0)$

One of the benefits of the choice $t_0 = 0$ is that we can cleanly impose the constraint $f_+(0) = f_0(0)$, because this simply requires $a_0^+ = a_0^0$ in the continuum. The way a_n is further broken down varies from decay to decay and will be discussed in the relevant chapters below, but in general it will always be the case that this constraint can be satisfied by setting some coefficients equal to one another. The real benefit comes in heavy-HISQ below, where a_n is a function of the heavy quark mass, and it is trivial to implement this for arbitrary mass, which is a powerful constraint.

For other choices of t_0 this is not so simple. The best alternative method is no basket of fruit - we add an artificial data point equal to zero, and add a model in the fit which says $f_+(0) - f_0(0)$ is equal to this data point. One problem with this is philosophical. For $a_n(M_H)$ we have to choose a value, or values of M_H to enforce this, which reduces the potency of a constraint which is true for all M_H. Another issue is numerical. We can't have a point 0(0) in our fit, so this point has to be given a small error on the order of 10^{-4}. This uncertainty cannot be too small, and leads to the requirement of an SVD cut. Happily, the wrong combination of SVD cut and error size leads to a catastrophic failure of the fit, which is otherwise stable, so, whilst this method isn't perfect, it is a reliable implementation of the constraint.

One final method we mention in passing is similar to that above, but is more passively enforced, without the requirement for an arbitrary data point. This is to redefine $f_+(q^2) \to f_+(q^2) + f_0(0) - f_+(0)$ in the fit. Again, this requires thought about which heavy masses to choose, and tests with this method found it to be no better than the artificial data point, so it was dropped in favour of the more intuitive option when fits were required for $t_0 \neq 0$. In this thesis, all results use $t_0 = 0$, so these other constraints are only used in cases where fit stability is tested against varying t_0.

Chiral perturbation theory

In Chapters 5 and 7, we require a further modification to Equations (4.61), (4.62) and (4.63) to allow for light quark mass mistuning. This involves multiplying by an overall $\mathcal{L}$, a function derived in chiral perturbation theory (χ-PT).

For the $B \to K$ and $D \to K$ transitions, which we are interested in, Equation (34) of [103] gives,

$$\mathcal{L}^{H\to K} = 1 + \frac{9}{8}g^2\frac{\overline{A}(M_\pi^2)}{F^2} + \Big(\frac{1}{2} + \frac{3}{4}g^2\Big)\frac{\overline{A}(M_K^2)}{F^2} + \Big(\frac{1}{6} + \frac{1}{8}g^2\Big)\frac{\overline{A}(M_\eta^2)}{F^2}, \tag{4.64}$$

where,

$$\overline{A}(m^2) = -\frac{m^2}{(4\pi)^2}\log\Big(\frac{m^2}{\mu^2}\Big). \tag{4.65}$$

Here, g is the coupling between H, H^* and light mesons, and F is the pion decay constant as defined in χ-PT [72] (that is with $f_\pi \approx 93$ MeV, as opposed to the experimental convention which includes a factor of $\sqrt{2}$). Putting all of this together, we can define $x_{\mathcal{M}} = \frac{M^2_{\mathcal{M}}}{(4\pi f_\pi)^2}$, for meson $\mathcal{M}$, where $4\pi f_\pi$ is the chiral scale, which we also take as a value for μ. Then rewrite Equation (4.64)

$$\mathcal{L}^{H\to K} = 1 - \frac{9}{8}g^2 x_\pi \log x_\pi - \Big(\frac{1}{2}+\frac{3}{4}g^2\Big)x_K \log x_K - \Big(\frac{1}{6}+\frac{1}{8}g^2\Big)x_\eta \log x_\eta. \tag{4.66}$$

In order to obtain x_π from quantities available in our simulation, we take the ratio of $4\pi f_\pi = 1364$ MeV to $M^{\text{phys}}_{\eta_s} = 688.5$ MeV [75], (inserting a factor of 2 to account for the different f_π definitions) and write,

$$x_\pi \approx 2\frac{m_{\text{spectator}}}{5.63 m_s}. \tag{4.67}$$

We can find $x_{\mathcal{M}}$ from this using $x_{\mathcal{M}} = x_\pi M^2_{\mathcal{M}}/M^2_\pi$ where we use proxies $M^2_\pi = (m_{\text{spectator}}/m_s)M^2_{\eta_s}$ and $M^2_\eta = (M^2_\pi + 2M^2_{\eta_s})/3$.

We also take this opportunity to account for finite volume corrections, arising from the fact that our lattice is not infinite. These are particularly important for light quarks and are considered in terms of the value $M_\pi L$ for each gluon ensemble. Following the method in [104], we account for these by modifying the x_π term in Equation (4.66) with the addition of δ_{FV} which is set to 0 in the infinite volume limit,

$$\mathcal{L}^{H\to K} = 1 - \frac{9}{8}g^2 x_\pi(\log x_\pi + \delta_{\text{FV}}) - \Big(\frac{1}{2}+\frac{3}{4}g^2\Big)x_K \log x_K - \Big(\frac{1}{6}+\frac{1}{8}g^2\Big)x_\eta \log x_\eta. \tag{4.68}$$

We calculate δ_{FV} using Equation (47) of [105] and the values on each of our 8 ensembles are listed in Table 4.1. We find that δ_{FV}, and indeed $\mathcal{L}$ more generally, has very little effect on any work we present here, as will be discussed below.

Table 4.1: δ_{FV} values used on each ensemble.

Set	δ_{FV}
1	0.12907825
2	0.04894993
3	0.06985291
4	0.05841163
5	0.018106911
6	0.020801419
7	0.020801419
8	0.027538708

4.5 Heavy-HISQ

So far, we have (hopefully) been careful to talk about the generic 'heavy' quark. This is partly to ensure that the discussion remains applicable to all three of the decays covered below, involving both b and c parent quarks, but also to allow for a seamless transition to heavy-HISQ.

4.5.1 The problem

As we have already discussed in Chapter 3, the HISQ action has very small discretisation errors. However, any power series in heavy quark mass am_h will cause problems if $am_h > 1$. In practice, we often take the upper limit for well behaved discretisation effects to be $am_h \approx 0.8$. This puts serious constraints on the masses we can reach with ensembles that are computationally feasible. For example, whilst the charm quark mass is accessible to all ensembles, on the finest ensembles used in this work, ultrafine 5 (set 8), the physical mass of the b quark corresponds to $am_b \approx 0.9$. This means that we cannot simply perform calculations at the b mass.

In the past (and maybe even the present), many calculations used effective theory descriptions of b quarks, like non-relativistic (NRQCD) [106], in combination with lighter quarks using an improved action like HISQ. This best of both worlds approach combines the precision of HISQ for light quarks with the ability to work at the physical b mass. However, like all lunches, it is not free. The cost is in a matching of the coefficients of the effective theory to continuum QCD [63, 107, 108], which is done in perturbation theory. Whilst errors are added to account for this, they are often hard to quantify accurately, as they require estimates of neglected higher order effects from the matching [108]. These uncertainties were becoming the dominant error for NRQCD (see e.g. [109] for the case of $B \to K$), signalling the need for a new approach.

4.5.2 The solution

The solution proffered by HPQCD, and adopted in this work, is heavy-HISQ. We perform our calculation at a number of heavy masses, typically $am_c \leq am_h \leq 0.8$ on each ensemble. This gives us information about the heavy mass dependence of the form factors, in much the same way that performing a calculation on several different ensembles tells us about the dependence on lattice spacing. Thus we can combine into the a_n coefficients of the z expansion, which allow for discretisation effects in a, extra terms which parameterise the am_h dependence of the form factors, which can then be evaluated at the physical point, i.e. the b mass. The exact structure of a_n varies from decay to decay and is given in the dedicated Chapters below, however, the heavy quark mass dependence is based on Heavy Quark Effective Theory (HQET) [110, 111] (see e.g. [112] for a good overview).

Similar approaches for extrapolating up to the b mass have also been adopted with other quark discretisations [113, 114, 115, 116, 117, 118, 119, 120, 121, 122]. These efforts, however, focused on the b mass and b meson decay constants, whereas here we are extrapolating hadronic form factors to the physical mass. The works above also do not take data as close to the physical b mass as we do here, using data points around the charm, and up to half the physical b mass. In our case, on our finest ensemble, we can get to about 0.9 of the physical b mass.

HQET

HQET [110, 111] takes advantage of the large separation of mass scales between the b (or c) quark and $\Lambda_{\rm QCD} \approx 0.5$GeV which is associated with the light degrees of freedom. This can be used to construct an effective theory of decays of heavy mesons like the B, which becomes increasingly good as $m_h \to \infty$.

As $m_h \to \infty$ the light degrees of freedom experience only the colour field of the heavy quark, which can be considered at rest in the rest frame of the hadron. This means that the configurations of light quarks in two systems which differ only in the flavour or spin of the heavy quark will be the same [123] (in the infinite heavy quark mass limit), allowing us to relate the behaviour of the B and D meson, for example. Such considerations lead to a power series in the inverse heavy quark mass, typically with scale set by $\Lambda_{\rm QCD}$, and specific examples will be shown in more detail in the chapters which follow.

As well as the simple power series, there is a logarithmic term, arising from the matching between HQET and QCD [124, 125], this term will be included in our fits where relevant, and is discussed

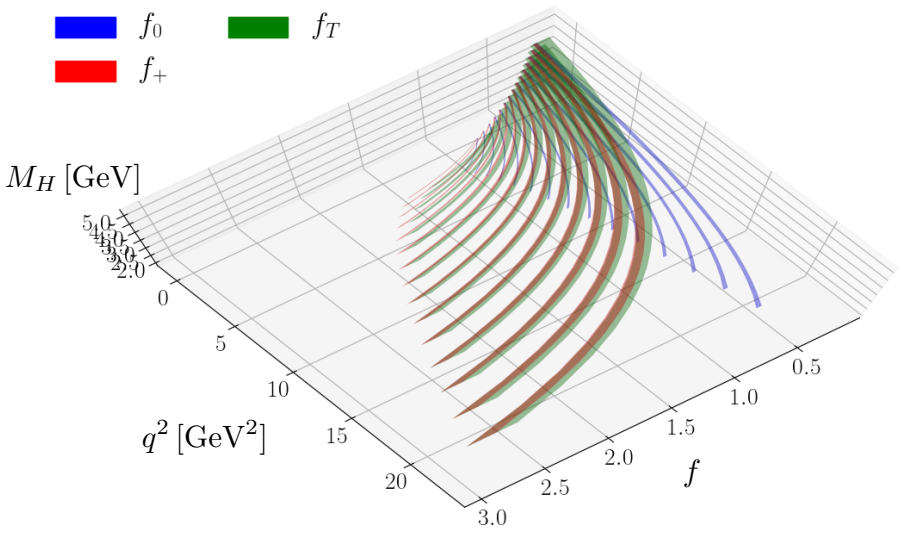

Figure 4.8: Scalar, vector and tensor form factors across the whole physical q^2 range evaluated at different masses M_H between M_D and M_B. This plot contains results from [3] which is the subject of Chapter 7.

in Chapters 6 and 7.

4.5.3 Heavy-HISQ in action

One offshoot of the heavy-HISQ approach is that we often get a second calculation 'for free'. For example, in the case of $B \to K$, we run our heavy mass from the physical charm up to close to the physical bottom, and extrapolate over this range. This means that we also calculate $D \to K$[14]. Figure 4.8 illustrates this point in a 3D plot [3]. We plot scalar, vector and tensor form factors across the whole physical q^2 range, with slices taken periodically to show the result at different heavy masses between M_D and M_B. The purpose of this plot is more to give a feel of the scope of heavy-HISQ, and not to provide accurate numerical values. It shows the form factors calculated in [3], which are the subject of Chapter 7, where the fit form which is used to generate this plot can be found.

[14] At least in the case where QED is ignored. A lattice calculation with QED would make this more complicated - what charge do you take for the heavy quark if you want it to be both a c and a b?

Chapter 5

$D \to K$ form factors: an improved determination of V_{cs}

In this chapter we will cover the work which is published in [1]. Many of the techniques used have been presented in Chapter 4 meaning we will focus primarily on the specific context and results of the calculation, as opposed to the calculation itself. Where appropriate, we will refrain from including lengthy numerical details which are tangential to the discussion and can easily be accessed in [1].

5.1 Introduction

We outlined the SM theory behind the semileptonic $D \to K\ell\nu$ decay in Section 2.3. To recap, flavour changing weak interactions involving the emission of a W boson involve elements of the CKM matrix [126, 127].

As shown in a Feynman diagram of the decay (Figure 5.1), the specific matrix element in this case is V_{cs}, which is an element on the leading diagonal of the matrix, i.e. a strongly favoured transition, and therefore $|V_{cs}| \approx 1$. In the SM, the CKM matrix is unitary, and so a good way to look for failures of the SM is via precise, independent determinations of CKM matrix elements, which we can then use to check for unitarity in the rows and columns of the matrix. As the leading diagonal elements are all close to unity, with off diagonal elements relatively small, it's important to determine these elements very precisely, as even a small relative uncertainty can

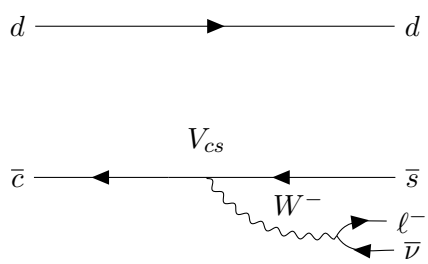

Figure 5.1: Feynman diagram for $D^- \to K^0 \ell^- \bar{\nu}$

dominate the unitarity test.

The accuracy with which unitarity tests can be performed varies substantially across the matrix. For example, the unitarity of the first row has been tested to a precision of 0.05% [29]. Tests of unitarity for other rows and columns of the CKM matrix are much less stringent, either because of larger experimental uncertainties, larger theoretical uncertainties or both [29]. The determination of V_{cs} proceeds most directly, as is similar for V_{us}, either through a study of leptonic decays of the D_s meson or through D semileptonic decay to $K\ell\nu$, which is what we shall focus on. Our improved determination of V_{cs} targets a particularly poorly determined element, which dominates the uncertainty in both of its unitarity tests [29];

$$|V_{cd}|^2 + |V_{cs}|^2 + |V_{cb}|^2 = 1.025(2)_{V_{cd}}(22)_{V_{cs}}, \tag{5.1}$$

$$|V_{us}|^2 + |V_{cs}|^2 + |V_{ts}|^2 = 1.026(22)_{V_{cs}}, \tag{5.2}$$

where only 2 significant figures of uncertainty are quoted and the $|V_{cs}|$ value used here is the average of the leptonic and semileptonic results before our calculation, as given by the PDG [29].

5.1.1 Previous status of V_{cs}

Here we briefly summarise the status of V_{cs} determinations available prior to our calculation. For a more detailed summary, see [1].

Leptonic determinations

The leptonic decay $D_s \to \ell\nu$ would seem to be the most intuitive way to access $|V_{cs}|$ as there is no need for q^2 dependent form factors; everything can be parameterised in terms of the decay

constant. However, the experimental measurement of the branching fraction for D_s leptonic decay has been challenging, with the average drifting downwards slowly with time as newer results are added (see [128, 129]). The current situation is reviewed in 'Leptonic decays of charged pseudoscalar mesons' by the PDG [29, 128]. They obtain an average from experiment of

$$|V_{cs}| f_{D_s} = 245.7(4.6) \text{ MeV}, \tag{5.3}$$

(their Equation (71.24) [29]) with similar findings in HFLAV [129], who do not follow the same approach to radiative corrections.

Early full[1] lattice QCD calculations of the D_s decay constant suffered from large uncertainties. The first calculation for $N_f = 2$ was done by ETM [130] in 2009 and for $N_f = 2 + 1$ by FNAL/MILC [131] in 2005. These have been improved over the years (see Table 34 of [132]), in particular by making use of the HISQ action [10]. HPQCD used this to obtain a 1% accurate result for f_{D_s} [133, 134] in 2010. Combined with the higher experimental average for the branching fraction at that time it led to a V_{cs} result with a central value above 1. More recently, results from the Fermilab/MILC collaboration [79] using HISQ give a 0.2% uncertainty on f_{D_s}. Recent non-staggered results have also been produced by ETM [135], but they have failed to compete with the accuracy of the HISQ results, being an order of magnitude less precise. 'Leptonic decays of charged pseudoscalar mesons' [29] then give a leptonic determination,

$$|V_{cs}|_{\text{lept}} = 0.983(13)(14)(2), \tag{5.4}$$

where the first uncertainty is from experiment, the second from radiative corrections and the third from f_{D_s}. It is clear that, in the case of leptonic decays, it now is the experimental uncertainty and radiative corrections which dominate the uncertainty, with the lattice uncertainty insignificant at the current level of precision.

In the 'CKM Quark-Mixing Matrix' section of the PDG [29] they quote their leptonic decay average as,

$$|V_{cs}|_{\text{lept}} = 0.992(12), \tag{5.5}$$

though this does not include uncertainties for electroweak (EW) or long distance QED effects.

[1] I.e. unquenched - including dynamical sea quarks.

Semileptonic determinations

The situation with semileptonic $D \to K$ decays is somewhat different. Smaller experimental uncertainties have been available for some time but form factors from lattice QCD calculations have proved to be more difficult. Additionally, uncertainties from electroweak radiative corrections have not previously been considered, something we will address in our calculation.

Experimental results for $D \to K\ell\overline{\nu}$ are available from BaBar, Belle, BES III and CLEO-c [129] and will be discussed in more detail below. They are either given in the form of a differential distribution in bins of q^2 or, following a fit to the distribution combined with an analysis of radiative bremsstrahlung corrections using PHOTOS [136], a value for $\eta_{\rm EW}|V_{cs}|f_+(0)$[2] or the branching fraction $\mathcal{B}$.

The lattice side of the calculation follows a similar trajectory to the leptonic case above. Full lattice QCD calculations of the $D \to K$ form factors again began before experimental results were available [137] and were again limited in precision. The use of the HISQ action by HPQCD brought a big improvement [94] and HPQCD preceded to a determination of f_+ across the full physical q^2 range in [138], allowing a 1.6%-accurate determination of $|V_{cs}|$ using a bin-by-bin comparison of the differential distribution with experiment. Recently the European Twisted Mass Collaboration (ETMC) determined the full shape of the $D \to K$ form factors [139, 140] using the twisted mass formalism [141, 142, 143] and combined that with experimental results to obtain a 3.5% accurate result for $|V_{cs}|$. Work is also underway by other groups; see, for example, [144, 145].

The ETMC result for $f_+^{D\to K}(0)$ is used in the 'CKM Quark-Mixing Matrix' review in [29] (quoting [146]) to give a semileptonic determination of V_{cs} as

$$|V_{cs}|_{\rm semi} = 0.939(38), \tag{5.6}$$

where, as noted above for the leptonic case, uncertainties from electromagnetic effects in the final state, and electroweak radiative corrections have not been considered. Combining their results over the full range of q^2 with experiment, ETMC [140] instead obtains

$$|V_{cs}|_{\rm semi} = 0.978(35), \tag{5.7}$$

with similar uncertainty. We note that these semileptonic determinations are somewhat less

[2] $\eta_{\rm EW}$ is an electroweak correction which is addressed below.

precise than the leptonic one (Equation (5.5)), which dominates the PDG 'CKM Quark-Mixing Matrix' average (ignoring EW and QED corrections) of $|V_{cs}|_{\text{average}} = 0.987(11)$.

Below we will improve upon the semileptonic determination of V_{cs}, using new form factor determinations in three semi-independent ways, and taking advantage of the full q^2 range and multiple sources of experimental data. We will arrive at a semileptonic determination of 0.9663(80), which is more precise even than the existing leptonic average, and a big improvement on the semileptonic one.

5.2 Lattice calculation

Much of the lattice calculation is based on methods outlined in Chapter 4, however, there are of course peculiarities unique to the production of data and fitting procedure, which will be covered below.

5.2.1 Simulation details

For this calculation, data from all eight of the MILC $N_f = 2 + 1 + 1$ gluon ensembles discussed in Section 3.4 and listed in Table 3.1 were used. In particular, it is worth noting that the first three ensembles, sets $1, 2$ and 3 include physical (though degenerate) light quark masses. These ensembles are important in the case of calculations such as this one, involving valence light quarks. For $D \to K$, we require parent (c) and spectator (l) quarks, both at the tuned valence masses given in Table 3.1. We also require daughter s quarks, again at their tuned valence masses, but with a number of twists applied, so as to cover the full q^2 range of the decay. Charm and strange masses were tuned in [80] to give physical η_c and η_s masses, and the light quarks are tuned using $m_s^{\text{tuned}}/m_l^{\text{tuned}} = 27.18(10)$ [79]. The twists applied on each ensemble, as well as the T values (i.e. the source sink separation in lattice units) are given in Table 5.1, along with the statistics used on each ensemble in terms of the number of configurations n_{cfg} and the number of time source (t_0) values n_{src} used on each configuration.

5.2.2 Correlation functions

The form factors required to extract V_{cs} are f_0 and f_+, as defined in Equations (4.56) and (4.55). In the language of Chapter 4, we require a Goldstone K and D with spin-taste $\gamma_5 \otimes \xi_5$, and also a non-Goldstone $\hat{D}$, with spin-taste $\gamma_5\gamma_0 \otimes \xi_5\xi_0$. These are combined with currents with scalar $(1 \otimes 1)$ and temporal vector $(\gamma_0 \otimes \xi_0)$ spin-taste structure to generate the required two-

Table 5.1: Details of the statistics, T values and K meson momenta used on each ensemble. The number of configurations n_{cfg} and time sources n_{src} are given, and momenta can be obtained from twist, θ, via $\theta = |a\vec{p}_K|N_x/(\sqrt{3}\pi)$, where N_x is the spatial dimension of the lattice, given in Table 3.1.

Set	$n_{\text{cfg}} \times n_{\text{src}}$	θ	T
1	998×16	0, 2.013, 3.050, 3.969	9, 12, 15, 18
2	985×16	0, 2.405, 3.641, 4.735	12, 15, 18, 21
3	620×8	0, 0.8563, 2.998, 5.140	14,17,20
4	1020×16	0, 0.3665, 1.097, 1.828	9, 12, 15, 18
5	1053×16	0, 0.441, 1.323, 2.205, 2.646	12, 15, 18, 21
6	499×16	0, 0.4281, 1.282, 2.141, 2.570	14, 17, 20
7	415×8	0, 1.261, 2.108, 2.946, 3.624	20, 25, 30
8	375×4	0, 0.706, 1.529, 2.235, 4.705	24, 33, 40

and three-point functions. See Section 4.2.1 for details on how these operators are implemented in the MILC code.

5.2.3 Correlator fits

Once we have two- and three-point functions on each ensemble, correlator fits are performed according to the method outlined in Section 4.3. The fits are performed in a standard way, as described, however, specific choices are to be made, which we will outline here. Firstly, we must decide which values of t_{min}, the number of initial time slices discarded from the fit, and N_{exp}, the number of exponentials in the fit, to take. These two values are somewhat related, as discussed above, and are determined through trial and error, paying attention to the goodness of fit tests in Section 4.3.5. Across our ensembles, t_{min} takes values in the range 2 to 5, and $N_{\text{exp}} = 4$ on the four coarsest lattices (sets 1, 2, 4 and 5) and 5 elsewhere. Secondly, priors (denoted $\mathcal{P}[...]$) are estimated as described in Section 4.3.3. Those which cannot be estimated using M_{eff}, A_{eff} or J_{eff} (Equations (4.44), (4.45) and (4.46)) must be provided. These are given conservative widths, and their rough size informed by log(GBF) using the empirical Bayes method described in Section 4.3.5. In general, from past HPQCD experience of similar fits [147], amplitude priors are set at $\mathcal{O}(0.1)$ and three point amplitudes at roughly 0(1). All energy splitting are always taken at 0.50(25)GeV in this thesis. Additionally, some amplitudes are adjusted by trial and error in order prevent the fit finding 'ghost' states, with zero amplitude, which interfere with the ground state. Different prior widths for large groups of priors are then trialled using empirical

Table 5.2: Priors used in the fit on each set. $A^{D,K}_{i\neq 0}$ indicates the amplitudes for non-oscillating and oscillating D mesons and for non-oscillating K mesons. $A^{K,\mathrm{o}}_i$ is the amplitude for oscillating K, which we expect to be smaller because the oscillation vanishes at zero momentum when the quark masses are the same. Parameters denoted S and V refer to the J^{kl}_{ij} parameters (see Equation (4.40)) for the scalar and temporal vector currents respectively. Columns 4 and 5 then give the priors for the ground-state to ground-state parameter cases where at least one of the states is an oscillating state. For the cases where at least one state is an excited state, $\mathcal{P}[S^{kl}_{ij\neq 00}] = \mathcal{P}[V^{kl}_{ij\neq 00}] = 0.0(5)$ in all cases.

set	$\mathcal{P}[A^{D}_{i\neq 0}]$	$\mathcal{P}[A^{K,o}_{i}]$	$\mathcal{P}[S^{kl\neq \mathrm{nn}}_{00}]$	$\mathcal{P}[V^{kl\neq \mathrm{nn}}_{00}]$
1	0.15(20)	0.05(5)	0.0(1.0)	0.0(1.0)
2	0.15(10)	0.05(5)	0.0(1.0)	0.0(1.0)
3	0.10(10)	0.05(5)	0.0(1.5)	0.0(1.5)
4	0.20(20)	0.05(5)	0.0(1.5)	0.0(1.5)
5	0.20(20)	0.03(3)	0.0(1.0)	0.0(1.0)
6	0.10(10)	0.05(5)	0.0(1.5)	0.0(1.5)
7	0.05(5)	0.02(2)	0.0(1.0)	0.0(2.0)
8	0.08(10)	0.01(2)	0.0(1.0)	0.0(1.5)

Bayes, for example changing all kaon amplitude priors widths at once. This same procedure is followed for all the correlator fits in this work. The values taken for most of these priors on each ensemble are given in Table 5.2, and can be found in [1]. The oscillating ground states are taken to be 0.4 GeV and 0.25 GeV above the non-oscillating ones for D and K respectively, based on experimental splittings [29]. Priors which are estimated from effective calculations above are given an uncertainty of $20-50\%$. As discussed in Section 4.3.3, we use the relativistic dispersion relation to inform our K priors, and the $\tilde{C}$ and $\tilde{D}$ coefficients in Equations (4.47) and (4.48) are given priors 0 ± 1. All priors are confirmed to be much (typically 5-10 times) broader then final fit parameters, and log(GBF) confirms that they are conservative, favouring narrower choices. The effect of narrowing and broadening priors on our fit is demonstrated below in Figure 5.2.

As described in Section 4.3, the required SVD cuts are calculated and applied. We also confirm that all fits have a χ^2 value close to 1 when SVD and prior noise is included.

Further to these choices and checks, tests of the fits can confirm they are stable and behaving as expected. We carry out 4 such tests.

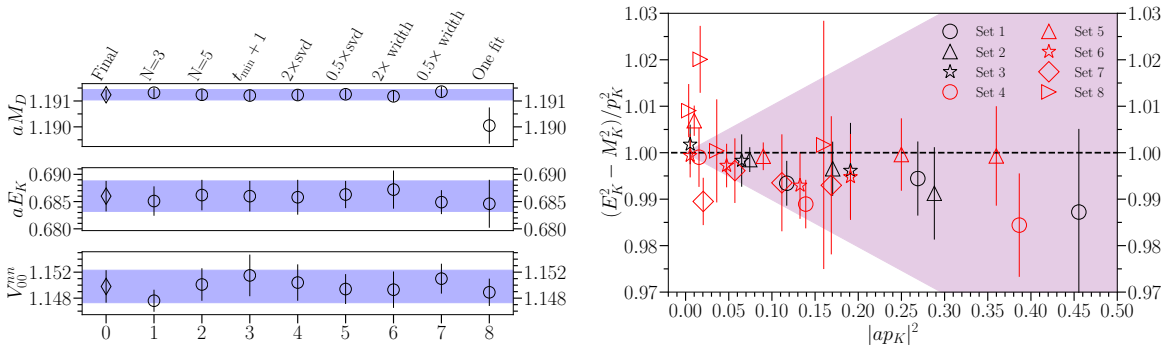

Figure 5.2: Left: stability plot for our fit on the set 5 ($a = 0.12$ fm) lattice, with our preferred fit using $N = 4$ exponentials, shown as the value at location 0. The different panels show (from the top) the mass of the D (parameter $E_0^{D,\mathrm{n}}$), the ground state energy of the K (parameter $E_0^{K,\mathrm{n}}$) with the largest twist for this set of 2.646, and current matrix element for the temporal vector current V_{00}^{nn} for twist 0.441. Tests 1 and 2 give the results from including one fewer and one more exponential respectively. Test 3 increases t_{min} by 1 across the whole fit. Tests 4 and 5 double and halve the svd cut and tests 6 and 7 double and halve all prior widths. The final test, 8, shows the results when the single correlator is fit on its own or, in the case of V_{00}^{nn}, just with the D and K two-point correlation functions required, rather than as part of one big simultaneous fit. Right: for each ensemble, we plot the ratio $(E_K^2 - M_K^2)/|\vec{p}_K|^2$ from our fit results against $|a\vec{p}_K|^2$ to check that the K meson energy agrees with that expected from the spatial momentum given to the meson in the lattice calculation. The points for gluon field configurations with physical sea u/d quark mass are in black. The ratio agrees with the expected value of 1 throughout the range of momenta and lattice spacing values. The purple wedge shows $1 \pm |a\vec{p}_K/\pi|^2$.

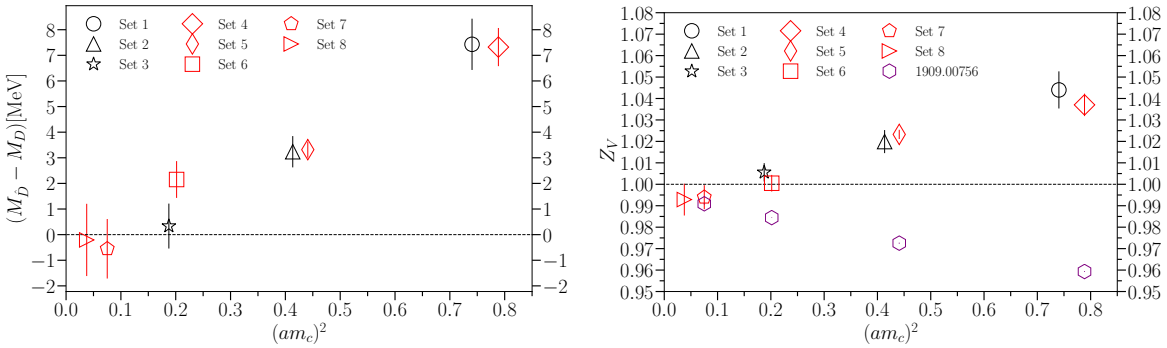

Figure 5.3: Left: the difference between the non-Goldstone $\hat{D}$ and Goldstone D meson masses, from our fit results, as a function of lattice spacing. The results show clearly that the splitting is a discretisation effect and is only a few MeV even on the coarsest lattices. Right: the renormalisation factor for the local temporal vector current, Z_V, plotted as a function of lattice spacing. The purple hexagons give results for Z_V values for the local vector current determined in a symmetric momentum-subtraction scheme on the lattice [95]. The two sets of Z_V values differ at finite lattice spacing by discretisation effects. In both cases the points in black correspond to gluon field configurations with physical u/d sea quark mass.

Figure 5.2 shows on its left hand side the stability of three parameters from one correlator fit under a variety of changes (see caption). We find that the fit is stable against these changes, and that log(GBF) favours the fit with narrower priors, indicating that we are being conservative with our choices. Other correlators are found to be similarly well behaved. On the right of this figure, we test the relativistic dispersion relation of the K, confirming that the fitted energies agree with the momenta given to the kaons, up to discretisation effects. This is indeed the case, with points mostly falling within the purple region of $1 \pm |a\vec{p}/\pi|^2$. This reassures us that our discretisation effects are small, and that our choice of fit form in Equations (4.47) and (4.48) is reasonable.

Finally, we look closer at discretisation effects in Figure 5.3, firstly by plotting the mass difference between the Goldstone and non-Goldstone D masses (left), which confirms that these effects are small and purely a lattice spacing dependent discretisation effect, as expected. We also confirm that the vector normalisation Z_V differs from that calculated in [95] only by discretisation effects. All these discretisation effects will be dealt with in our modified z expansion below.

The tabulated numerical results of these correlator fits are given in [1].

5.2.4 z expansion and continuum extrapolation

The form factors are extracted from the three-point correlator fit results as in Section 4.4, taking $H = D$ and $L = K$. Once we have form factors on each ensemble, we need to extrapolate to the continuum.

The basic form of the modified z expansion that we use to fit the form factors data across the eight ensembles is that described in Section 4.4.2, and we make the choice $t_0 = 0$[3]. We use the basic format of Equations (4.61) and (4.62), and modify them for the $D \to K$ form factors,

$$f_0(q^2) = \frac{1 + L(m_l)}{1 - \frac{q^2}{M^2_{D^*_{s0}}}} \sum_{n=0}^{N-1} a^0_n z(q^2)^n, \tag{5.8}$$

$$f_+(q^2) = \frac{1 + L(m_l)}{1 - \frac{q^2}{M^2_{D^*_s}}} \sum_{n=0}^{N-1} a^+_n \Big(z(q^2)^n - \frac{n}{N}(-1)^{n-N} z(q^2)^N \Big). \tag{5.9}$$

Here $M_{D^*_{s0}}$ is the mass of the scalar D_s meson, and $M_{D^*_s}$ that of the vector. These appear in the poles as they are the only masses below the $M_D + M_K = 2.36$GeV threshold. We obtain these masses in lattice unit by using the physical splittings [29] e.g. $aM_{D^*_{s0}} = aM_D + a(M^{\rm phys}_{D^*_{s0}} - M^{\rm phys}_D)$, where aM_D is the D meson mass obtained from our correlator fits. The PDG [29] gives $M^{\rm phys}_{D^*_{s0}} = 2.3178(5)$GeV, $M^{\rm phys}_{D^*_s} = 2.1122(4)$GeV and $M^{\rm phys}_D = 1.867240(35)$GeV, the average of M_{D^0} and M_{D^+}. As with all other z expansions in this thesis, we take $N = 3$.

We must include valence light quark mass dependence, in order to link data from our ensembles with physical light quark masses (sets 1, 2 and 3) and those with $m_s/m_l = 5$, and extrapolate to the physical light quark limit. This is achieved using chiral perturbation theory, which we outlined in Section 4.4.2. In this case (following the notation in [1]) we parameterise this dependence using $1 + L(m_l)$, an overall multiplicative factor in the form factor expressions, defined,

$$L(m_l) = -\frac{9g^2}{8} x_\pi (\log x_\pi + \delta_{\rm FV}), \tag{5.10}$$

where we take the $DD^*\pi$ coupling $g = 0.570(6)$ [148], and x_π and $\delta_{\rm FV}$ are discussed in more detail in Section 4.4.2. In this case we write $x_\pi = m_l/(5.63 m_s^{\rm tuned})$. Here we cover a small range of light mass values, so we only include the terms of Equation (4.68) associated with the π meson

[3] Though other cases are tested below.

mass.

In Equations (5.8) and (5.9)

$$a_n^{0,+} = (1 + \mathcal{N}_n^{0,+}) \times \sum_{j=0}^{N_j - 1} d_{jn}^{0,+} \Big(\frac{a m_c^{\text{val}}}{\pi} \Big)^{2j}, \tag{5.11}$$

where we take $N_j = 3$ and sum allows for different powers of discretisation effects in am_c^{val}. Mistuning of quark masses is accounted for using

$$\mathcal{N}_n^{0,+} = c_{s,n}^{\text{val},0,+} \delta_s^{\text{val}} + c_{l,n}^{\text{val},0,+} \delta_l^{\text{val}} + c_{s,n}^{\text{sea},0,+} \delta_s^{\text{sea}} + 2 c_{l,n}^{\text{sea},0,+} \delta_l^{\text{sea}} + c_{c,n}^{0,+} \Big(\frac{M_{\eta_c} - M_{\eta_c}^{\text{phys}}}{M_{\eta_c}^{\text{phys}}} \Big). \tag{5.12}$$

In the first four terms we use,

$$\delta_q = \frac{m_q - m_q^{\text{tuned}}}{10 m_s^{\text{tuned}}}, \tag{5.13}$$

where the tuned strange mass is obtained using $M_{\eta_s}^{\text{phys}} = 0.6885(20)$ GeV [75] via

$$m_s^{\text{tuned}} = m_s^{\text{val}} \left(\frac{M_{\eta_s}^{\text{phys}}}{M_{\eta_s}} \right)^2. \tag{5.14}$$

Once this has been calculated, we can obtain m_l^{tuned} using $m_s^{\text{tuned}}/m_l^{\text{tuned}} = 27.18(10)$ [79]. A detailed discussion of our prior choices for the z expansion follows below. Note that we trivially impose the constraint $f_+(0) = f_0(0)$ (see Section 4.4.2) by setting $d_{00}^+ = d_{00}^0$.

When it comes to evaluating our form factors at physical quark masses, in the continuum, we simply set the lattice spacing a, and mistuning terms $\mathcal{N}_n^{0,+}$ to nought. We can then obtain the form factors using the $a_n^{0,+}$ that result from this in Equations (5.8) and (5.9), combined with physical masses from the PDG [29] where required.

z expansion priors and posteriors

Table 5.3 provides the priors and posteriors for our preferred z expansion. All posteriors fall within 1σ of their priors, with the exception of $c_{s,0}^{\text{val},0}$, which is slighlty over one standard deviation. This indicates that our priors are conservative.

Our priors are chosen initally based on a_n being of order unity in the z expansion, and given values

Table 5.3: Priors and their posterior values for the z expansion fit.

Parameter	Prior	Posterior	Parameter	Prior	Posterior
d^0_{00}	0.0(2.0)	0.7292(43)	d^+_{00}	0.0(2.0)	0.7292(43)
d^0_{01}	0.0(2.0)	0.825(80)	d^+_{01}	0.0(2.0)	-0.95(10)
d^0_{02}	0.0(2.0)	0.72(50)	d^+_{02}	0.0(2.0)	1.1(1.3)
d^0_{10}	0.0(1.0)	0.094(90)	d^+_{10}	0.0(1.0)	0.088(91)
d^0_{11}	0.0(1.0)	0.98(67)	d^+_{11}	0.0(1.0)	0.05(94)
d^0_{12}	0.0(1.0)	-0.07(99)	d^+_{12}	0.0(1.0)	0.01(1.00)
d^0_{20}	0.0(1.0)	0.26(62)	d^+_{20}	0.0(1.0)	0.31(62)
d^0_{21}	0.0(1.0)	0.01(1.00)	d^+_{21}	0.0(1.0)	0.01(1.00)
d^0_{22}	0.0(1.0)	-0.001(1.000)	d^+_{22}	0.0(1.0)	0.0004(1.0000)
$c^{\mathrm{sea},0}_{s,0}$	0.00(50)	0.09(30)	$c^{\mathrm{sea},+}_{s,0}$	0.00(50)	0.10(39)
$c^{\mathrm{sea},0}_{s,1}$	0.00(50)	-0.06(50)	$c^{\mathrm{sea},+}_{s,1}$	0.00(50)	-0.007(500)
$c^{\mathrm{sea},0}_{s,2}$	0.00(50)	0.005(500)	$c^{\mathrm{sea},+}_{s,2}$	0.00(50)	-0.0006(5000)
$c^{\mathrm{sea},0}_{l,0}$	0.00(50)	-0.35(36)	$c^{\mathrm{sea},+}_{l,0}$	0.00(50)	-0.35(36)
$c^{\mathrm{sea},0}_{l,1}$	0.00(50)	0.08(47)	$c^{\mathrm{sea},+}_{l,1}$	0.00(50)	0.06(50)
$c^{\mathrm{sea},0}_{l,2}$	0.00(50)	0.003(500)	$c^{\mathrm{sea},+}_{l,2}$	0.00(50)	0.006(500)
$c^0_{c,0}$	0.00(50)	0.06(41)	$c^+_{c,0}$	0.00(50)	0.16(41)
$c^0_{c,1}$	0.00(50)	-0.02(50)	$c^+_{c,1}$	0.00(50)	-0.001(500)
$c^0_{c,2}$	0.00(50)	0.002(500)	$c^+_{c,2}$	0.00(50)	0.0001(5000)
$c^{\mathrm{val},0}_{s,0}$	0.0(1.0)	1.05(84)	$c^{\mathrm{val},+}_{s,0}$	0.0(1.0)	-0.15(93)
$c^{\mathrm{val},0}_{s,1}$	0.0(1.0)	-0.08(1.00)	$c^{\mathrm{val},+}_{s,1}$	0.0(1.0)	-0.002(1.000)
$c^{\mathrm{val},0}_{s,2}$	0.0(1.0)	0.006(1.000)	$c^{\mathrm{val},+}_{s,2}$	0.0(1.0)	-0.0002(1.0000)
$c^{\mathrm{val},0}_{l,0}$	0.0(1.0)	-0.70(71)	$c^{\mathrm{val},+}_{l,0}$	0.0(1.0)	-0.70(71)
$c^{\mathrm{val},0}_{l,1}$	0.0(1.0)	0.17(95)	$c^{\mathrm{val},+}_{l,1}$	0.0(1.0)	0.12(99)
$c^{\mathrm{val},0}_{l,2}$	0.0(1.0)	0.007(0.999)	$c^{\mathrm{val},+}_{l,2}$	0.0(1.0)	0.01(1.00)

of 0(1). We then tighten the priors pertaining to sea quark mistuning, as we know this is a small effect. After such adjustments, we apply the empirical Bayes method discussed in Section 4.3.5 to broad groups of priors that cover different aspects of the fit (for example all priors $d^{0,+}_{j!=0,n}$ which pertain to discretistation effects). We find all such groups to be conservative, except $d^{0,+}_{0,n}$, which we find to be roughly optimal for maximising log(GBF). Taking a more conservative view, we double the width of these priors, though this has minimal effect on the final fit. This approach of taking priors of order unity, and adjusting down where they are known to be small, before checking they are conservative usuing log(GBF) is the standard method we shall use throughout this thesis to choose z expansion priors. We aim for a light touch here, only adjusting prior widths down if the empircal Bayes study suggets widths less than half of what they are.

We do not include in Table 5.3 other parameters which do technically appear as priors in the fit, but have very well informed values (i.e. not priors we have chosen ourselves). For example the physical meson masses, such as M_D^{phys}. Because such priors are known to a very high degree of accuracy, the fit tends not to move them at all - all such posteriors agree with their priors to within 1σ and most are unchanged by the fit.

Our preferred fit has a χ^2/d.o.f. of 0.666, for 64 degrees of freedom, corresponding to a Q value of 0.982. log(GBF) = 205.35. The continuum form factors resulting from this fit, as well as tests of the fit stability against reaosnable variations in the fit form and parameters will be given below in Section 5.3.

5.2.5 A more agnostic approach; cubic splines

As an affirmation of the suitability of our z expansion, we also experiment with the use of a model independent cubic spline fit to fit the form factors directly in q^2 space. This is a much more intuitive approach, and includes no assumptions about the shape of the form factors, or choices regarding, for example, t_0. We use a Steffen spline [149], which is straightforwardly implemented using lsqfit [85].

After treating the pole and chiral logarithm identically to those above, we start with the basis functions $g_0^{0,+}(q^2)$, which will constitute our continuum form factors, and add to these other splines $g_i^{0,+}(q^2)$ to describe the same discretisation effects and quark mistuning terms that we

considered in our z expansion.

$$f_0(q^2) = \frac{1 + L(m_l)}{1 - \frac{q^2}{M^2_{D^*_{s0}}}} \Big(g_0^0(q^2) + \mathcal{N}^0 + \sum_{j=1}^{N_j - 1} g_j^0(q^2) \Big(\frac{am_c}{\pi} \Big)^{2j} \Big),$$
$$f_+(q^2) = \frac{1 + L(m_l)}{1 - \frac{q^2}{M^2_{D^*_s}}} \Big(g_0^+(q^2) + \mathcal{N}^+ + \sum_{j=1}^{N_j - 1} g_j^+(q^2) \Big(\frac{am_c}{\pi} \Big)^{2j} \Big). \tag{5.15}$$

We take $N_j = 2$ but taking $N_j = 3$ gives no significant difference. For $\mathcal{N}$ we use further spline functions:

$$\mathcal{N}^{0,+} = g_s^{\text{val},0,+} \delta_s^{\text{val}} + g_l^{\text{val},0,+} \delta_l^{\text{val}} + g_s^{\text{sea},0,+} \delta_s^{\text{sea}} + 2 g_l^{\text{sea},0,+} \delta_l^{\text{sea}} + g_c^{0,+} \Big(\frac{M_{\eta_c} - M_{\eta_c}^{\text{phys}}}{M_{\eta_c}^{\text{phys}}} \Big), \tag{5.16}$$

where the δs in $\mathcal{N}$ take the same form as in the z expansion (Equation (5.13)). We choose the knots in our splines to span the q^2 range of our data evenly, giving $\{-3.25, -1.5, 0.25, 2.0\}$ GeV, and provide priors on the spline values at these knots. In the continuum limit, the form factors (excluding $L(m_l)$ and the pole) are represented by $g_0^{0,+}$, and we give this a prior of 0.75(15), based on the raw data with the pole removed, (see Figure 4 in [1]). Priors for g_j, g_s and g_l are taken as 0.0(5) and for g_c we take 0.0(1.0), using the same empiricaly Bayes method as above. Slightly tighter priors are prefrerred here as the fit form is much less constrained. We also know that discretisation effects are small in the HISQ action. We will discuss and compare the outcome from the two fit forms below.

5.3 Form factor results

The form factor results in q^2 are shown as coloured bands in Figure 5.4. These results represent our preferred z expansion fit in the continuum and at physical quark masses. Data from each of the eight ensembles is included as data points, joined by black lines to guide the eye. The lines through these data points are the fit result to those points, that is to say the final fit evaluated at finite lattice spacing with quark mistunings in place. The fact that all these fits are within error of the continuum result is a testament to the tuning of the quark masses, and the strong suppression of discretisation effects in the HISQ formalism. There are 64 data points in this fit, and 180 fit parameters, each with a prior.

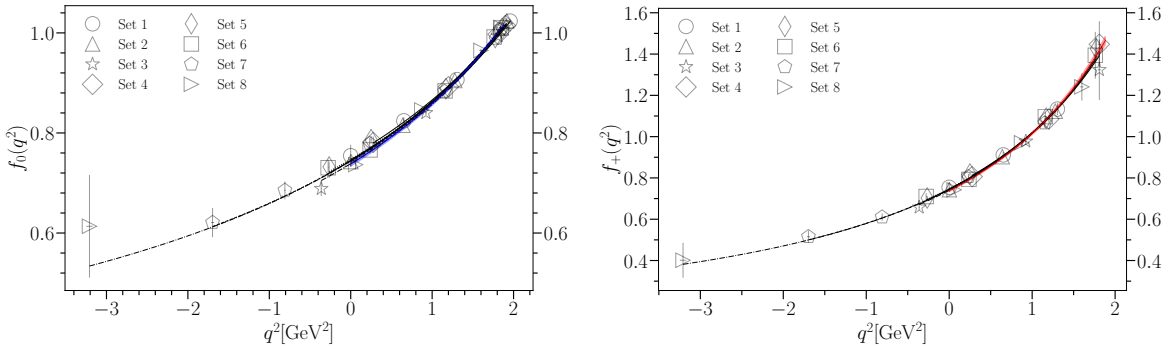

Figure 5.4: $D \to K$ f_0 and f_+ form factors in q^2 space using the z expansion, our preferred fit. Data points from the eight gluon ensembles are included, and are joined by lines corresponding to the fit at that lattice spacing (Equations (5.8) and (5.9)). The continuum, physical point results are the coloured bands.

We compare this result with our spline fit in Figure 5.5. We find that the spline fit agrees well, but is slightly less accurate. This is because the splines explore non-analytic functions of q^2, which we do not expect to contribute. Nonetheless, the good agreement with this model independent method is an excellent affirmation of our preferred z expansion fit.

We test the stability of our fit against a variety of changes in Figure 5.6. Our preferred fit is stable against all tests, and in particular we note that setting $L(m_l) = 0$ has almost no effect on the fit. See figure caption and [1] for finer details of the tests conducted.

A comparison of our results with previous work [94, 138, 139, 140] is given in Figure 5.7 (we include information on reconstructing our form factors below). We see broad agreement and a significant improvement in uncertainty over previous work, with the possible exception of a slight (2σ) tension with ETMC's q^2_{max} form factors [139] (see Figure 5.7 and Table 7.8 below for numerical values).

We provide an error breakdown of the form factors in q^2 space in Figure 5.8 (left). We see that the main contribution to uncertainty in both cases is statistics, with discretisation effects and quark mistunings playing a relatively minor role across the q^2 range. This is encouraging, as it suggests that we have not reached the end of the line with this method, and the form factor uncertainty could be improved somewhat by simply computing data on more configurations, perhaps reducing uncertainty by a factor of two in places with enough statistics. The right hand side of Figure 5.8 further breaks down the statistical error into contributions from each ensemble.

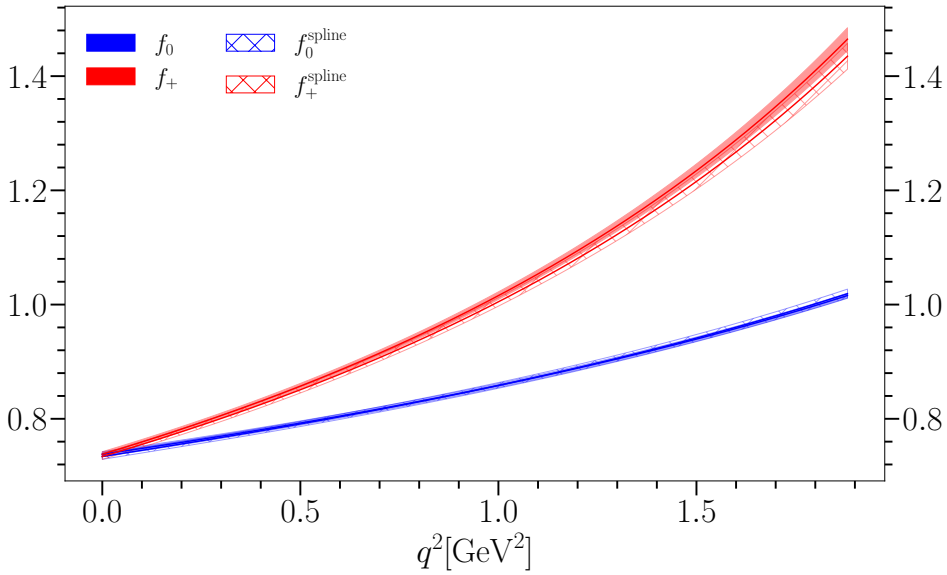

Figure 5.5: Comparison between our preferred z expansion fit described in Section 5.2.4 (solid bands) and the spline fit described in Section 5.2.5 (hatched bands).

The bars here represent the statistical varience contributions from each of the eight ensembles, normalised by the total statistical varience. Unlike in the $B \to K$ calculation (see both $B \to K$ and $D \to K$ from this heavy-HISQ calculation in Figure 7.13), there is no clear pattern to the dominant uncertainties here with the contributions well spread across the ensembles, particularly at $q^2 = 0$. Nevertheless, a good place to start improving statistics at q^2_{max} would be on fine-physical (set 3) or coarse-5 (set 5).

Our continuum form factors can be reconstructed from the z expansion coefficients, pole masses, and chiral logaritm term given in Table 5.4. The correlation matrix for these values is also provided.

5.3.1 Form factor shape

One simple and important test of QCD is to check the shape of our f_+ form factor in q^2 space against that measured experimentally. This is easily achieved by plotting a_1^+/a_0^+ and a_2^+/a_0^+. Unfortunately, experimentalists use a different (BGL [100]) form of the z expansion, featuring a rather baroque 'outer function', which means we cannot simply use the a_n^+ values extracted

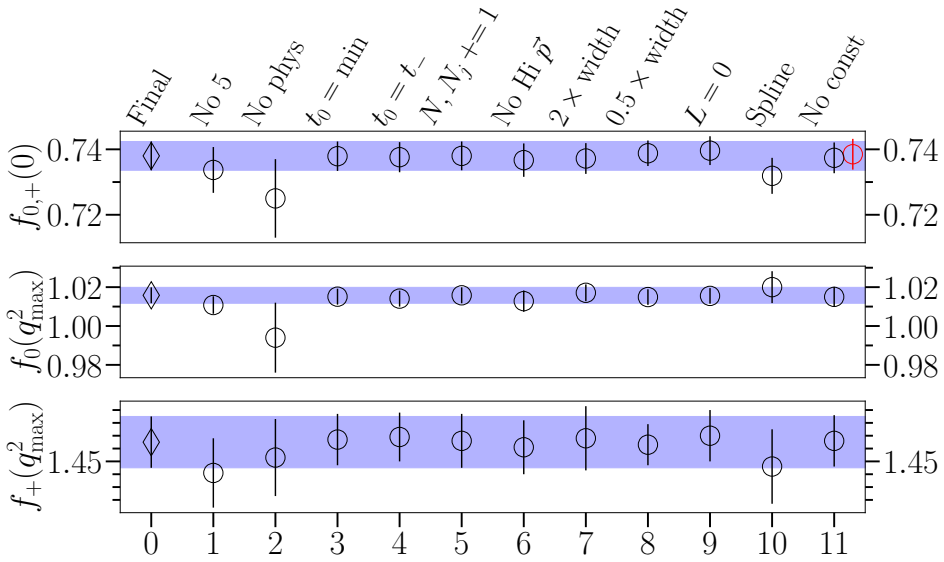

Figure 5.6: Stability test of the z-expansion fit; 0 marks our final result. Test 1 removes all the results from gluon field configurations with $m_s/m_l = 5$, so that only sets 1, 2 and 3 remain. Test 2 removes the results from sets 1, 2 and 3 and fits the others. Test 3 takes t_0 in the q^2 to z mapping to the 'minimum' prescription described in Section 4.4.2. Test 4 sets t_0 to t_-. Test 5 includes an extra term in the sums over n up to N and over j up to N_j (Equations (5.8), (5.9) and (5.11)). Test 6 removes the highest momentum data point for each gluon field ensemble (and highest two on set 7 so that there are no results included with $q^2 < 0$). Test 7 doubles the width of all 'd' priors (this decreased the Gaussian Bayes Factor), and 8 halves them. Test 9 sets the logarithmic factor $L(m_l)$ to zero (Equation (5.10)). Test 10 shows the results of a completely different kind of fit, a cubic spline fit in q^2 discussed in Section 5.2.5. Test 11 removes the $f_0(0) = f_+(0)$ constraint, in this case the black point is $f_0(0)$ and the red is $f_+(0)$.

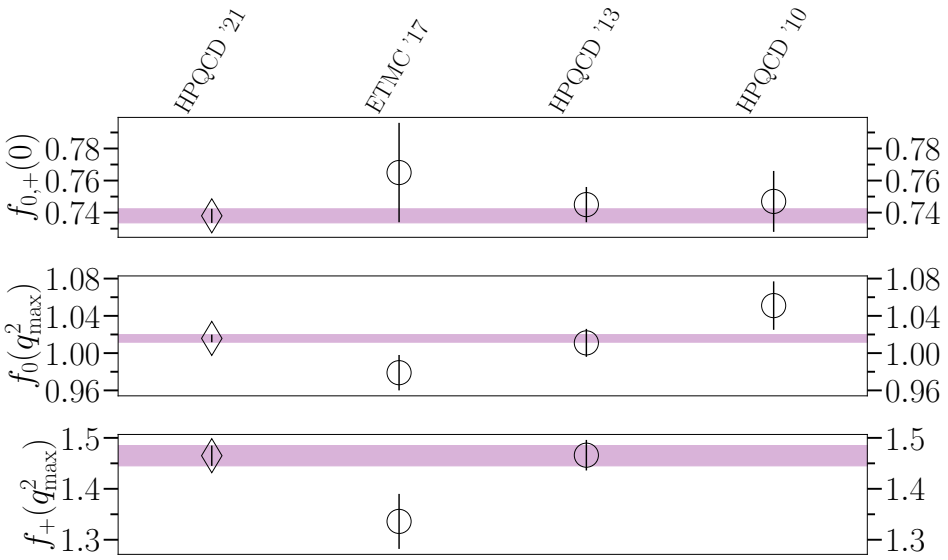

Figure 5.7: Comparison of our lattice form factors at $q^2 = 0$ and $q^2_{\rm max}$ with earlier lattice QCD calculations. The points marked 'HPQCD '10' are from [94]; the points marked 'HPQCD '13' from [138] and the points marked 'ETMC '17' from [139, 140]. A preliminary analysis of the scalar form factor in [145] gives $f_0(0) = 0.768(16)$, but we have not plotted that point. Our new results are labelled 'HPQCD '21' and demonstrate a significant improvement in uncertainty over earlier values.

Table 5.4: Values and uncertainties for the fit coefficients $a_n^{0,+}$, pole masses, and chiral logarithmic term $(1 + L(m_l))$ for the reconstruction of our form factors at the physical point as a function of q^2 from Equations (5.8) and (5.9). The correlation matrix between these parameters is given below the row with their values. The pole masses are in GeV. The pole masses and $L(m_l)$ are very slightly correlated due to the way the fit function is constructed. These correlations are too small to have any meaningful effect on the fit, but we include them for completeness in reconstructing our results.

a_0^0	a_1^0	a_2^0	a_0^+	a_1^+	a_2^+	$M^{\rm phys}_{D_{s0}^*}$	$M^{\rm phys}_{D_s^*}$	$(1 + L(m_l))$
0.7292(43)	0.825(80)	0.72(50)	0.7292(43)	-0.95(10)	1.1(1.3)	2.31780(50)	2.11220(40)	1.01200(26)
1.00000	0.73103	0.51757	1.00000	0.29251	0.02299	-0.00023	-0.00005	-0.04904
	1.00000	0.90723	0.73103	0.49742	0.01488	-0.01619	0.00001	-0.00795
		1.00000	0.51757	0.52335	0.00600	0.00368	0.00003	-0.00222
			1.00000	0.29251	0.02299	-0.00023	-0.00005	-0.04904
				1.00000	0.49065	0.00007	-0.01488	0.00553
					1.00000	0.00019	0.00362	-0.00017
						1.00000	-0.00000	-0.00000
							1.00000	0.00000
								1.00000

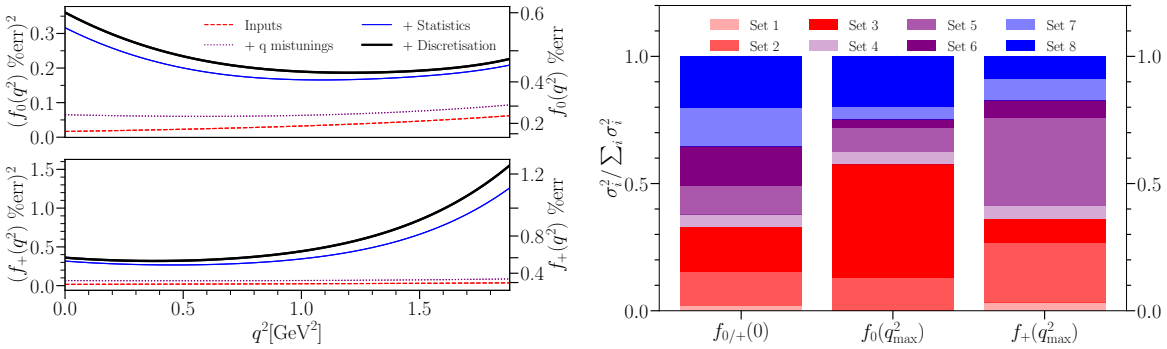

Figure 5.8: Left: uncertainties for f_0 and f_+ (for $m_u = m_d$) as a function of q^2. The red line 'Inputs' shows the uncertainties coming from fixed inputs, such as experimental meson masses used in the analysis. The purple line 'q mistunings' adds in uncertainties arising from mistuning of valence and sea quark masses. The blue 'Statistics' line further adds the statistical uncertainties from the lattice results (correlator fits). Finally, the black line ('Discretisation') gives the total uncertainty, now including the contribution from discretisation effects. These uncertainties add in quadrature, so we plot the squared percentage error and include an axis showing the corresponding percentage error on the right for clarity. Right: breakdown of the contributions to statistical uncertainty of the form factors at extremal q^2 values, from each of the eight ensembles (Table 3.1). The variance σ_i^2 from each ensemble is added linearly, and normalised by the total variance.

from our fit. The BGL form is,

$$f_+(q^2) = \frac{1}{z(q^2, t_0 = M_{D_s^*}^2)\phi(q^2)} \sum_{n=0}^{N-1} a_n^+ z^n, \tag{5.17}$$

where the outer function,

$$\begin{aligned}\phi(q^2, t_0) = \sqrt{\frac{\pi}{3}} m_c \left(\frac{z(q^2,0)}{-q^2}\right)^{5/2} \left(\frac{z(q^2,t_0)}{t_0 - q^2}\right)^{-1/2} \\ \times \left(\frac{z(q^2,t_-)}{t_- - q^2}\right)^{-3/4} \frac{t_+ - q^2}{(t_+ - t_0)^{1/4}}.\end{aligned} \tag{5.18}$$

The q^2 to z mapping uses $t_0 = t_+(1 - (1 - t_-/t_+)^{1/2})$ (for $t_{+/-} = (M_D \pm M_K)^2$). This is the prescription that minimises the maximum value of z over the q^2 range of the decay. The parameter $m_c = 1.25$GeV. We apply the fit form of Equation (5.17) to our form factors at the physical point, generating synthetic data. We used 20 evenly spaced points but changing the number of points makes no difference. This gives us the parameters a_n^+ for this fit form, along with their correlation matrix and these are the values plotted on the right of Figure 5.9.

The left hand side of Figure 5.9 shows this refitted data alongside the original results, as well as the ratio of the two, confirming that we are accurately fitting the data points and the uncertainties are unaffected.

The right hand side of Figure 5.9 compares our form factor shapes directly with experiment [150, 151, 152, 153], and shows good agreement. The agreement with the 2016 HFLAV average [153] is particularly compelling, and this is a strong test of QCD.

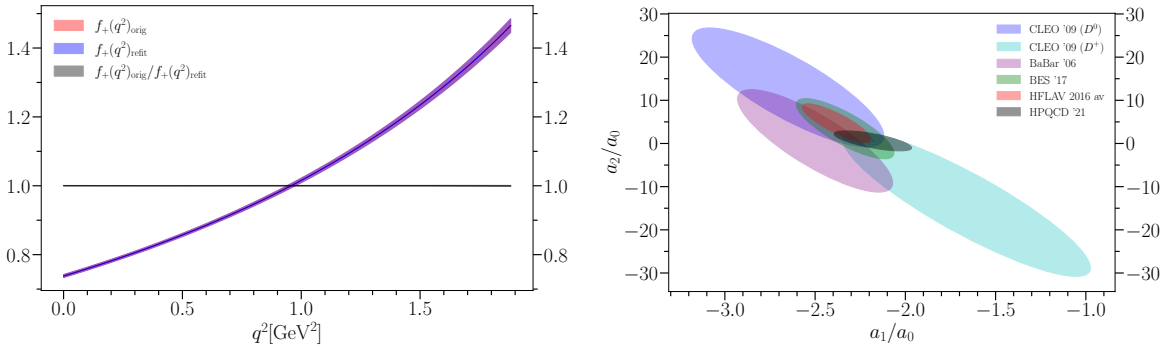

Figure 5.9: Left: the original f_+ form factor, as well as the result using the refitting procedure described in Section 5.3.1. The black line (and grey uncertainty band, barely visible) is the ratio of the two. We see that the refitting reproduces the original form factor and its uncertainty accurately. Right: comparison of the shape of $f_+(q^2)$ for $D \to K$ expressed in terms of ratios of the z-expansion coefficients a_1^+ and a_2^+ to a_0^+. Ellipses give the 68% confidence limits ($\Delta\chi^2 = 2.3$). Experimental results are from [150, 151, 152, 153]. CLEO results are for $D^0 \to K^- e^+ \nu_e$ (dark blue) and $D^+ \to \bar{K}^0 e^+ \nu_e$ (light blue); all other experimental data is for $D^0 \to K^- e^+ \nu_e$. The HFLAV experimental average [153] is given as the red ellipse. Our results here are given by the black ellipse, showing good agreement.

5.4 Extracting the differential decay rate and V_{cs} determination

The full expression for the differential decay rate for the $D \to K\ell\nu$ decay (see Section 2.3.3) is given by Equation (2.38):

$$\frac{d\Gamma}{dq^2} = \frac{(G_F \eta_{\rm EW})^2}{24\pi^3} |V_{cs}|^2 (1-\epsilon)^2 (1+\delta_{\rm EM}) \left[|\vec{p}_K|^3 (1+\frac{\epsilon}{2}) |f_+(q^2)|^2 + |\vec{p}_K| M_D^2 \left(1 - \frac{M_K^2}{M_D^2}\right)^2 \frac{3\epsilon}{8} |f_0(q^2)|^2 \right] \tag{5.19}$$

where $G_{\rm F} = 1.1663787(6) \times 10^{-5}$ GeV^{-2} [29] is the Fermi constant and the f_0 contribution is highly suppressed by the lepton mass m_ℓ through $\epsilon = m_\ell^2/q^2$.

5.4.1 Isospin breaking, $\eta_{\rm EW}$ and $\delta_{\rm EM}$

Equation (5.19) contains a number of subtleties which require unpacking. The first is to do with meson charges. In our form factor calculation, we used degenerate light quarks $m_l = (m_u + m_d)/2$, meaning that our resulting form factors are neither specific to $D^0 \to K^-$ nor

$D^+ \to K^0$, but rather a decay with average masses $M_D = (M_{D^0} + M_{D^+})/2 = 1.867240(35)\text{GeV}$ and $M_K = (M_{K^0} + M_{K^-})/2 = 0.4956440(92)\text{GeV}$ [29]. In our V_{cs} determination, we will wish to use experimental data which is explicitly from one decay or the other. As such, we will use the correct D and K masses in Equation (5.19) for the decay we are interested in, and we also include an uncertainty on the form factors to allow for this isospin breaking. We can test the effect of having a different light quark mass (corresponding to u or d) by changing the ratio $m_s^{\text{tuned}}/m_l^{\text{tuned}} = 27.18(10)$ [79] in our code. We take $m_d/m_u \approx 2$ [29] so that $m_u/m_s \approx 2/(3 \times 27.18)$ and $m_d/m_s \approx 4/(3 \times 27.18)$ and compare to our original form factors. This is an overestimate of any effects as it changes both valence and sea light quark masses. The result is a shift of, at most, 0.15% in our form factors, and so we add this conservative uncertainty across the q^2 range to both form factors.

We also include extra corrections η_{EW} and δ_{EM}, which have previously been ignored[4], in Equation (5.19). The first of these corrections,

$$\eta_{\text{EW}} = 1 + \frac{\alpha_{\text{QED}}}{\pi} \log\left(\frac{M_Z}{M_D}\right) = 1.009(2), \tag{5.20}$$

accounts for universal short-distance corrections to G_{F} from box diagrams in the standard model [154]. The uncertainty here allows for a factor of two variation in the central value of M_D.

Experimental data is treated using PHOTOS [136], which provides an analysis of radiative bremsstrahlung corrections, which are then removed, however, final state electromagnetic interactions still need attention. The second correction, δ_{EM}, is present to account for these possible electromagnetic interactions in the final state, and so is expected to be larger in the case of the charged kaon. We take our cue from the semileptonic decay $K \to \pi$, where electromagnetic corrections have been calculated [155], and are found to be $\delta_{\text{EM}} \approx 0$ for neutral and $\delta_{\text{EM}} \approx 0.7\%$ for charged final state mesons. We take independent $\delta_{\text{EM}} = \pm 0.5\%$ and $\delta_{\text{EM}} = \pm 1.0\%$ uncertainties for these cases respectively.

5.4.2 Tests of flavour universality

One test which we can conduct easily using Equation (5.19) is a check of lepton flavour universality. In the SM, aside from mass effects, the three lepton flavours are treated identically, and

[4] And indeed are negligible on the scale of previous V_{cs} uncertainties.

we can demonstrate this by taking the ratio,

$$R\mu/e = \frac{\int \frac{d\Gamma^{\mu}}{dq^2} dq^2}{\int \frac{d\Gamma^{e}}{dq^2} dq^2} \tag{5.21}$$

Experimental work in B decays has shown hints of disagreement here [9] (see discussion in Chapter 7), and the theoretical calculation is extremely precise, as most uncertainties cancel when the ratio is taken.

Our calculations yields $R\mu/e = 0.9779(2)_{\text{latt}}(50)_{\text{EM}}$, the uncertainty being dominated by that from δ_{EM}. This agrees well with the less precise experimental result of 0.974(7)(12) from BES [156], and is in slight tension with the ETMC lattice determination of $0.975(1)_{\text{latt}}$ [140], if we neglect the EM uncertainty.

We can also plot $R_{\mu/e}$ in q^2 space[5], and compare with binned experimental results. This is what is shown on the left of Figure 5.10, where we again compare with BES [156]. Our uncertainties here are plotted (black line), but are too small to be seen, whilst the experimental uncertainties are at present too large to discriminate between our SM model result, and our result with the added effect of a possible new physics scalar coupling, given as examples in red and blue. These red and blue lines are just one possible example of new physics, in this case we show the effect of a scalar coupling in the μ sector, $C_S^{(\mu)}$, which would multiply a contribution to the Lagrangian with a scalar $\bar{s}c$ current multiplied by a $\bar{\nu}_{\mu}\mu$. This would feature in Equation (5.19), multiplying the scalar form factor term by $|1 + C_S^{(\mu)} q^2/(m_{\mu}(m_s - m_c))|^2$ [157]. We define,

$$\zeta_S \equiv \frac{C_S^{(\mu)}}{m_s - m_c} \tag{5.22}$$

and show values of $\zeta_S = \pm 0.1$ GeV^{-1}, which roughly covers the possible range of agreement with the BES results.

Another source of SM tests is to be found in angular variables. In particular, we are able to access the forward backward asymmetry

$$A_{\text{FB}}^{(\ell)}(q^2) = -\frac{b_{\ell}}{d\Gamma^{(\ell)}/dq^2} \tag{5.23}$$

[5] I.e. don't integrate.

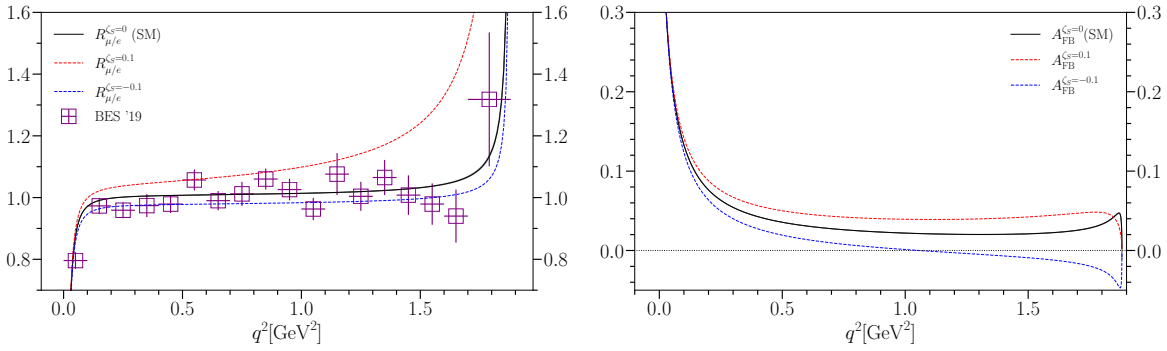

Figure 5.10: Left: lepton flavour universality tests in $D \to K$ decay. The solid black curve as a function of q^2 shows the Standard Model ratio of branching fractions for a muon in the final state to that for an electron obtained from our form factors using Equation (5.19). The width of the curve gives the (very small) uncertainty from our results. Possible QED effects are not included here. The points, with error bars, are from the BES experiment [156]. For illustration the red and blue dashed lines show what the curve would look like in the presence of a new physics scalar coupling for the μ case (see Equation (5.23) for definition of ζ_S). Right: the forward-backward asymmetry of the muon produced in $D \to K$ decay. This is defined with respect to the angle θ_ℓ in the W rest-frame. The solid black line shows the SM result derived from our form factors, including the lattice QCD uncertainty but ignoring any uncertainty from possible QED corrections. For illustration the red and blue dashed lines show what the curve would look like in the presence of a new physics scalar coupling for the μ case (see text and Equation (5.23)).

where

$$\frac{d\Gamma^{(\ell)}}{dq^2 d\cos\theta_\ell} = a_\ell(q^2) + b_\ell(q^2)\cos\theta_\ell + c_\ell(q^2)\cos^2\theta_\ell, \tag{5.24}$$

and b_ℓ is defined in [157]. Here, θ is the angle between the charged lepton momentum in the W rest frame and the W momentum in the D rest frame (see [1] for diagram). $A_{\rm FB} = 0$ for massless leptons, because only the helicity zero component of the W can contribute. We plot $A_{\rm FB}^{(\mu)}$, again including the effects of the scalar coupling $\zeta_S = \pm 0.1$ GeV^{-1}, on the right hand side of Figure 5.10. In this case we can see that negative values of ζ_S actually change the sign of $A_{\rm FB}^{(\mu)}$ at large values of q^2.

5.4.3 V_{cs} determination

Below, we will describe three semi-independent methods for determining $|V_{cs}|$. These methods all use the same form factors (our form factors), but they each make use of different parts of the q^2 range, and different combinations of experimental results, allowing access to different sets of the four available channels: $D^+ \to \bar{K}^0 e^+ \nu_e$, $D^0 \to K^- e^+ \nu_e$, $D^+ \to \bar{K}^0 \mu^+ \nu_\mu$ and $D^0 \to K^- \mu^+ \nu_\mu$.

Via the binned differential decay rate

Using Equation (5.19), we can construct the differential decay rate $d\Gamma/dq^2$ across the whole q^2 range from our form factors. We can isolate $|V_{cs}|$ by comparing this with the experimentally determined $d\Gamma/dq^2$, which is available in multiple q^2 bins, for multiple experiments [150, 151, 152, 158].

We use experimental results for which a covariance matrix is provided for the partial rates between q^2 bins. We add covariance matrices for statistical and systematic uncertainties where they are provided separately (effectively adding the uncertainties in quadrature). In some cases an overall uncertainty on each bin is given along with the percentage breakdown into systematic and statistical uncertainty. We use this, along with the correlation matrices given, to obtain the separate covariance matrices and add them.

CLEO results are taken from [150], where both $D^0 \to K^- e^+ \nu_e$ and $D^+ \to \overline{K}^0 e^+ \nu_e$ differential rates are measured and the correlations between them given. Partial rates were taken from Table V, and $\sigma_i^{\rm stat}$, $\sigma_i^{\rm syst}$ and their covariance matrices were calculated using these, the percentage error breakdowns in Tables VII and VIII and the correlation matrices in Tables XVI and XVII. These

covariance matrices are then easily included in our calculation using the gvar package [87]. Our determination of V_{cs} on a bin-by-bin basis is shown for the CLEO results in Figure 5.11. The fit for the weighted average gives a χ^2/d.o.f. of 0.64 in the D^0 case and 1.7 in the D^+ case. In both cases there are nine degrees of freedom. The q^2 bins with the minimum total uncertainty are at the small q^2 end of the range, where the experiment is most accurate. BaBar results are taken from [151]; these are for the $D^0 \to K^- e^+ \overline{\nu}_e$ decay normalised by the branching fraction for $D^0 \to K^- \pi^+$. Table II gives the normalised decay distribution and total correlation matrix. The leading diagonal values of the matrix give the σ_i. The distribution has been normalised so that the sum over all bins equals unity. A value is also given for

$$R = \frac{\mathcal{B}(D^0 \to K^- e^+ \nu_e)}{\mathcal{B}(D^0 \to K^- \pi^+)}, \tag{5.25}$$

which is included in the correlation matrix. Using this value, and multiplying by the global average for $\mathcal{B}(D^0 \to K^- \pi^+) = 0.03950(31)$ [29], we determine $\mathcal{B}(D^0 \to K^- e^+ \nu_e)$. This allows us to extract the branching fractions per bin from the decay distribution and convert these to partial rates by dividing by the D^0 lifetime $\tau_{D^0} = 4.101(15) \times 10^{-4}$ ns [29]. We drop the largest q^2 bin from our weighted average fit (because it is equal to one minus the sum of the others from the normalisation constraint). We include the normalisation uncertainty after averaging to avoid normalisation bias.

Our determination of V_{cs} from the BaBar results is shown in Figure 5.12 and has a χ^2/d.o.f. of 0.9 with nine degrees of freedom.

BES results are taken from [152] for the D^0 decay channel. The data can be found in Table V, and the breakdown of the percentage errors and correlation matrices for systematic and statistical uncertainty are given in Tables IX and XI. BES results for the D^+ channel are given in [158] (Table VI). Our determination of V_{cs} on a bin-by-bin basis is shown for these two sets of BES results in Figures 5.12 and 5.13, with χ^2/d.o.f. 1.1 (d.o.f. = 18) and 0.9 (d.o.f. = 9) respectively.

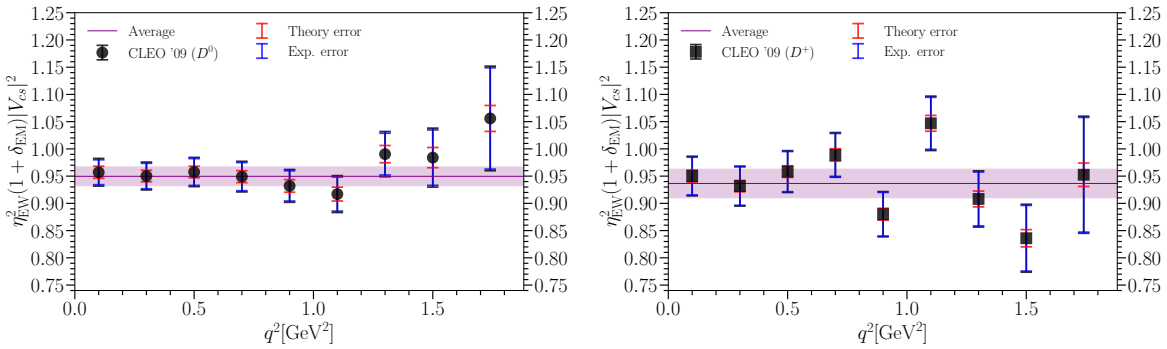

Figure 5.11: Plots of the determination of $|\eta_{\rm EW}V_{cs}|^2(1+\delta_{\rm EM})$ per q^2 bin for CLEO D^0 and D^+ results [150]. The total uncertainty for each bin is given in black and this is broken down into experimental (blue) and theoretical (red) contributions, the latter coming from our form factors. Each data point is centred on the q^2 bin it corresponds to. Note that the uncertainties are correlated between q^2 bins. The purple band gives the weighted average for these data points, with all correlations included.

We integrate over each bin (q_i^2, q_{i+1}^2) using the trapezoidal rule to obtain.

$$\begin{aligned}\Delta_i\Gamma =& \int_{q_i^2}^{q_{i+1}^2} \frac{d\Gamma}{dq^2} dq^2 \\ =& \frac{G_F^2|\eta_{\rm EW}V_{cs}|^2(1+\delta_{\rm EM})}{24\pi^3} \int_{q_i^2}^{q_{i+1}^2} dq^2 \Bigg[|\vec{p}_K|^3(1-\epsilon)^2(1+\frac{\epsilon}{2})|f_+(q^2)|^2 \\ &+ |\vec{p}_K|(1-\epsilon)^2 M_D^2 \left(1-\frac{M_K^2}{M_D^2}\right)^2 \frac{3\epsilon}{8}|f_0(q^2)|^2 \Bigg],\end{aligned} \tag{5.26}$$

and then compare with the experimental binned measurements $\Delta_i\Gamma$ to isolate a $|V_{cs}|$ value for each experimental q^2 bin. We then average over the bins in each experiment, and finally average these four experimental values to give a final answer for $|V_{cs}|$. Figures 5.11, 5.12 and 5.13 show the $|\eta_{\rm EW}V_{cs}|^2(1+\delta_{\rm EM})$ values for each q^2 bin for each experiment we considered, as well as their average. We then average these numbers together, not including the BES D^+ results [158], as we are unable to correlate these with the BES D^0 [152] results. The $|V_{cs}|$ value that results from the average of four other averages is

$$|V_{cs}|^{{\rm d}\Gamma/{\rm d}q^2} = 0.9663(53)_{\rm latt}(39)_{\rm exp}(19)_{\eta_{EW}}(40)_{\rm EM}, \tag{5.27}$$

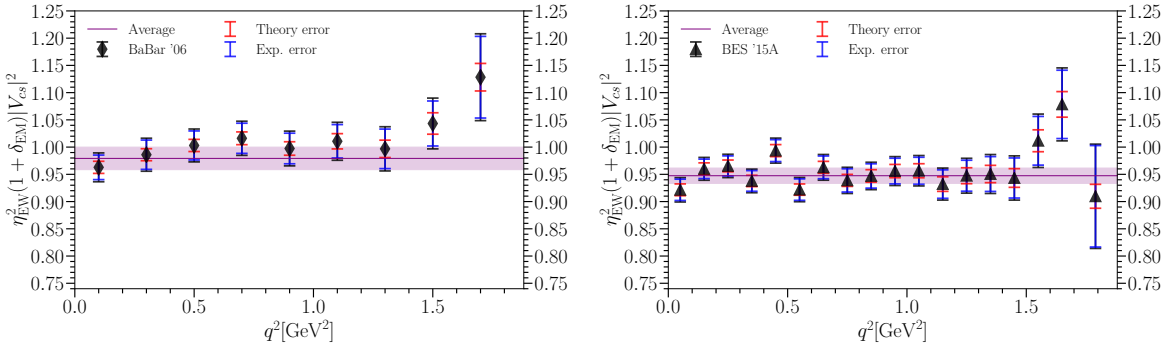

Figure 5.12: Plots of the determination of $|\eta_{\rm EW} V_{cs}|^2(1+\delta_{\rm EM})$ per q^2 bin for BaBar D^0 [151] and BES D^0 [152] results. The total uncertainty for each bin is given in black and this is broken down into experimental (blue) and theoretical (red) contributions, the latter coming from our form factors. Each data point is centred on the q^2 bin it corresponds to. Note that the uncertainties are correlated between q^2 bins. The purple band gives the weighted average for these data points, with all correlations included.

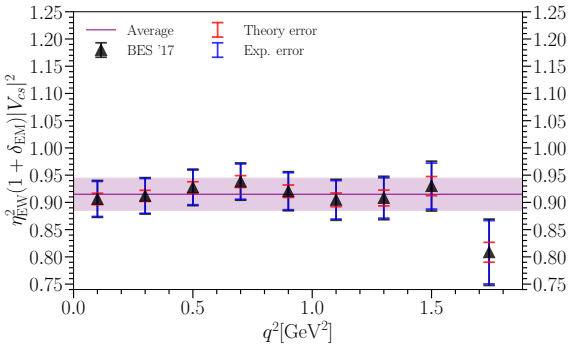

Figure 5.13: Plot of the determination of $|\eta_{\rm EW} V_{cs}|^2(1+\delta_{\rm EM})$ per q^2 bin for BES D^+ results [158]. The total uncertainty for each bin is given in black and this is broken down into experimental (blue) and theoretical (red) contributions, the latter coming from our form factors. Each data point is centred on the q^2 bin it corresponds to. Note that the uncertainties are correlated between q^2 bins. The purple band gives the weighted average for these data points, with all correlations included.

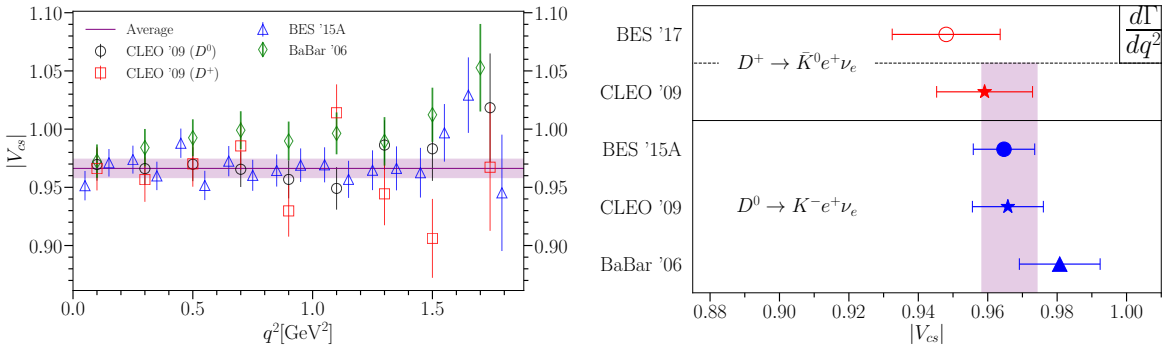

Figure 5.14: Left: plot of $|V_{cs}|$ per bin for CLEO, BaBar and BES results from [150, 151, 152]. Each data point is centred on the q^2 bin it corresponds to and the error bars plotted include the uncertainties from $\eta_{\rm EW}$ and $\delta_{\rm EM}$. The purple line and band give the result from our total weighted average for $|V_{cs}|^2$, with all correlations included. Right: comparison plot of the determination of $|V_{cs}|$ using the differential decay rate for CLEO, BaBar and BES results from [150, 151, 152, 158] for D^0 and D^+ decays. The purple band gives the total weighted average for V_{cs}, not including the BES '17 result.

where the uncertainties are from our form factors, experimental data, $\eta_{\rm EW}$ and $\delta_{\rm EM}$ respectively, and the result is shown in Figure 5.14. The left hand side of this figure shows the binned $|V_{cs}|$ values which contribute to the final result, as well as $|V_{cs}|$, the purple band. The right hand side shows the averages over bins for each of the five experiments, as well as the final $|V_{cs}|$, an average of the four results shown below the dotted line.

As a note on averaging with correlations: it's important in situations when averaging over correlated data to proceed with caution. One thing to be aware of is D'Agostini or normalisation bias [159]. This appears in certain situations when taking correlated, weighted averages of data which contains a common factor. In our case, this is why we average $|\eta_{\rm EW} V_{cs}|^2(1+\delta_{\rm EM})$ over q^2 bins, only dividing through by the $\eta^2_{\rm EW}(1+\delta_{\rm EM})$, which is a overall factor applied to each experiment, after the bin averaging[6]. Our final result is obtained by a fit (using [85]) to the four averages, with $|V_{cs}|$ as a fit parameter, as well as $\eta^2_{\rm EW}$ and $(1+\delta^{e,0/+}_{\rm EM})$, the latter two parameters[7] with priors of 1.00(1) and 1.000(5) for the charged and neutral K cases respectively. This fit is formally equivalent to a simple weighted average, but allows us to obtain posteriors for the corrections $\eta^2_{\rm EW}$ and $(1+\delta^{e,0/+}_{\rm EM})$, which we find agree well with their priors. $\eta^2_{\rm EW}$ sits on its prior,

[6] Due to a subtlety with the BaBar data, (see [1]), we have an overall normalisation for this, which we also keep out of the average over bins.

[7] A separate fit parameter is assigned to each of the charges.

Table 5.5: Total width for $D \to K$ semileptonic decay up to a factor of $|\eta_{EW} V_{cs}|^2(1 + \delta_{EM})$, determined from our form factors. We give results for all 4 modes that we consider. They differ slightly in the parent and daughter meson masses and in the mass of the lepton in the final state; these affect the kinematic factors in the differential rate and the end-points of integration for the total width. These values can be combined with experimental values of the relevant branching fraction and D meson lifetime to determine $|V_{cs}|$.

	$\Gamma/(\lvert\eta_{EW} V_{cs}\rvert^2(1 + \delta_{EM}))$ (ns^{-1})
$D^+ \to \overline{K}^0 \mu^+ \nu_\mu$	88.30(99)
$D^+ \to \overline{K}^0 e^+ \nu_e$	90.3(1.0)
$D^0 \to K^- \mu^+ \nu_\mu$	87.57(98)
$D^0 \to K^- e^+ \nu_e$	89.5(1.0)

whilst $(1 + \delta_{\text{EM}}^{e,+}) = 1.0022(92)$ and $(1 + \delta_{\text{EM}}^{e,0}) = 0.9995(49)$.

This is our preferred method for determining $|V_{cs}|$, as it uses information from across the full q^2 range.

Via the total branching fraction

A separate method for determining $|V_{cs}|$ is to integrate Equation (5.19) across the full q^2 range to determine $\Gamma/(|\eta_{\text{EW}} V_{cs}|^2(1 + \delta_{\text{EM}}))$ (shown in Table 5.5), and then convert to the branching fraction $\mathcal{B}$ using the appropriate D meson lifetime [29].

This gives us four different $\mathcal{B}^{e/\mu,0/+}/(|\eta_{\text{EW}} V_{cs}|^2(1 + \delta_{\text{EM}}^{e/\mu,0/+}))$ values, for the different channels. We can then compare these directly with experimental $\mathcal{B}$ values to isolate $|V_{cs}|$. In this case, there are more experimental determinations, covering all four of the charge-lepton combinations. In our calculation, we assume 100% correlation of systematic uncertainties on results from the same experiments, a conservative move as no correlation data is available.

Again, we do the averaging by way of a fit, including the parameters η_{EW}^2 and $(1 + \delta_{\text{EM}}^{e/\mu,0/+})$, as well as the D lifetimes, and our values for $\mathcal{B}^{e/\mu,0/+}/(|\eta_{\text{EW}} V_{cs}|^2(1 + \delta_{\text{EM}}^{e/\mu,0/+}))$. The priors for $(1 + \delta_{\text{EM}}^{e/\mu,0/+})$ are again 1.00(1) and 1.000(5) for the charged and neutral K cases respectively, but now we have four distinct fit parameters, allowing for a difference in EM effects not only with kaon charge, but also with lepton flavour. We obtain posteriors $(1 + \delta_{\text{EM}}^{e,+}) = 1.0077(74)$, $(1+\delta_{\text{EM}}^{e,0}) = 0.9938(45)$, $(1+\delta_{\text{EM}}^{\mu,+}) = 1.0015(68)$ and $(1+\delta_{\text{EM}}^{\mu,0}) = 1.0038(46)$, which are somewhat more precise than their priors. Whilst the precision on these numbers means it's impossible to draw any firm conclusions here, it seems that both the binned and total branching fraction

methods lean towards $\delta_{\rm EM}^{e,+}$ being positive and $\delta_{\rm EM}^{e,0}$ negative.

The result,

$$|V_{cs}|^{\mathcal{B}} = 0.9686(54)_{\rm latt}(39)_{\rm exp}(19)_{\eta_{EW}}(30)_{\rm EM}, \tag{5.28}$$

agrees well with our differential decay rate method and is shown alongside the experimental data in Figure 5.15.

Via $|V_{cs}|f_+(0)$

Finally, many experiments perform a fit and extrapolation to $q^2 = 0$, and provide a value for $|V_{cs}|f_+(0)\eta_{\rm EW}\sqrt{1+\delta_{\rm EM}}$. These numbers have helpfully been averaged by HFLAV [129], giving $|V_{cs}|f_+(0)\eta_{\rm EW}\sqrt{1+\delta_{\rm EM}} = 0.7180(33)$. We can simply divide this average by our value for $f_+(0)\eta_{\rm EW}\sqrt{1+\delta_{\rm EM}}$ to isolate $|V_{cs}|$,

$$|V_{cs}|^{f_+(0)} = 0.9643(57)_{\rm latt}(44)_{\rm exp}(19)_{\eta_{EW}}(48)_{\rm EM}, \tag{5.29}$$

which is shown in Figure 5.15, alongside the equivalent $|V_{cs}|$ value for each of the experiments that goes into the average. In this case, as the averaging has already taken place, we simply allow a 1% uncertainty for $\delta_{\rm EM}$. This method agrees well with both of the others.

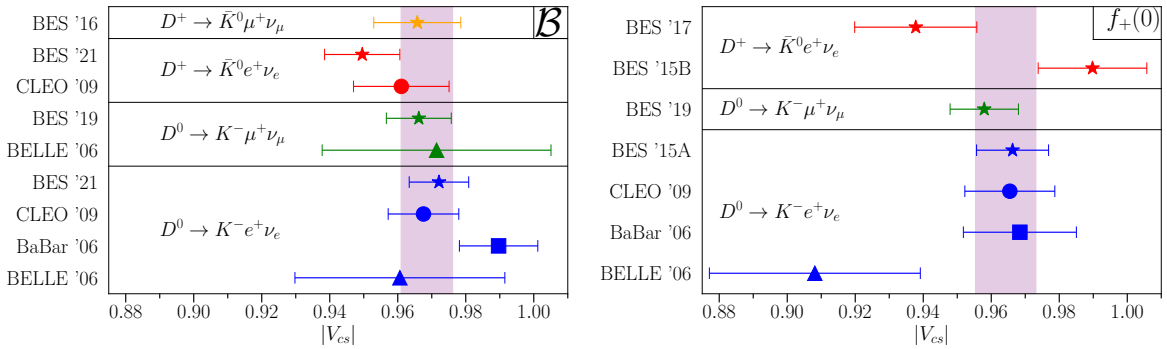

Figure 5.15: Left: comparison plot of the determination of $|V_{cs}|$ using the total branching fraction. Experimental results are from [160] for $D^+ \to \overline{K}^0 \mu^+ \nu_\mu$, [150, 161] for $D^+ \to \overline{K}^0 e^+ \nu_e$, from [156, 162] for $D^0 \to K^- \mu^+ \nu_\mu$ and from [150, 151, 161, 162] for $D^0 \to K^- e^+ \nu_e$ decays. Note that the BES results for final state e are the quoted averages for results from [161] and [152, 158]. The purple band gives the total average for V_{cs}, assuming 100% correlation of systematic uncertainties for results from a given experiment. The width of the purple band encompasses all uncertainties, including those from $\eta_{\rm EW}$ and $\delta_{\rm EM}$. Right: comparison plot of the determination of $|V_{cs}|$ using the extrapolation of experimental results to $q^2 = 0$. Experimental results are from [158, 163] for $D^+ \to \overline{K}^0 e^+ \nu_e$, from [156] for $D^0 \to K^- \mu^+ \nu_\mu$ and [150, 151, 152, 162] for $D^0 \to K^- e^+ \nu_e$. The purple band gives the weighted average result for V_{cs} obtained from the HFLAV weighted average [129] of the experimental results but including a correction for η_{EW} and an additional uncertainty from QED corrections.

5.5 Discussion and conclusions

Our three semi-independent $|V_{cs}|$ determinations (Equations (5.27), (5.28) and (5.29)) are shown in Figure 5.16, along side previous determinations. We see that all three methods agree very well, and provide a considerable improvement in uncertainty over previous work. Indeed, the uncertainty on our preferred result, $|V_{cs}|^{\mathcal{B}} = 0.9686(54)_{\rm latt}(39)_{\rm exp}(19)_{\eta_{EW}}(30)_{\rm EM}$, is below 1% for the first time, and is the first $|V_{cs}|$ determination to be significantly below unity. How's that for a slice of fried gold? Importantly, theoretical uncertainty has been brought down in line with experimental uncertainty, whilst $\delta_{\rm EM}$ has become a significant error in the final result, which needs further attention from the theory community. The unitarity constraints,

$$\begin{aligned} |V_{cd}|^2 + |V_{cs}|^2 + |V_{cb}|^2 &= 0.9826(22)_{V_{cd}}(155)_{V_{cs}}(1)_{V_{cb}}, \\ |V_{us}|^2 + |V_{cs}|^2 + |V_{ts}|^2 &= 0.9859(2)_{V_{us}}(155)_{V_{cs}}(1)_{V_{ts}}, \end{aligned} \tag{5.30}$$

are still dominated by $|V_{cs}|$, but with an improved uncertainty. They are both consistent with unity.

The left hand side of Figure 5.17 plots the $\pm 1\sigma$ band for our determination of V_{cs} from Equation (5.27) as the darker blue band. This is compared to the result (red band) from D_s leptonic decay of 0.983(18) from the 'Leptonic decays of charged pseudoscalar mesons' review in [29]. This result uses lattice QCD results for the D_s decay constant and includes uncertainties for η_{EW} and long-distance QED effects. The 'CKM Quark-Mixing Matrix' [29] review gives a value of 0.992(12) but without including these effects. This value would then lie in the upper half of the V_{cs} leptonic band plotted in Figure 5.17. In either case it is clear that our new result for V_{cs} is more accurate than that from leptonic decay and has a lower central value.

The plot also shows the constraints currently available on V_{cd}. The 'CKM Quark-Mixing Matrix' review in [29] quotes a value for V_{cd} from semileptonic $D \to \pi$ decay from combining experimental results with the form factor at $q^2 = 0$ determined in $N_f = 2+1+1$ lattice QCD by ETMC [139]. This gives $V_{cd} = 0.2330(136)$. The value quoted in the same review from D^+ leptonic decays is 0.2173(51). This combines experimental results with the D^+ decay constant determined in $N_f = 2+1+1$ lattice QCD by the Fermilab/MILC collaboration [79]. Another constraint follows from the ratio of D_s to D leptonic decay rates [118] combined with the ratio of D_s and D decay constants. Using ratios of $V_{cs} f_{D_s}$ and $V_{cd} f_{D^+}$ averaged over experimental results from [29] and the lattice QCD result for f_{D_s}/f_{D^+} from [79] gives the constraint $|V_{cd}|/|V_{cs}| = 0.2209(56)$ if we assume that electromagnetic corrections to the leptonic rates will largely cancel.

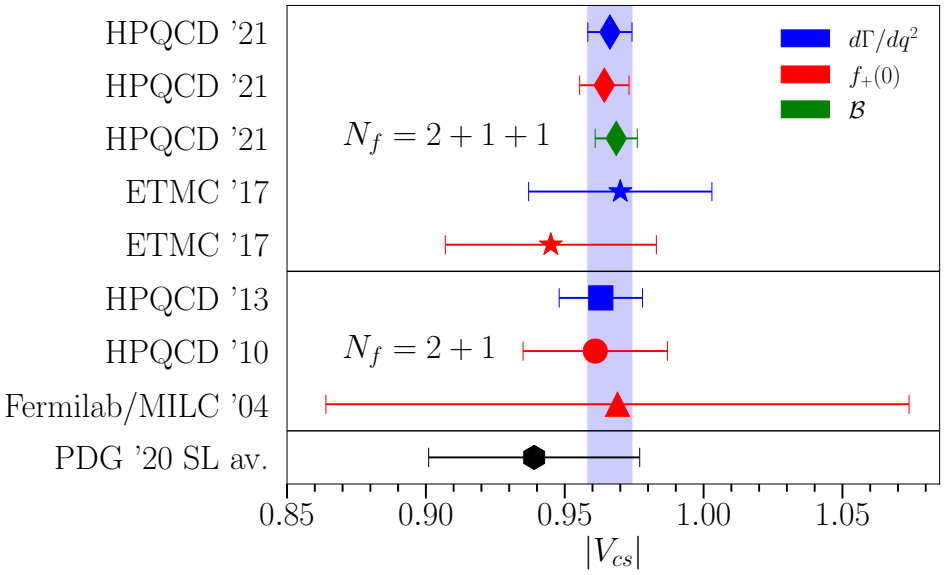

Figure 5.16: Our $|V_{cs}|$ result compared with other $N_f = 2+1+1$ and $N_f = 2+1$ results using lattice QCD. Different symbols indicate different lattice calculations, whilst different colours indicate the method used. Blue indicates use of the differential rate in q^2 bins, red indicates use of the $f_+(0)$ method and green indicates use of the total branching fraction for the decay. Points marked 'HPQCD '21' come from this work, 'ETMC '17' is from [139, 140], 'HPQCD '13' is from [138], 'HPQCD '10' is from [94] and 'Fermilab/MILC '04' is from [137]. For comparison we give at the bottom the value currently quoted in the Particle Data Tables [29] from semileptonic $D \to K$ decay. The blue band carries our preferred result, $V_{cs}^{\mathrm{d}\Gamma/\mathrm{d}q^2}$, down the plot.

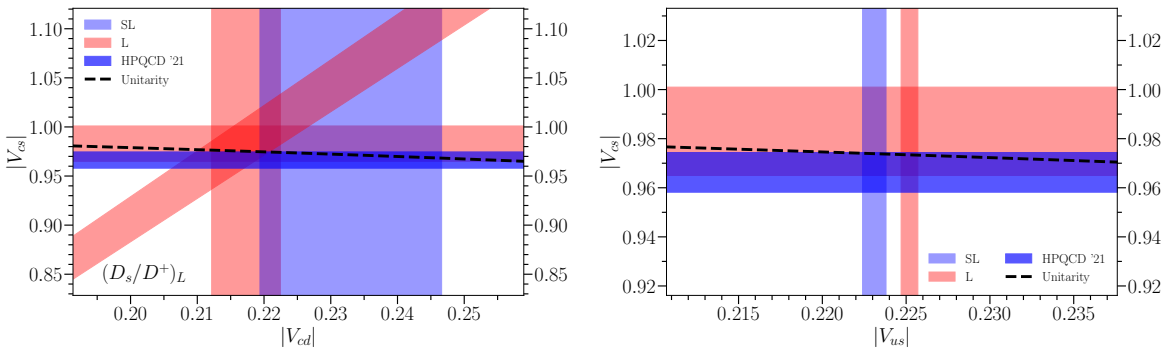

Figure 5.17: Left: a comparison of constraints on V_{cs} and V_{cd} with the expectation from CKM unitarity. Red bands show the $\pm 1\sigma$ range for the determination of V_{cs} and V_{cd} from leptonic decays of D_s and D^+ combined with decay constants from lattice QCD. The diagonal red band is the constraint from the ratio of leptonic rates for D_s and D^+ combined with the lattice QCD ratio of decay constants. The solid light blue band shows the result for V_{cd} from the $D \to \pi \ell \overline{\nu}$ decay combined with lattice QCD form factor results. See text for a discussion of the values used. The darker blue band shows our new determination here of V_{cs} from $D \to K \ell \overline{\nu}$ with $\pm 1\sigma$ uncertainties. For comparison the black dashed line gives the unitarity constraint curve of $|V_{cd}|^2 + |V_{cs}|^2 + |V_{cb}|^2 = 1$. Right: a similar comparison to above, this time of constraints on V_{cs} and V_{us}. Red bands show the $\pm 1\sigma$ range for the determination of V_{cs} and V_{us} from leptonic decays of D_s and K^+ combined with decay constants from lattice QCD. The light blue band shows the result for V_{us} from $K \to \pi \ell \overline{\nu}$ decay combined with lattice QCD form factor results. See text for a discussion of the values used. The darker blue band again shows our new determination of V_{cs} and the black dashed line gives the unitarity constraint curve, $|V_{us}|^2 + |V_{cs}|^2 + |V_{ts}|^2 = 1$.

The black dashed line in the left hand Figure 5.17 corresponds to the unitarity constraint $|V_{cd}|^2 + |V_{cs}|^2 + |V_{cb}|^2 = 1$. V_{cb} has little impact on this curve; we use the average value of 0.0410(14) from [29]. Our result for V_{cs} is in good agreement with the unitarity curve for values of V_{cd} in the range given by the leptonic and semileptonic constraints.

The right hand side of Figure 5.17 gives the same picture for the V_{us}, V_{cs}, V_{ts} column of the CKM matrix, showing constraints in the $|V_{cs}| - |V_{us}|$ plane. $|V_{cs}|$ values are as for the left hand figure but plotted over a smaller range because of the higher accuracy of $|V_{us}|$ (we scale x and y axis ranges together).

We take $|V_{us}|$ values from the review 'V_{ud}, V_{us}, Cabibbo angle and CKM unitarity' in [29]. This gives $|V_{us}| = 0.2252(5)$ from leptonic decays of K^+ and 0.2231(7) from $K \to \pi$ semileptonic decay. The leptonic result uses an average [164] of lattice QCD results for the K decay constants dominated by that from [79]. The semileptonic result uses an average [164] of lattice QCD results for the $K \to \pi$ form factor at $q^2 = 0$ from [165, 166]. The current most accurate lattice QCD results for the form factor are given in [167].

The right hand side of Figure 5.17 shows the tension developing between leptonic and semileptonic determinations of V_{us} [29, 167]. The black dashed line in the figure shows the unitarity constraint $|V_{us}|^2 + |V_{cs}|^2 + |V_{ts}|^2 = 1$. $|V_{ts}|$ has little impact on this curve; we use the current most accurate determination of $|V_{ts}| = 0.04189(93)$ from the measured oscillation rate of B_s mesons [29] and HPQCD's lattice QCD determination [90] of the matrix elements of the 4-quark operators that give the mass difference between the B_s eigenstates. Our improved accuracy for $|V_{cs}|$, along with the unitarity curve, is not sufficient to distinguish between the two values for $|V_{us}|$.

Chapter 6

$B_s \to \eta_s$ form factors

In this chapter, we will cover the work published in [2], which represents a progression of the heavy-HISQ method from $b \to c$ to $b \to s$ semileptonic decays, ultimately leading to the $B \to K$ calculation outlined in Chapter 7. Whilst the $B_s \to \eta_s$ calculation presented here was actually published before the $D \to K$ work discussed in Chapter 5, it embodies a transition from lattice calculations at a fixed, physical heavy mass (m_c), to the heavy-HISQ method, using multiple non-physical heavy masses m_h. For this reason it is natural to present these chapters in the order that we do, and not chronologically. We also note at this point that the $\eta_s = \bar{s}s$ is not a physical particle, but is straightforwardly constructed on the lattice. If this is causing an irresistible raising of your eyebrows, consider the η_s a kaon with an artificially heavy light quark - there's nothing fantastic here.

6.1 Introduction

The heavy-HISQ technique, which is described in detail in Section 4.5, is a method which allows all HISQ calculations of b meson form factors, even though the b mass is too high to be reached on almost all of our ensembles. To briefly recap Section 4.5, the method involves calculating form factors at a variety of heavy masses, m_h, typically such that $am_c \leq am_h \leq 0.8$ on each ensemble. A HQET inspired fit form (see Section 4.5.2) can then be folded into the z expansion, allowing an extrapolation in heavy mass to take place at the same time as the extrapolation to $a = 0$. The result is the form factors for b mesons without the matching errors associated with other methods such as NRQCD. There is a logarithmic matching between HQET and QCD, (see

Section 4.5.2) which we take account of in our fits. Whilst others [113, 114, 115, 116, 117, 118, 119, 120, 121, 122] have used a similar approach to calculate decay constants, we apply it to hadronic form factors.

At the time of the $B_s \to \eta_s$ calculation we will discuss here, heavy-HISQ had only recently been developed, and had only been used for $b \to c$ transitions, where it had been very successful [93, 147, 168, 169]. In order to proceed to $b \to s$ decays such as $B \to K$, it was important to check that the method performed equally well for lighter daughter quarks. Whilst the η_s is unphysical, $B_s \to \eta_s$ provides an intermediate step, with the $b \to s$ transition, but with a s spectator quark, which is much cheaper to compute than light quarks, and does not require costly ensembles with physical light sea quark masses in order to perform a good calculation. Additionally, strange quarks are less noisy than light quarks, and so a calculation with lower statistics can provide small enough uncertainties to check that the method has been successful.

As an additional bonus, once it was confirmed that heavy-HISQ was an appropriate method to study $b \to s$ decays, the $B_s \to \eta_s$ data calculated here was readily included in the $B \to K$ calculation detailed in Chapter 7, recycling the core hours expended, as well as improving the light quark mass extrapolation in $B \to K$.

6.2 Calculation details

As in the previous chapter, most of the general form of the calculation here is outlined in Chapter 4, so we will focus on details which are unique to $B_s \to \eta_s$

6.2.1 Simulation details

For this calculation, data from the three finest of the MILC $N_f = 2 + 1 + 1$ gluon ensembles discussed in Section 3.4 and listed in Table 3.1 were used, that is fine 5, superfine 5 and ultrafine 5, sets 6, 7 and 8. Note that in [2] these sets are labelled 1, 2 and 3 respectively. The twists applied on each ensemble, as well as the T values (i.e. the source sink separation in lattice units) are given in Table 6.1, along with the statistics used on each ensemble in terms of the number of configurations n_{cfg} and the number of time source (t_0) values n_{src} used on each configuration. Additionally, we now have four different heavy mass (am_h) values on each ensemble, in each case with the smallest being the tuned charm mass, and the largest being 0.8.

Table 6.1: Details of the statistics, T values, heavy masses, and η_s meson twists used on each ensemble. The number of configurations n_{cfg} and time sources n_{src} are given, and momenta can be obtained from twist, θ, via $\theta = |a\vec{p}_{\eta_s}|N_x/(\sqrt{3}\pi)$, where N_x is the spatial dimension of the lattice, given in Table 3.1. In each case, the smallest am_h value corresponds to the tuned am_c mass on that ensemble.

Set	$n_{\text{cfg}} \times n_{\text{src}}$	θ	am_h	T
6	504×16	0, 0.4281, 1.282, 2.141, 2.570	0.449, 0.566, 0.683, 0.8	14, 17, 20
7	454×8	0, 1.261, 2.108, 2.946, 3.624	0.274, 0.45, 0.6, 0.8	20, 25, 30
8	118×4	0, 0.706, 1.529, 2.235, 4.705	0.194, 0.45, 0.6, 0.8	33, 40

6.2.2 Correlation functions

As in the chapter above, the form factors calculated here are f_0 and f_+, as defined in Equations (4.56) and (4.55). As before, in the language of Chapter 4, we require a Goldstone η_s and H_s with spin-taste $\gamma_5 \otimes \xi_5$, and also a non-Goldstone $\hat{H}_s$, with spin-taste $\gamma_5\gamma_0 \otimes \xi_5\xi_0$. These are combined with currents with scalar $(1 \otimes 1)$ and temporal vector $(\gamma_0 \otimes \xi_0)$ spin-taste structure to generate the required two- and three-point functions. See Section 4.2.1 for details on how these operators are implemented in the MILC code.

6.2.3 Correlator fits

Two- and three-correlator fits are, as usual, performed using the recipe in Section 4.3, using the standard goodness of fit metrics of the χ^2 and log(GBF) value described there. A difference with the previous chapter, (and indeed the next chapter) is that here we make use of chained and marginalised fits, which are variations on the naive simultaneous fits, and are easily facilitated in our fitting packages [85, 86].

Chained and marginalised fits

Chaining and marginalisation are two methods to speed up large fits with many models. We will give an overview of both methods here, but they are detailed extensively in the documentation for corrfitter [86] and we direct the reader there for further details.

Chaining is perhaps the most obvious way to break up large fits. In our case, we have a large number of models which we wish to fit at once. We have four m_h values, each of which has a H_s and $\hat{H}_s$, giving a total of eight H_s type two-point correlators. We then have five different

daughter momenta, giving five η_s two point correlators. In terms of three-points, there is a scalar and vector current insertion for each mass-twist combination and T, giving $2 \times 4 \times 5 \times 3 = 120$ three-point functions in total[1]. Large numbers of models drive up the required SVD cut (see Section 4.3.4) and can be very slow to fit (in our case taking days on a standard machine). A solution is to break the fit down into smaller chunks, for example fitting the two points first, and then the scalar three-points, and then the vector three-points. At each stage in this chain, the fit parameters from the previous fit are included, even if not required, with their best fit values as priors. This allows their correlations to other fit parameters to be (approximately) accounted for, but speeds the fit up because the data is not included. In the limit of large statistics, the results from chained fits are identical to those from simultaneous fits.

Marginalisation is a different approach, aimed at reducing the number of fit parameters without changing the number of models. In marginalised fits, one effectively integrates out non-leading parameters. As a concrete example, let's imagine we want to fit a correlation function to $N = 6$ exponentials labelled $n = 0, 1, ..., 5$. We need to include all six exponentials in order to get a good fit, but higher order terms are poorly determined, and we are only actually interested in the ground state results ($n = 0$). Marginalising this case, we could take the priors for the $n = 3, 4, 5$ exponentials, and subtract the fit function evaluated using these priors from the data. We could then fit this modified data to a fit form with $N = 3$ exponentials, giving what is effectively a $N = 6$ exponential fit, but in the time (and with the resources) taken to do a $N = 3$ exponential fit. The obvious limitation to this method is that we are effectively forcing the marginalised fit parameters to take their prior values. This could lead to large uncertainties, however, with correlator fits, where the higher order exponentials die off rapidly and the priors can be reasonably well determined, this approach can be effective.

Chaining and marginalisation are readily combined, and this is the technique we employ here. On each ensemble, we have a two link chain, with the first link including models for all two-point fits, and the second for all three-points. The two-point fits are fitted to $N_{\text{exp}}^{\text{2pts}}$, which is 5 for set 6 and 6 for sets 7 and 8. These two-point fits are then chained into the three-point fits, which we apply marginalisation to. We fit to $N_{\text{exp}}^{\text{3pts}}$ exponentials, which is 2, 3 and 2 in the cases of sets 6, 7 and 8 [2] respectively, and then marginalise over the higher order exponentials up to $N_{\text{exp}}^{\text{2pts}}$ on each ensemble. This is effective because the two-point data are typically much more precise

[1] 80 for set 3 where there are only 2 T values.

[2] Set 8 contains only 2 T values, so the ratio of statistics in the two- to three-points is less than on the other ensembles, which explains why our goodness of fit tests prefer a fit to 2 exponentials, which is not in line with what might be expected based on the two-points.

Table 6.2: Priors used in the fit on each set. Priors are based on previous experience of similar fits [147] (see text) and given conservative widths. In some places, adjustment is made for lattice spacing, and priors are tuned using an increase in the GBF. The effect of doubling and halving the standard deviation on all priors on the final fit result is shown in Figure 6.1.

Set	$\mathcal{P}[A^{\rm n}_{i\neq 0}]$ and $\mathcal{P}[A^{\rm o}_i]$	$\mathcal{P}[S_{00}^{kl\neq nn}]$	$\mathcal{P}[V_{00}^{kl\neq nn}]$	$\mathcal{P}[S_{ij\neq 00}^{kl}]$	$\mathcal{P}[V_{ij\neq 00}^{kl}]$
6	0.10(10)	0.0(8)	0.2(1.0)	0.0(3)	0.0(3)
7	0.10(10)	0.0(8)	0.0(1.0)	0.0(3)	0.0(4)
8	0.05(05)	0.0(8)	0.0(1.0)	0.0(3)	0.0(4)

than three-points, and have fewer fit parameters, meaning that a larger number of exponentials is required for a good fit (for the same $t_{\rm min}$).

As in the previous chapter, the recommended SVD cut (see Section 4.3.4) is applied, and priors for ground state energies and amplitudes are estimated using $M_{\rm eff}$, $A_{\rm eff}$ or $J_{\rm eff}$ (Equations (4.44), (4.45) and (4.46)). All energy spacings between excited states are given a prior of $\Lambda_{\rm QCD} = 0.5{\rm GeV} \pm 50\%$, (this is common to all correlator fits in this thesis) and the oscillating ground states are taken to be 0.5 GeV above the non-oscillating ones. It is worth noting that unlike in $D \to K$ and $B \to K$, we do not make use of the dispersion relation to constrain the priors for η_s mesons (as in Equations (4.47) and (4.48)), though there is no reason we could not have done this. The remaining priors are listed in Table 6.2. As in Chapter 5, based on experience of similar fits [147], these are taken initially as being 0.1(1) for two-point amplitudes and 0(1) for three-points. They are then adjusted to the values in Table 6.2 using empirical Bayes.

Fit results and tests

We present in Figure 6.1 an example of fit stability and a test of the relativistic dispersion relation. The fit is tested against a variety of changes (see caption) and the stability of a particular parameter $V_{00}^{\rm nn}(am_h = 0.45, a|\vec{p}_{\eta_s}| = 0.143)$ (set 7) is plotted as an example. We also include the χ^2/d.o.f. for each fit, though it should be noted that this is artificially reduced by prior widths and SVD cuts (see Section 4.3.5) and so is only of use in a relative sense for fits with the same priors and data. The energies of the η_s mesons of different momenta are also plotted, and found to agree well with the relativistic dispersion relation, with no discernible discretisation effects at this level of uncertainty.

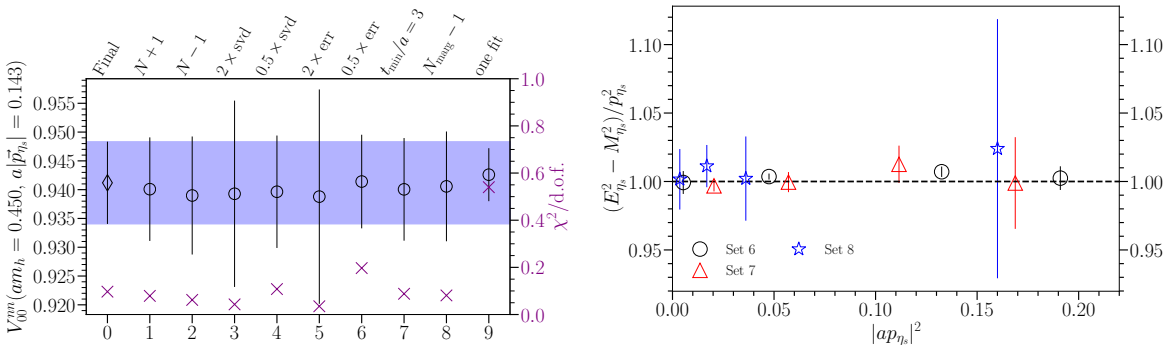

Figure 6.1: Left: stability tests of the chained, marginalised fit used on a typical three-point correlator. Test 0, the final result, shows the value of $V_{00}^{\rm nn}$ for $am_h = 0.45$, $a|\vec{p}| = 0.1430$ on set 7, with $N_{\rm exp}^{\rm 3pts} = 3$ exponential terms and three additional states marginalised (as discussed in the text), with $t_{\rm min} = 2$, the number of data points removed from the fit at the start and end of the data. Tests 1 and 2 show the effects of increasing and decreasing the number of fitted exponentials by 1, tests 3 and 4 show the effect of doubling and halving the SVD cut, and 5 and 6 show the effect of doubling and halving the error on all priors. Test 7 shows the effect of an increase on $t_{\rm min}$ by 1, and test 8 shows the reduction of the marginalised exponentials from 6 to 5. Finally, test 9 shows the result of just fitting the vector 3 point correlator for this mass and twist, and the relevant 2 points; this gives a reduced error since the smaller fit requires a smaller SVD cut. Fitting like this does not preserve correlations, however, so we use a global fit. Other two and three-point correlators behaved similarly well under the same tests. The χ^2/d.o.f. values (purple ×s) are also plotted for reference. Note that these are the raw values and hence artificially small (see text) and the degrees of freedom are not the same across all tests. Right: we plot the ratio $(E_{\eta_s}^2 - M_{\eta_s}^2)/|\vec{p}_{\eta_s}|^2$ from our fit results against $|a\vec{p}_{\eta_s}|^2$ to check that the η_s energy in our final fit results agrees with the momentum given to the meson in the lattice calculation.

Numerical details of all correlator fit results are given in the Appendix to [2].

6.2.4 z expansion: incorporating heavy mass dependence

The z expansion used this time will again be based on that described in Section 4.4.2, and we again make the choice $t_0 = 0$. We use the basic format of Equations (4.61) and (4.62), as with $D \to K$, but there are key differences from Equations (5.8), (5.9) and (5.11). The first is that we don't need a chiral logarithm term, as we have no valence light quark. All our valence quarks are well tuned, so any slight mistuning can be accounted for in the quark mistuning term $\mathcal{N}$. The second is that we now have to fold in an extrapolation in the heavy mass, which is inspired by HQET. Taking this into account we arrive at the prescription that follows:

$$f_0(q^2) = \frac{1}{1 - \frac{q^2}{M^2_{H^*_{s0}}}} \sum_{n=0}^{N-1} a_n^0 z(q^2)^n, \tag{6.1}$$

$$f_+(q^2) = \frac{1}{1 - \frac{q^2}{M^2_{H^*_s}}} \sum_{n=0}^{N-1} a_n^+ \Big(z(q^2)^n - \frac{n}{N}(-1)^{n-N} z(q^2)^N \Big). \tag{6.2}$$

Here we take $N = 3$ and we have modified the labelling of the heavy-strange scalar and vector meson pole masses from those appearing in [2] to be consistent with other notation in this thesis. The subtlety here is that we can't simply use a fixed pole mass, we need one that slides with our heavy mass. The mass $M_{H^*_{s0}}$ is taken as $M_{H_s} + 0.4$ GeV, which is consistent with lattice results in [170] and experimental results [29] for the axial vector–vector splitting, $M_{B_s}(1^+) - M_{B_s}(1^-)$.

The position of the $M_{H^*_s}$ pole can be estimated, as in [147], using the fact that $M_{H^*_s} - M_{H_s} \to 0$ as $m_h \to \infty$, with the ansatz $M_{H^*_s} = M_{H_s} + x/M_{H_s}$. We find x from the Particle Data Group (PDG [29]) value of $M^{\rm phys}_{B^*_s} - M^{\rm phys}_{B_s} = x/M^{\rm phys}_{B_s} = 0.0489(15)$ GeV. We go one step further to ensure that this ansatz also gives the correct PDG value for $M^{\rm phys}_{D^*_s} = 2.1122(4)$ GeV, using

$$\begin{aligned} M_{H^*_s} = M_{H_s} &+ \frac{M^{\rm phys}_{D_s}}{M_{H_s}} \Delta(D_s) \\ &+ \frac{M^{\rm phys}_{B_s}}{M_{H_s}} \Big[\frac{M_{H_s} - M^{\rm phys}_{D_s}}{M^{\rm phys}_{B_s} - M^{\rm phys}_{D_s}} \Big(\Delta(B_s) - \frac{M^{\rm phys}_{D_s}}{M^{\rm phys}_{B_s}} \Delta(D_s) \Big) \Big], \end{aligned} \tag{6.3}$$

with splittings $\Delta(B_s) = 0.0489(15)$ GeV and $\Delta(D_s) = 0.14386(41)$ GeV, from the PDG.

Another subtlety is the masses which appear in t_+, which goes into z (Equation 4.60). $t_+ = (M_1 + M_2)^2$, where M_1 and M_2 are the mesons with the lowest mass combination and the same quantum numbers as the current. That is to say, we have a $h \to s$ current, so we can combine this with light quarks to get the lowest mass combination. In the case of $D \to K$ and $B \to K$, the mesons here are exactly those used in the decay M_H (or M_D) and M_K, so we already have values for these masses. For $B_s \to \eta_s$, however, these masses are also M_H and M_K, which we do not determine in our calculation. Hence, we use $M_H = M_{H_s} + a(M_B^{\text{phys}} - M_{B_s}^{\text{phys}})$ and $M_K = M_{\eta_s} + a(M_K^{\text{phys}} - M_{\eta_s}^{\text{phys}})$ using the PDG and $M_{\eta_s}^{\text{phys}} = 0.6885(20)$ GeV [75].

In Equations (6.1) and (6.2), we have a more complicated expression for $a_n^{0,+}$ than in the case of $D \to K$,

$$a_n^{0,+} = (1 + \mathcal{N}_n^{0,+})\Big(1 + \rho_n^{0,+} \log\Big(\frac{M_{H_s}}{M_{D_s}}\Big)\Big) \sum_{j=i,j,k}^{N_{ijk}-1} d_{ijkn}^{0,+} \Big(\frac{a\Lambda_{\text{QCD}}}{M_{H_s}}\Big)^i \Big(\frac{am_h^{\text{val}}}{\pi}\Big)^{2j} \Big(\frac{a\Lambda_{\text{QCD}}}{\pi}\Big)^{2k}, \quad (6.4)$$

where we take $N_{ijk} = 3$ (after trialling different values using empirical Bayes) and the sum allows for different powers of discretisation effects in am_h^{val} as well as mass independent discretisation effects in $a\Lambda_{\text{QCD}}$. The fit to the heavy mass is taken care of via the (HQET inspired) polynomial in $\frac{a\Lambda_{\text{QCD}}}{M_{H_s}}$, with an additional logarithmic term $\rho_n^{0,+} \log\left(\frac{M_{H_s}}{M_{D_s}}\right)$, which allows for the matching of HQET to QCD [124, 125]. From [125], we expect the coefficient of the log term to be of order unity, so we use a prior of 0 ± 1. Initial guesses for priors for $d_{ijkn}^{0,+}$ coefficients are taken as 0(1), with the exception of $\mathcal{P}[d_{i10n}] = 0.0(5)$, which multiplies the second order mass depended discretisation effects, highly suppressed in the HISQ formalism. These are then adjusted using empirical Bayes (the standard procedure for priors in this thesis), which is described in more details below, with all priors being given in Tables 7.2-7.4. The constraint $f_+(0) = f_0(0)$ (see Section 4.4.2) is again trivial to implement by setting $d_{i000}^+ = d_{i000}^0$ and $\rho_0^+ = \rho_0^0$.

Mistuning of quark masses is accounted for using

$$\mathcal{N}_n^{0,+} = c_{s,n}^{\text{val},0,+}\delta_s^{\text{val}} + c_{s,n}^{\text{sea},0,+}\delta_s^{\text{sea}} + 2c_{l,n}^{\text{sea},0,+}\delta_l^{\text{sea}} + c_{c,n}^{0,+}\Big(\frac{M_{\eta_c} - M_{\eta_c}^{\text{phys}}}{M_{\eta_c}^{\text{phys}}}\Big), \quad (6.5)$$

where δ_q and m_s^{tuned} are the same as usual (Equations (5.13) and (5.14)). Further detail on prior choices is given in [2].

Additional data point

In our fit, we also include a continuum data point. This is not necessary, but it helps to reduce uncertainty considerably at $q^2_{\rm max}$, and testing our fit with and without the additional data confirms that our continuum extrapolation and heavy-HISQ method is working properly. We take the result $f_+(q^2_{\rm max}) = 0.811(17)$ from a previous HPQCD calculation [171]. Whilst this data point was obtained using NRQCD, and so would ordinarily contain matching errors which would not make it suitable to include in our calculation, it was constructed using a ratio, so as to eliminate these errors, and hence can be included in our fit without additional uncertainty. In Chapter 7, we shall see that we can match this data point almost exactly, by evaluating our $B \to K$ results at spectator mass $m_{\rm spectator} = m_s$, but let's not get ahead of ourselves.

z expansion priors and posteriors

Tables 7.2-7.4 provide prior and posterior values for our final z expansion fit. We use our standard method for determining such priors, introduced in Chapter 5 above: the z expansion is constructed such that we expect a_n values to be of order unity. Thus, we take most priors as 0(1). In some cases we tighten up priors, for example when they relate to terms containing discretisation effect which go like a^2 (in this case $d^{0,+}_{i,1,0,n}$). We know such effect are highly suppressed in the HISQ action, so we give them narrower priors. Similarly, see quark mistuning effects are expect to be small, so their priors are narrowed. After such adjustments, we apply the empirical Bayes method discussed in Section 4.3.5 to broad groups of priors, for example $c^{\rm sea}$. Generally, we do not narrow priors unless the width favoured by log(GBF), is less than half of what we are using, which in this case leaves all priors as they were. The effect of broadening and narrowing priors on the final form factors will be discussed in the next section.

All but one of our priors and posteriors (d^0_{0001} is slightly over 1σ from its prior) agree to within 1 standard deviation, indicating that our priors are indeed conservative. We do not include other parameters which do technically appear as priors in the fit, but have very well informed values. For example the physical meson masses, such as $M^{\rm phys}_{B_s}$. Because such priors are known to a very high degree of accuracy, the fit tends not to move them at - all such posteriors agree with their priors to within 1σ, with most not moved off their priors.

Our preferred fit has a χ^2/d.o.f. of 0.14, for 109 degrees of freedom, corresponding to a Q value of 1.00. log(GBF) = 298.22. The continuum form factors resulting from this fit, as well as tests of the fit stability against reasonable variations in the fit form and parameters will be given

Table 6.3: Priors and their posterior values for the z expansion fit.

Parameter	Prior	Posterior
ρ_0^0	0.0(1.0)	-0.582(48)
ρ_1^0	0.0(1.0)	-0.63(49)
ρ_2^0	0.0(1.0)	-0.20(82)
ρ_0^+	0.0(1.0)	-0.582(48)
ρ_1^+	0.0(1.0)	0.13(33)
ρ_2^+	0.0(1.0)	-0.03(98)
d^0_{0000}	0.0(1.0)	0.690(70)
d^0_{0001}	0.0(1.0)	0.36(21)
d^0_{0002}	0.0(1.0)	0.35(57)
d^0_{0010}	0.0(1.0)	-0.60(72)
d^0_{0011}	0.0(1.0)	0.07(1.00)
d^0_{0012}	0.0(1.0)	-0.005(1.000)
d^0_{0020}	0.0(1.0)	-0.004(1.000)
d^0_{0021}	0.0(1.0)	0.0005(1.0000)
d^0_{0022}	0.0(1.0)	-2e-05 +- 1
d^0_{0100}	0.0(1.0)	0.08(19)
d^0_{0101}	0.0(1.0)	0.20(80)
d^0_{0102}	0.0(1.0)	-0.16(96)
d^0_{0110}	0.0(1.0)	0.02(1.00)
d^0_{0111}	0.0(1.0)	-0.0005(1.0000)
d^0_{0112}	0.0(1.0)	-0.0001(1.0000)
d^0_{0120}	0.0(1.0)	0.0001(1.0000)
d^0_{0121}	0.0(1.0)	-6e-06 +- 1
d^0_{0122}	0.0(1.0)	5e-07 +- 1
d^0_{0200}	0.0(1.0)	0.51(93)
d^0_{0201}	0.0(1.0)	-0.007(0.997)
d^0_{0202}	0.0(1.0)	-0.01(1.00)
d^0_{0210}	0.0(1.0)	0.003(1.000)
d^0_{0211}	0.0(1.0)	-0.0002(1.0000)
d^0_{0212}	0.0(1.0)	-9e-06 +- 1
d^0_{0220}	0.0(1.0)	2e-05 +- 1
d^0_{0221}	0.0(1.0)	-1e-06 +- 1
d^0_{0222}	0.0(1.0)	5e-08 +- 1
d^0_{1000}	0.0(1.0)	0.25(45)
d^0_{1001}	0.0(1.0)	0.46(77)
d^0_{1002}	0.0(1.0)	0.15(98)
d^0_{1010}	0.0(1.0)	-0.17(98)
d^0_{1011}	0.0(1.0)	0.02(1.00)
d^0_{1012}	0.0(1.0)	-0.001(1.000)
d^0_{1020}	0.0(1.0)	-0.001(1.000)
d^0_{1021}	0.0(1.0)	0.0001(1.0000)
d^0_{1022}	0.0(1.0)	-9e-06 +- 1
d^0_{1100}	0.00(50)	-0.07(49)
d^0_{1101}	0.00(50)	0.01(50)
d^0_{1102}	0.00(50)	-0.004(500)
d^0_{1110}	0.0(1.0)	0.0003(1.0000)
d^0_{1111}	0.0(1.0)	6e-05 +- 1
d^0_{1112}	0.0(1.0)	-1e-05 +- 1
d^0_{1120}	0.0(1.0)	6e-06 +- 1
d^0_{1121}	0.0(1.0)	1e-09 +- 1
d^0_{1122}	0.0(1.0)	8e-08 +- 1
d^0_{1200}	0.0(1.0)	0.08(1.00)
d^0_{1201}	0.0(1.0)	-0.002(1.000)
d^0_{1202}	0.0(1.0)	-0.001(1.000)
d^0_{1210}	0.0(1.0)	0.0005(1.0000)
d^0_{1211}	0.0(1.0)	-3e-05 +- 1
d^0_{1212}	0.0(1.0)	1e-07 +- 1
d^0_{1220}	0.0(1.0)	3e-06 +- 1
d^0_{1221}	0.0(1.0)	-2e-07 +- 1
d^0_{1222}	0.0(1.0)	1e-08 +- 1
d^0_{2000}	0.0(1.0)	-0.26(77)
d^0_{2001}	0.0(1.0)	0.25(95)
d^0_{2002}	0.0(1.0)	0.03(1.00)
d^0_{2010}	0.0(1.0)	-0.04(1.00)
d^0_{2011}	0.0(1.0)	0.007(1.000)
d^0_{2012}	0.0(1.0)	-0.0004(1.0000)
d^0_{2020}	0.0(1.0)	-0.0004(1.0000)
d^0_{2021}	0.0(1.0)	4e-05 +- 1
d^0_{2022}	0.0(1.0)	-3e-06 +- 1
d^0_{2100}	0.00(50)	-0.03(50)
d^0_{2101}	0.00(50)	0.003(500)
d^0_{2102}	0.00(50)	-0.0004(5000)
d^0_{2110}	0.0(1.0)	-0.0005(1.0000)
d^0_{2111}	0.0(1.0)	6e-05 +- 1
d^0_{2112}	0.0(1.0)	-2e-06 +- 1
d^0_{2120}	0.0(1.0)	-3e-06 +- 1
d^0_{2121}	0.0(1.0)	3e-07 +- 1
d^0_{2122}	0.0(1.0)	-1e-09 +- 1
d^0_{2200}	0.0(1.0)	0.01(1.00)
d^0_{2201}	0.0(1.0)	-0.0005(1.0000)
d^0_{2202}	0.0(1.0)	-9e-05 +- 1
d^0_{2210}	0.0(1.0)	8e-05 +- 1
d^0_{2211}	0.0(1.0)	-4e-06 +- 1
d^0_{2212}	0.0(1.0)	2e-07 +- 1
d^0_{2220}	0.0(1.0)	5e-07 +- 1
d^0_{2221}	0.0(1.0)	-3e-08 +- 1
d^0_{2222}	0.0(1.0)	2e-09 +- 1
$c^{\mathrm{sea},0}_{l,0}$	0.00(30)	0.001(212)
$c^{\mathrm{sea},0}_{l,1}$	0.00(30)	0.004(300)
$c^{\mathrm{sea},0}_{l,2}$	0.00(30)	-0.001(300)

Table 6.4: Priors and their posterior values for the z expansion fit.

Parameter	Prior	Posterior
$c_{s,0}^{\mathrm{sea},0}$	0.00(30)	0.03(26)
$c_{s,1}^{\mathrm{sea},0}$	0.00(30)	-0.002(300)
$c_{s,2}^{\mathrm{sea},0}$	0.00(30)	0.0002(3000)
$c_{c,0}^{0}$	0.00(30)	-0.008(299)
$c_{c,1}^{0}$	0.00(30)	0.001(300)
$c_{c,2}^{0}$	0.00(30)	-0.0001(3000)
$c_{s,0}^{\mathrm{val},0}$	0.0(1.0)	0.11(99)
$c_{s,1}^{\mathrm{val},0}$	0.0(1.0)	-0.008(1.000)
$c_{s,2}^{\mathrm{val},0}$	0.0(1.0)	0.0002(1.0000)
d_{0000}^{+}	0.0(1.0)	0.690(70)
d_{0001}^{+}	0.0(1.0)	-1.07(26)
d_{0002}^{+}	0.0(1.0)	0.87(83)
d_{0010}^{+}	0.0(1.0)	-0.35(75)
d_{0011}^{+}	0.0(1.0)	-0.01(1.00)
d_{0012}^{+}	0.0(1.0)	0.005(1.000)
d_{0020}^{+}	0.0(1.0)	-0.003(1.000)
d_{0021}^{+}	0.0(1.0)	-6e-05 +- 1
d_{0022}^{+}	0.0(1.0)	3e-05 +- 1
d_{0100}^{+}	0.0(1.0)	0.009(196)
d_{0101}^{+}	0.0(1.0)	-0.08(97)
d_{0102}^{+}	0.0(1.0)	0.03(1.00)
d_{0110}^{+}	0.0(1.0)	-0.003(1.000)
d_{0111}^{+}	0.0(1.0)	-0.001(1.000)
d_{0112}^{+}	0.0(1.0)	0.0002(1.0000)
d_{0120}^{+}	0.0(1.0)	-2e-05 +- 1
d_{0121}^{+}	0.0(1.0)	-7e-06 +- 1
d_{0122}^{+}	0.0(1.0)	1e-06 +- 1
d_{0200}^{+}	0.0(1.0)	-0.04(98)
d_{0201}^{+}	0.0(1.0)	-0.002(1.000)
d_{0202}^{+}	0.0(1.0)	0.0008(1.0000)
d_{0210}^{+}	0.0(1.0)	-9e-05 +- 1
d_{0211}^{+}	0.0(1.0)	-6e-05 +- 1
d_{0212}^{+}	0.0(1.0)	9e-06 +- 1
d_{0220}^{+}	0.0(1.0)	-3e-07 +- 1
d_{0221}^{+}	0.0(1.0)	-4e-07 +- 1
d_{0222}^{+}	0.0(1.0)	6e-08 +- 1
d_{1000}^{+}	0.0(1.0)	0.25(45)
d_{1001}^{+}	0.0(1.0)	-0.15(95)
d_{1002}^{+}	0.0(1.0)	0.22(99)
d_{1010}^{+}	0.0(1.0)	-0.12(98)
d_{1011}^{+}	0.0(1.0)	-0.0007(1.0000)
d_{1012}^{+}	0.0(1.0)	0.001(1.000)
d_{1020}^{+}	0.0(1.0)	-0.0008(1.0000)
d_{1021}^{+}	0.0(1.0)	-4e-06 +- 1
d_{1022}^{+}	0.0(1.0)	6e-06 +- 1

Parameter	Prior	Posterior
d_{1100}^{+}	0.00(50)	-0.02(49)
d_{1101}^{+}	0.00(50)	-0.006(500)
d_{1102}^{+}	0.00(50)	0.002(500)
d_{1110}^{+}	0.0(1.0)	-0.001(1.000)
d_{1111}^{+}	0.0(1.0)	-0.0002(1.0000)
d_{1112}^{+}	0.0(1.0)	5e-05 +- 1
d_{1120}^{+}	0.0(1.0)	-9e-06 +- 1
d_{1121}^{+}	0.0(1.0)	-1e-06 +- 1
d_{1122}^{+}	0.0(1.0)	3e-07 +- 1
d_{1200}^{+}	0.0(1.0)	-0.003(0.999)
d_{1201}^{+}	0.0(1.0)	-0.001(1.000)
d_{1202}^{+}	0.0(1.0)	0.0003(1.0000)
d_{1210}^{+}	0.0(1.0)	-1e-05 +- 1
d_{1211}^{+}	0.0(1.0)	-1e-05 +- 1
d_{1212}^{+}	0.0(1.0)	2e-06 +- 1
d_{1220}^{+}	0.0(1.0)	-7e-08 +- 1
d_{1221}^{+}	0.0(1.0)	-7e-08 +- 1
d_{1222}^{+}	0.0(1.0)	1e-08 +- 1
d_{2000}^{+}	0.0(1.0)	-0.26(77)
d_{2001}^{+}	0.0(1.0)	0.02(99)
d_{2002}^{+}	0.0(1.0)	0.05(1.00)
d_{2010}^{+}	0.0(1.0)	-0.04(1.00)
d_{2011}^{+}	0.0(1.0)	0.0003(1.0000)
d_{2012}^{+}	0.0(1.0)	0.0003(1.0000)
d_{2020}^{+}	0.0(1.0)	-0.0002(1.0000)
d_{2021}^{+}	0.0(1.0)	2e-06 +- 1
d_{2022}^{+}	0.0(1.0)	1e-06 +- 1
d_{2100}^{+}	0.00(50)	-0.01(50)
d_{2101}^{+}	0.00(50)	-0.001(500)
d_{2102}^{+}	0.00(50)	0.0004(5000)
d_{2110}^{+}	0.0(1.0)	-0.0005(1.0000)
d_{2111}^{+}	0.0(1.0)	-3e-05 +- 1
d_{2112}^{+}	0.0(1.0)	1e-05 +- 1
d_{2120}^{+}	0.0(1.0)	-3e-06 +- 1
d_{2121}^{+}	0.0(1.0)	-2e-07 +- 1
d_{2122}^{+}	0.0(1.0)	5e-08 +- 1
d_{2200}^{+}	0.0(1.0)	-0.0004(1.0000)
d_{2201}^{+}	0.0(1.0)	-0.0004(1.0000)
d_{2202}^{+}	0.0(1.0)	7e-05 +- 1
d_{2210}^{+}	0.0(1.0)	-3e-06 +- 1
d_{2211}^{+}	0.0(1.0)	-2e-06 +- 1
d_{2212}^{+}	0.0(1.0)	4e-07 +- 1
d_{2220}^{+}	0.0(1.0)	-2e-08 +- 1
d_{2221}^{+}	0.0(1.0)	-1e-08 +- 1
d_{2222}^{+}	0.0(1.0)	2e-09 +- 1

Table 6.5: Priors and their posterior values for the z expansion fit.

$c_{l,0}^{\text{sea},+}$	0.00(30)	0.002(214)
$c_{l,1}^{\text{sea},+}$	0.00(30)	0.003(300)
$c_{l,2}^{\text{sea},+}$	0.00(30)	0.002(300)
$c_{s,0}^{\text{sea},+}$	0.00(30)	0.03(26)
$c_{s,1}^{\text{sea},+}$	0.00(30)	-0.00009(29998)
$c_{s,2}^{\text{sea},+}$	0.00(30)	-0.0002(3000)
$c_{c,0}^{+}$	0.00(30)	-0.004(299)
$c_{c,1}^{+}$	0.00(30)	0.0006(3000)
$c_{c,2}^{+}$	0.00(30)	0.0003(3000)
$c_{s,0}^{\text{val},+}$	0.0(1.0)	0.07(99)
$c_{s,1}^{\text{val},+}$	0.0(1.0)	-0.002(1.000)
$c_{s,2}^{\text{val},+}$	0.0(1.0)	-0.0006(1.0000)

below.

6.3 Results and discussion

6.3.1 Form factor results

The heavy-HISQ method is clearly demonstrated in Figure 6.2, which show the continuum, physical point f_0 and f_+ form factors in z space, with their poles removed, alongside the data.

Both plots show the data on each ensemble, with lines joining points of fixed m_h, cascading down the page as the heavy mass increases, towards the continuum result (solid band). In both cases the closest data point to the continuum is that on the finest ensemble (set 8) at the highest mass, which is close to the physical b mass. We can see that the z expansions are nearly linear in both cases, justifying our choice of $N = 3$ as being somewhat conservative. We can also see that the data points at the lowest mass on each ensemble (shown in dark blue), which is the charm mass in each case, are very consistent with each other. This suggest that the form factors[3] do not have strong lattice spacing dependence, equivalent to saying that discretisation effects are small in the HISQ action, as is expected. Additionally, the continuum data point is plotted in the case of f_0, providing a visual link between the data points at finite lattice spacing and unphysical heavy mass, and the continuum physical point.

The top of Figure 6.3 shows our $B_s \to \eta_s$ form factors, evaluated in continuum and at physical masses. They compare well with the results in [104], though f_+ is significantly less precise at large q^2, something which could straightforwardly be improved with more statistics, as will be demonstrated below. The agreement with [104], which is a calculation using an NRQCD b quark at the physical mass, is an excellent affirmation of the suitability of heavy-HISQ for carrying out $b \to s$ calculations such as this one.

At the bottom, Figure 6.3 presents tests of the stability of the z expansion against a variety of changes (see caption for detail). Most notably, we see that the removal of the continuum data (Test 4) only affects $f_0(q^2_{\text{max}})$, and whilst the error at this point roughly doubles without the external input, the central value remains within 1σ of the final result. We find that f_0 and f_+ are stable against all of the changes.

A breakdown in the error on the form factors across the q^2 range is given in Figure 6.4. We can see that statistics is the dominant source of uncertainty in both cases, followed by the extrapolation in heavy quark mass used to reach M_{B_s}. Improving statistics is straightforward (though costly),

[3] At least at the $D_s \to \eta_s$ end.

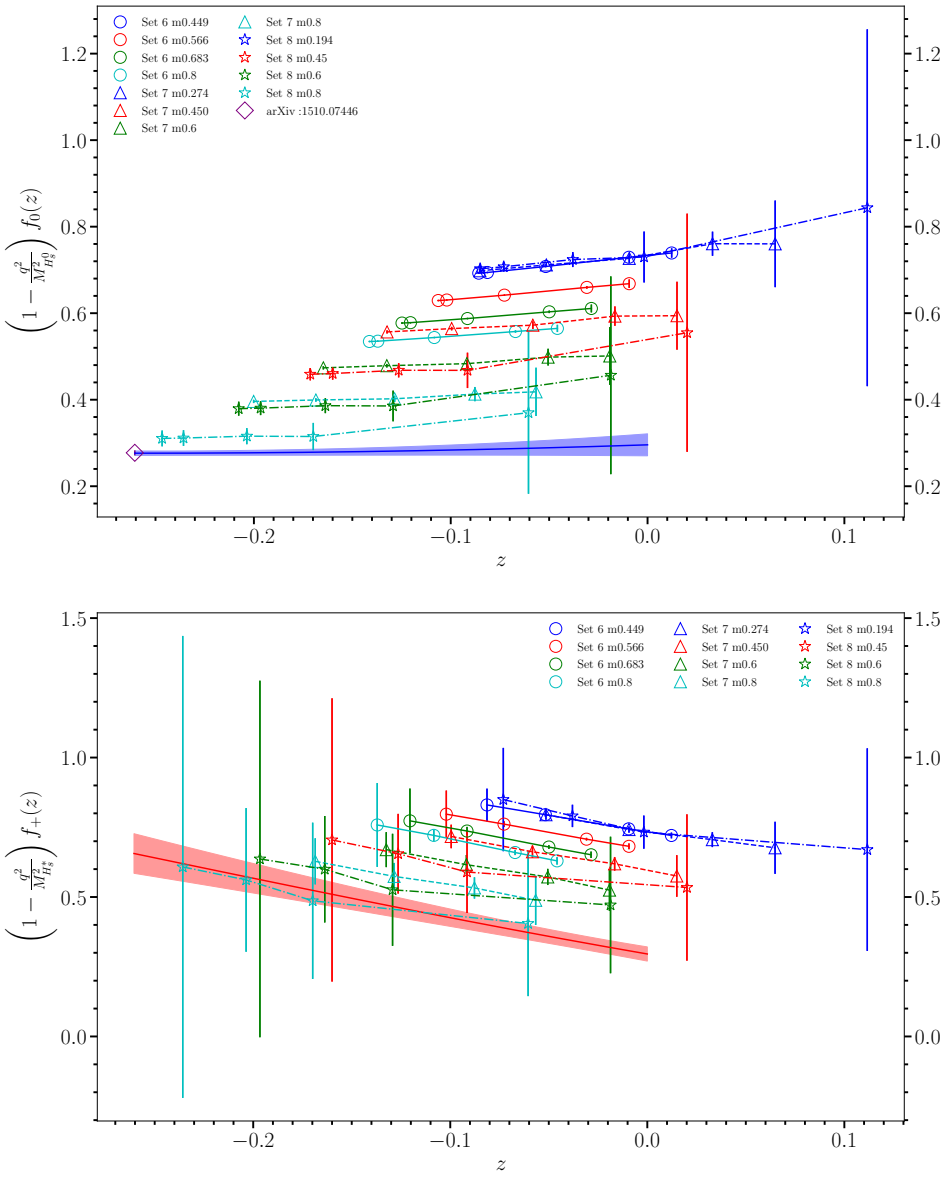

Figure 6.2: $\left(1-\frac{q^2}{M^2_{H^*_{s0}}}\right) f_0(z)$ (top) and $\left(1-\frac{q^2}{M^2_{H^*_s}}\right) f_+(z)$ (bottom) data points and final result at the physical point (coloured bands). Data points are labelled by mass for sets 6, 7 and 8 respectively, where e.g. m0.8 indicates $am_h = 0.8$ on that ensemble. Lines between data points of a given heavy mass over the full z range are there to guide the eye. For f_0, the additional continuum data point from [171] is shown in purple, and helps to pin down the form factor in the high q^2 limit.

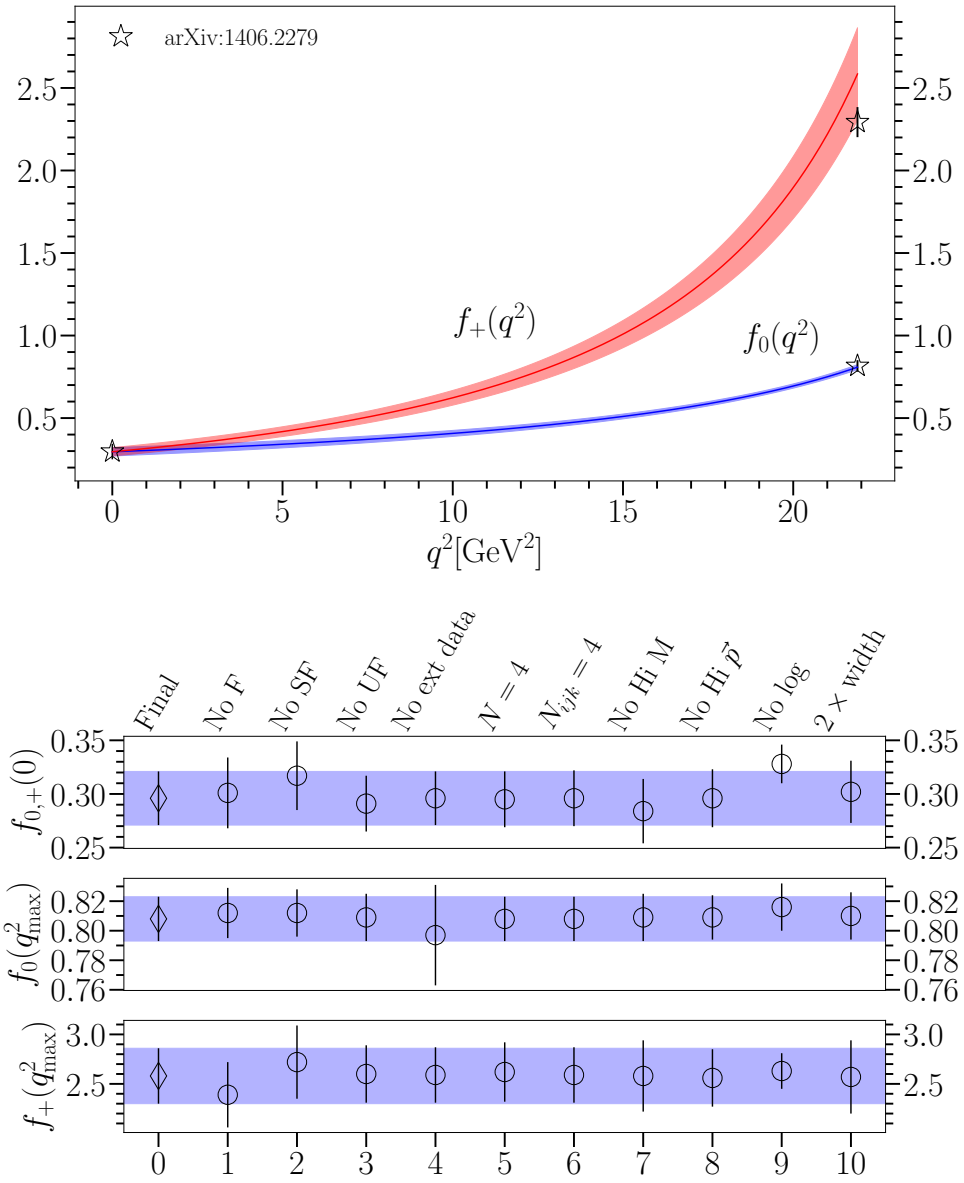

Figure 6.3: Top: final form factor results for $f_0(q^2)$ and $f_+(q^2)$. Results from [104] at $q^2 = 0$ and $q^2 = q^2_{\rm max}$ are also shown. Bottom: stability tests of the fit of the form factors $f_{0,+}(0)$, $f_0(q^2_{\rm max})$ and $f_+(q^2_{\rm max})$. Test 0 is the final result, shown throughout by the blue band. Tests 1, 2 and 3 are the results if the fine, superfine and ultrafine data are removed respectively. Test 4 is the fit without the data point from [171]. Test 5 adds a cubic term in the z expansion (Equations (6.1) and (6.2)). Test 6 shows the effect of extending the i, j, k sum in Equation (6.4) Tests 7 and 8 remove the highest masses and momenta for all lattice spacings respectively. Test 9 is without the log term in Equation (6.4), here we find that d_{i000} terms change to mimic the Taylor expansion of the log, and we require much larger priors (0.0(5.0)) to account for this. Test 10 shows the effect of doubling the width of all d_{ijkn} priors. We see that our extrapolation is stable to all of the above modifications. GBF values again tell us that our priors are conservative.

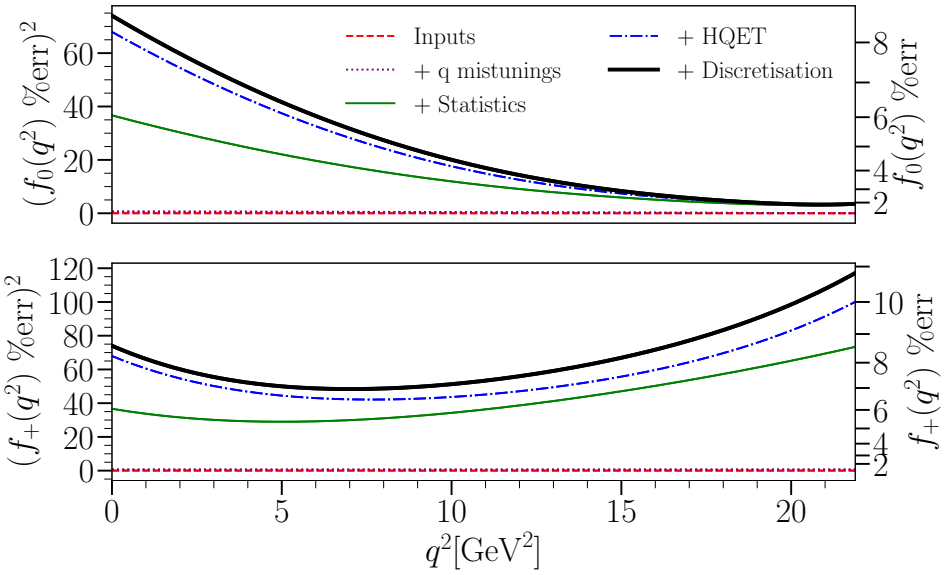

Figure 6.4: The contributions to the total percentage error (black line) of $f_0(q^2)$ (top) and $f_+(q^2)$ (bottom) from different sources, shown as an accumulating error. The red dashed line ('inputs') includes values for masses taken from the PDG [29] and used in the fit as described above. The purple dotted line ('q mistunings') adds, negligibly, to the inputs the error contribution from the quark mistunings associated with c fit parameters, whilst the solid green line ('statistics') further adds the error from our correlator fits. The blue dot-dash line ('HQET') includes the contribution from the expansion in the heavy quark mass, and, finally, the thick black line ('Discretisation'), the total error on the form factor, also includes the discretisation errors. The percentage variance adds linearly and the scale for this is given on the left hand axis. The percentage standard deviation, the square root of this, can be read from the scale on the right-hand side.

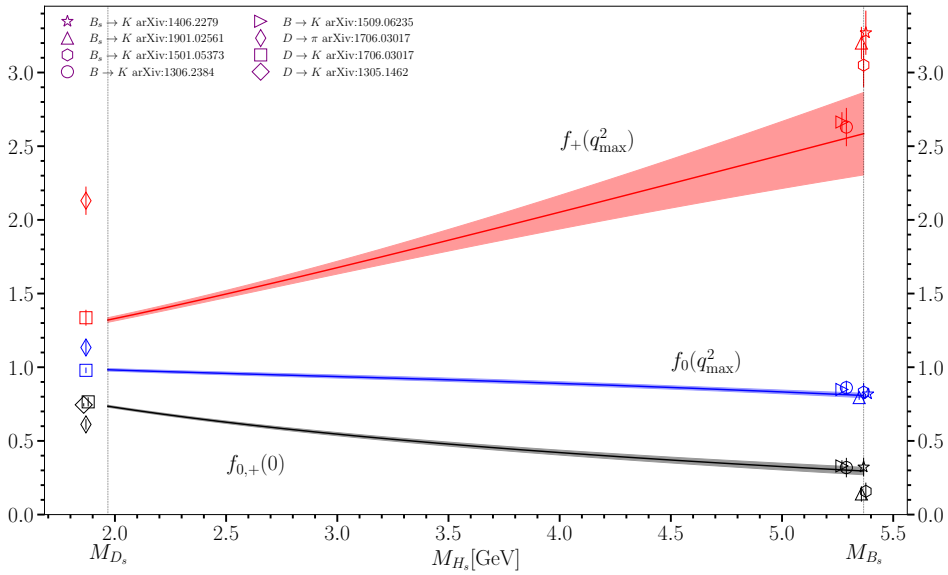

Figure 6.5: The form factors $f_{0,+}(0)$, $f_0(q^2_{\rm max})$ and $f_+(q^2_{\rm max})$ over the range of heavy masses from the physical D_s to the physical B_s. Results are included for $f_{0,+}(0)$, $f_0(q^2_{\rm max})$ and $f_+(q^2_{\rm max})$ (in their respective colours) for several other decays related by SU(3) flavour symmetry [104, 138, 139, 172, 173]. Data points are plotted at the x axis values corresponding to their physical heavy meson mass, not the mass that would result from their heavy quark and a strange quark (which would put them all at M_{D_s} or M_{B_s}). In the case of M_B and M_D some of the points are offset slightly either side of the mass for clarity.

and a reduction in the uncertainty from the heavy mass would be a similar exercise, running data for extra heavy mass values, and perhaps pushing the highest value on the finest lattice closer to the b mass. As the form factors are unphysical, and acting as a test bed for heavy-HISQ, we did not invest more computer time in reducing these uncertainties - the fact that the uncertainty is dominated by sources which are easily addressed is what we wanted to take from this.

Whilst the dependence on heavy mass is apparent in Figure 6.2, the heavy-HISQ method allows us to be even more explicit. Instead of fixing the M_{H_s} mass (at M_{B_s}) and plotting the form factors in q^2, we can fix q^2 and plot the form factors against varying M_{H_s}. Given the range of our $m_c \leq m_h$ values used to extract the data, we can reasonably do this over the range $M_{D_s} \leq M_{H_s} \leq M_{B_s}$, giving the form factors across the range $D_s \to \eta_s$ to $B_s \to \eta_s$. This is what is shown in Figure 6.5, taking extremal q^2 values of the form factors. These results are compared with other decays which are related by $SU(3)$ flavour symmetry at both end of the mass range:

Table 6.6: Values of fit coefficients $a_n^{0,+}$ and pole masses at the physical point for the $B_s \to \eta_s$ decay with correlation matrix are given below. Form factors can be reconstructed by evaluating Equations (6.1) and (6.2) using these coefficients and pole masses. Note that $M_{B_{s0}}$ is set to $M_{B_s} + 0.4\,\text{GeV}$. Masses are in GeV. The pole masses are very slightly correlated due to the way the fit function is constructed. These correlations are too small to have any meaningful effect on the fit, but we include them for completeness in reconstructing our results.

a_0^0	a_1^0	a_2^0	a_0^+	a_1^+	a_2^+	$M_{B_{s0}^*}$	$M_{B_s^*}$
0.296(25)	0.15(20)	0.29(47)	0.296(25)	−1.22(32)	0.9(1.2)	5.76688(17)	5.4158(15)
1.00000	0.90818	0.72266	1.00000	0.30483	0.09764	−0.00042	0.00021
	1.00000	0.93763	0.90818	0.38642	0.09064	0.00002	−0.00009
		1.00000	0.72266	0.40724	0.07271	0.00012	−0.00036
			1.00000	0.30483	0.09764	−0.00042	0.00021
				1.00000	0.51317	0.00179	−0.01229
					1.00000	−0.00045	0.00248
						1.00000	0.00000
							1.00000

$B_s \to K$, $B \to K$, $D \to K$ and $D \to \pi$. (Spoiler alert: a more robust comparison, across the full q^2 range, will be carried out in Chapter 7). We note that changing the spectator quark from strange to light ($B \to K$, $D \to K$) has almost no effect on form factors, whilst the same change in the daughter quark ($B_s \to K$) has a modest effect, pushing up $f_+(q^2_{\text{max}})$ and down $f_{0/+}(0)$, but not significantly. Changing both the spectator and daughter quarks to light ($D \to K$) makes a much larger difference, showing a clear breaking of $SU(3)$ flavour symmetry. $B \to \pi$ results are in even worse agreement, and are not included.

Finally, Table 6.6 provides the continuum z expansions coefficients and pole Masses required to reconstruct our form factors, together with their correlation matrix.

6.3.2 Tests of HQET

Using our form factors, we can perform a number of tests of Heavy Quark Effective Theory (HQET), which is outlined in Section 4.5.2. The fact that the heavy mass is a knob we can twiddle in our results makes them particularly suited to this endeavour. A caveat to all of this is that our fit form was of course motivated by HQET, so our form factors implicitly contain polynomial M_H^{-1} dependence. This does not prevent us from exploring some HQET predictions,

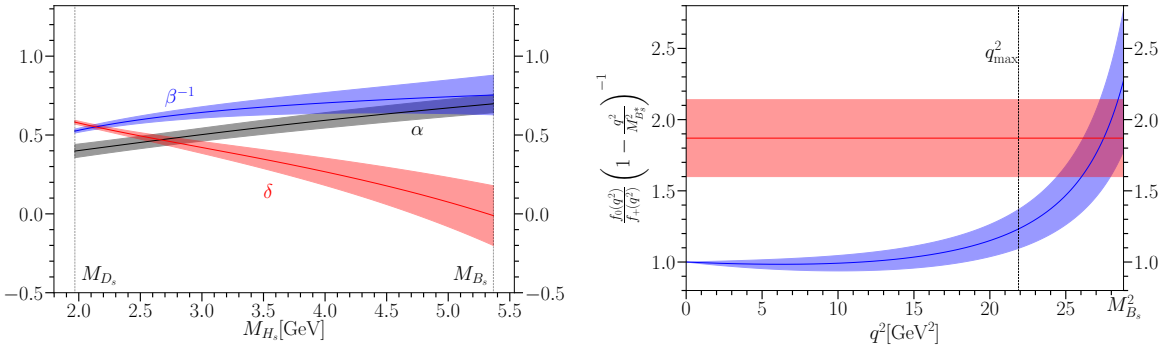

Figure 6.6: Left: the quantities α, β^{-1} and δ, defined in Equations (6.6), (6.7) and (6.8), over the range of heavy masses from the physical D_s to the physical B_s. Right: the form factor ratio, $\frac{f_0(q^2)}{f_+(q^2)}\left(1-\frac{q^2}{M^2_{B^*_s}}\right)^{-1}$ over the range $0 \le q^2 \le M^2_{B_s}$ (blue band), as compared with the HQET expectation in the limit $q^2 \to M^2_{B_s}$ (red band), defined in Equation (6.9).

however.

The first thing we can address is the form factor shape, characterised by the quantities α, δ and β^{-1} [174, 175], which are related to the slope of the form factors at $q^2 = 0$ and the value at high q^2,

$$\frac{1}{1-\alpha} = \frac{1}{M^2_{H^*_s}} \text{Res}_{q^2=M^2_{H^*_s}} \frac{f_+(q^2)}{f_+(0)}, \tag{6.6}$$

$$\delta = 1 - \frac{M^2_{H_s} - M^2_{\eta_s}}{f_+(0)} \left(\left.\frac{df_+}{dq^2}\right|_{q^2=0} - \left.\frac{df_0}{dq^2}\right|_{q^2=0} \right), \tag{6.7}$$

$$\frac{1}{\beta} = \frac{M^2_{H_s} - M^2_{\eta_s}}{f_+(0)} \left.\frac{df_0}{dq^2}\right|_{q^2=0}, \tag{6.8}$$

where in Equation (6.6) the residue is defined in the usual way, in this case the effect being to remove the pole at $q^2 = M^2_{H^*_s}$.

The shape of form factors for physical processes where experimental measurements exist is a strong test of QCD, as we saw for $D \to K$ in Figure 5.9, and in this case we can use our full QCD calculation to test HQET.

The left hand side of Figure 6.6 shows our results for these quantities, plotted across the full range of heavy masses from D_s to B_s. Our results for α and β are qualitatively in agreement

with expectations from HQET [174] with α and β close to unity at the heaviest masses and differing further from one as the heavy quark mass falls. We also see that δ is close to nought at the B_s end of the plot but clearly nonzero at the D_s end. For numerical results, see [2].

The form factor ratio $\frac{f_0(q^2)}{f_+(q^2)}\Big(1-\frac{q^2}{M_{B_s^*}^2}\Big)^{-1}$ is also shown in Figure 6.6 (on the right), where it is compared with the HQET expectation [176]

$$\lim_{q^2\to M_{B_s}^2}\frac{f_0(q^2)}{f_+(q^2)}\Big(1-\frac{q^2}{M_{B_s^*}^2}\Big)^{-1}=\Big(\frac{f_{B_s}}{f_{B_s^*}}\Big)\frac{1}{g_{B_s^*B_s\eta_s}}. \tag{6.9}$$

This is included in [176] as a $B \to \pi$ expectation; to test it here in $B_s \to \eta_s$ we replace B with B_s. We take the coupling $g_{B_s^*B_s\eta_s} \approx g_{B^*B\pi} = 0.56(8)$ [177], because again the light quark mass dependence seen in [177] is mild. This leads us to expect little impact from SU(3) flavour symmetry breaking in our test of Equation (6.9). This is also consistent with our observation in Figure 6.5 that SU(3) flavour symmetry breaking effects in the daughter quark affect both f_0 and f_+ at large q^2, and so there will be some cancellation of the effects in their ratio. The right hand side of Figure 6.6 shows reasonable agreement with Equation (6.9) in the limit $q^2 \to M_{B_s}^2$, as is found for $B \to \pi$ in [172]. It is worth noting that this q^2 value is of course larger than $q^2_{\rm max}$ and so extrapolating our form factors to this point is somewhat questionable, though reasonable in this context of a rough comparison.

The left hand side of Figure 6.7 tests the relationships between form factors for a changing initial state mass, (H_s) but fixed final state (η_s) with a fixed energy. In [174] it is shown that the f_0 form factor for a pseudoscalar heavy meson decay to a pseudoscalar light meson at fixed energy is inversely proportional to the square root of the heavy meson mass. This scaling should work both at small energy, close to zero recoil, and also at large energy, high recoil. In [174] this is used to compare $B \to \pi$ and $D \to \pi$ decay. Here we compare $B_s \to \eta_s$ to $H_s \to \eta_s$ for variable H_s mass from D_s upward.

Figure 6.7 (left) compares $f_0(H_s \to \eta_s(E))/f_0(B_s \to \eta_s(E))$ to the expectation $\sqrt{M_{B_s}/M_{H_s}}$ given by the black line. We include an error in the HQET expectation from higher-order HQET terms of $\pm\sqrt{\frac{M_{B_s}}{M_{H_s}}}\Lambda_{\rm QCD}|M_{H_s}^{-1}-M_{B_s}^{-1}|$. Results are shown at two energies: the blue line and error band give results at zero recoil ($E_{\rm min} = M_{\eta_s}$) and the red line and error band give results at a higher energy, the maximum energy available to an η_s in a D_s decay [$E_{\rm max} = (M_{D_s}^2 + M_{\eta_s}^2)/2M_{D_s} = 1.105\,\text{GeV}$]. Our results at both energies are flatter than the $\sqrt{1/M_{H_s}}$ expectation, indicating that sizeable

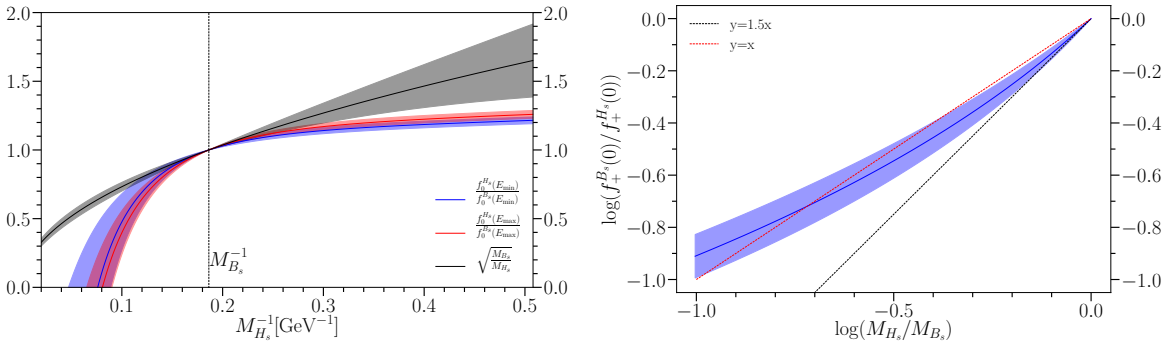

Figure 6.7: Left: the form factor ratio $\frac{f_0^{Hs}(q^2(E))}{f_0^{Bs}(q^2(E))}$ evaluated at η_s energy $E = E_{\rm min} = M_{\eta_s} = 0.6885(22)\,\text{GeV}$ (blue line and error band) and at $E_{\rm max}$ corresponding to the largest energy available to the η_s in a D_s decay (red line and error band). Both ratios are plotted over a range of inverse heavy meson masses up to $M_{D_s}^{-1}$. The black dashed line marks $M_{H_s}^{-1} = M_{B_s}^{-1}$. Results are compared with the expectation of $\sqrt{\frac{M_{Bs}}{M_{Hs}}}$ [174], given by the black band (see text). Right: the form factor ratio $\frac{f_+^{Bs}(0)}{f_+^{Hs}(0)}$ plotted against the meson mass ratio M_{H_s}/M_{B_s} in a log-log plot. Our results are shown as a blue curve with error band. The HQET expectation that the form factor ratio should depend on the 3/2 power of the mass ratio is shown as a black dashed line. In contrast, the red dashed line shows linear dependence on the mass. Results for the D_s meson correspond to the left-hand end of the plot, $\log(M_{D_s}/M_{B_s}) = -1.003$.

corrections are needed to this expectation to describe the physical behaviour, possibly owing to SU(3) symmetry breaking.

Finally, large-recoil scaling laws [178, 179] give the prediction $\frac{f_+^{B_s}(0)}{f_+^{H_s}(0)} = (\frac{M_{H_s}}{M_{B_s}})^{3/2}$ at leading order. We examine this on the right hand side of Figure 6.7, showing our results as a blue band and the HQET expectation as a black dashed line. We see that indeed the HQET expectation is borne out in the large heavy mass region close to the M_{B_s}. There are large corrections away from this region, however. We find $\frac{f_+^{B_s}(0)}{f_+^{D_s}(0)} = 0.402(33)$ which is almost twice the size of the expected $(\frac{M_{D_s}}{M_{B_s}})^{3/2} = 0.222$ [29].

As an additional observation, we find that the $n = 0$ coefficient of the HQET-QCD matching, $\rho_0^{0,+} = -0.582(48)$ is well determined and $\mathcal{O}(1)$, as expected [125]. Other $\rho_n^{0,+}$ values are consistent with 0.

6.4 Conclusions

To conclude, we have performed a fully relativistic heavy-HISQ calculation of the scalar and vector form factors for the unphysical $B_s \to \eta_s$ decay, which provides a proof of concept for the extension to other $b \to s$ decays with light spectators. We find that our modified z expansion and HQET inspired fit form suitably describes the data, and the final form factor uncertainties are statistics dominated and of the order 10% and below. All of this provides the reassurance necessary to proceed with a much more expensive, higher statistics calculation of the $B \to K$ form factors, which is outlined in the next Chapter.

Additionally, our form factors allow for some examination of SU(3) symmetry breaking in the daughter and spectator quarks; which we shall examine further in the next Chapter; As well as comparisons with HQET predictions, mainly for $B \to \pi$, with which we find varying degrees of agreement.

Chapter 7

$B \to K$ form factors

As with the previous two chapters, the basis of this chapter is published [3] and [4]. The $B \to K$ calculation presented here draws heavily on the $D \to K$ calculation in Chapter 5, with which it shares data, as well as parallels in the chiral perturbation theory used. Additionally, the heavy-HISQ set up used in the $B_s \to \eta_s$ decay (Chapter 6) is exactly analogous to that used here, and we will not describe it again. Indeed, the $B_s \to \eta_s$ decay was treated as a feasibility study for this very calculation, and we are able to incorporate the data from that calculation here, via the chiral extrapolation.

7.1 Introduction

Fundamentally, our heavy-HISQ approach here relies on the smooth transition in heavy mass from $c \to s$ to $b \to s$. Whilst this is a legitimate way to calculate form factors for heavy mass

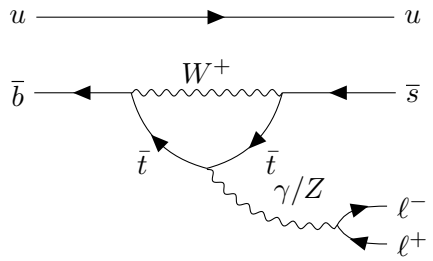

Figure 7.1: Example of a Feynman diagram for $B^+ \to K^+\ell^+\ell^-$.

$c \leq h \leq b$, it's interesting to note that the physics does change significantly. This is because there are no flavour changing neutral currents in the SM, equivalent to saying that there is no V_{bs} element in the CKM matrix. Because of this, whilst our calculation for $D \to K$ (FCCC) and $B \to K$ (FCNC) is identical, the process we are describing is somewhat different. Compare the straightforward Feynman diagram for $D \to K$ in Figure 5.1 with that[1] for $B \to K$ in Figure 7.1 This is often referred to as a penguin diagram, owing to what I'm sure you'll agree is its uncanny resemblance to everyone's favourite order of flightless Austrodyptornithes. Rather than a simple weak decay involving one instance of a CKM element, we have a process involving loops. As a result of this, the $B \to K$ is a highly suppressed 'rare' decay in the Standard Model. This means that the branching fractions are very small, and that any new physics, for example in the form of lepto-quarks or a Z', could have a relatively large effect on them. Historically, the phenomenology associated with $B \to K$ decays has been in strong tension with the experimental results, but lattice calculations have been hampered by large uncertainties, owing to the double whammy of the presence of heavy and light quarks.

The most precise, up to date results come on the theory side from two places. HPQCD '13 [180], uses form factors calculated on the lattice in [173], using HISQ light valence quarks, the MILC $N_f = 2+1$ asqtad ensembles, and a nonrelativistic QCD (NRQCD) prescription for the b quark. FNAL/MILC '15 [181] use form factors they calculated in [182], a lattice calculation using the $N_f = 2 + 1$ MILC ensembles, with light Asqtad valence and sea quarks, and the Fermilab interpretation for the b quark. Away from the lattice, other theoretical work [183, 184, 185, 186, 187, 188, 189] uses form factors from light cone sum rules [190, 191] and combinations of light cone sum rules with lattice QCD [173, 182], also including dispersive bounds [192].

On the experimental side, the most recent results, to which we will compare, come from Belle '19 [193], LHCb '14A [194], LHCb '16 [195] and LHCb '21 [9]. Previous determinations are also available [196, 197, 198, 199, 200, 201, 202, 203, 204], many of which have been superseded by the above.

[1] This is just one example.

Table 7.1: Details of the statistics, T values, heavy masses, and K meson twists used on each ensemble. The number of configurations n_{cfg} and time sources n_{src} are given, and momenta can be obtained from twist, θ, via $\theta = |a\vec{p}_K|N_x/(\sqrt{3}\pi)$, where N_x is the spatial dimension of the lattice, given in Table 3.1. In each case, the smallest am_h value corresponds to the tuned am_c mass on that ensemble, i.e. the $D \to K$ data in Chapter 5.

Set	$n_{\text{cfg}} \times n_{\text{src}}$	θ	am_h	T
1	998×16	0, 2.013, 3.050, 3.969	0.8605	9, 12, 15, 18
2	985×16	0, 2.405, 3.641, 4.735	0.643	12, 15, 18, 21
3	998×16	0, 0.8563, 2.998, 5.140	0.433, 0.683 0.8	14, 17, 20
4	985×16	0, 0.3665, 1.097, 1.828	0.888	9, 12, 15, 18
5	1053×16	0, 0.441, 1.323, 2.205, 2.646	0.664, 0.8, 0.9	12, 15, 18, 21
6	499×16	0, 0.4281, 1.282, 2.141, 2.570	0.449, 0.566, 0.683, 0.8	14, 17, 20
7	413×8	0, 1.261, 2.108, 2.946, 3.624	0.274, 0.45, 0.6, 0.8	20, 25, 30
8	375×4	0, 0.706, 1.529, 2.235, 4.705	0.194, 0.45, 0.6, 0.8	24, 33, 40

7.2 Calculation details

7.2.1 Simulation details

For this calculation, data from all eight of the MILC $N_f = 2 + 1 + 1$ gluon ensembles discussed in Section 3.4 and listed in Table 3.1 was included. That means five ensembles where $m_s/m_l = 5$ and three (with approximately the same lattice spacings as the coarsest of the five), where m_l is physical. As with $D \to K$ above, these physical ensembles are important as we have a valence light quark, the mass of which will need to be extrapolated to the physical point using chiral perturbation theory, more details of which will be given below. The twists applied on each ensemble, as well as the T values (i.e. the source sink separation in lattice units) are given in Table 7.1, along with the statistics used on each ensemble in terms of the number of configurations n_{cfg} and the number of time source (t_0) values n_{src} used on each configuration. As in Chapter 6, we have various different heavy mass (am_h) values on each ensemble, in each case with the smallest being the tuned charm mass, data at this mass is identical to that used for $D \to K$, with the exception that we calculate the tensor form factor here as well as the scalar and vector[2].

[2]No tensor data was calculated on sets 1 and 2.

7.2.2 Correlation functions

Here we calculate f_0, f_+ and additionally, the tensor form factor f_T (Equations (4.56), (4.55) and (4.57)). As before, in the language of Chapter 4, we require a Goldstone K and H with spin-taste $\gamma_5 \otimes \xi_5$, and also a non-Goldstone $\hat{H}$, with spin-taste $\gamma_5\gamma_0 \otimes \xi_5\xi_0$. These are combined with currents with scalar $(1 \otimes 1)$ and temporal vector $(\gamma_0 \otimes \xi_0)$ spin-taste structure to generate the required two- and three-point functions. Additionally, we now require a non-Goldstone $\hat{K}$ with spin-taste $\gamma_5 \otimes \xi_5\xi_1$, which is combined with $\hat{H}$ and the T^{10} tensor current $\gamma_1\gamma_0 \otimes \xi_1\xi_0$. See Section 4.2.1 for details on how these operators are implemented in the MILC code.

7.2.3 Spatial and temporal vector components

In the chapters above, we have always calculated matrix elements only for the temporal element of the vector current, V^0. This is because this element is sufficient to calculate Z_V at zero recoil (see Equation (4.58) and (4.59)). However, this component presents challenges towards q^2_{max}, owing to the construction of Equation (4.55), which can be rewritten (for $B \to K$),

$$f_+(q^2) = \frac{1}{A^\mu - B^\mu}(Z_V\langle K|V^\mu|\hat{H}\rangle - f_0(q^2)B^\mu) \tag{7.1}$$

where $A^\mu = p^\mu_H + p^\mu_K$ and $B^\mu = \frac{M^2_H - M^2_K}{q^2}q^\mu$. As the kaon momentum vanishes (i.e. towards q^2_{max}), A^0 and B^0 both $\to M_H + M_K$. The numerator of Equation (7.1) also vanishes, thanks to Equations (4.59) and (4.55). The problem with this is that the numerator becomes a small number with a large relative uncertainty, which is then divided by a small number, causing the uncertainty on f_+ to blow up. This is clear in Figure 6.2 in the above chapter, and will be demonstrated again here. It is the main reason that our uncertainty in the $B_s \to \eta_s$ and $D \to K$ calculations above is at its maximum for $f_+(q^2_{\text{max}})$.

If we take a spatial component of the vector current, say V^1, then as momentum vanishes, A^1 and B^1 both vanish also. However, this does not happen as fast as for V^0. The numerator does not suffer from the same problem of shrinking rapidly and ending up with a large relative uncertainty, either, and both these effects lead to much smaller uncertainties in $f_+(q^2_{\text{max}})$. To this end, we include additional V^1 data for the lowest two twists on sets 7 and set 8, for the largest two and three masses respectively (see appendix of [3] for numerical results).

The comparison between f_+ values for the spatial and temporal data is made in Figure 7.2. We see that the spatial data is much more precise, but that the central values of both components

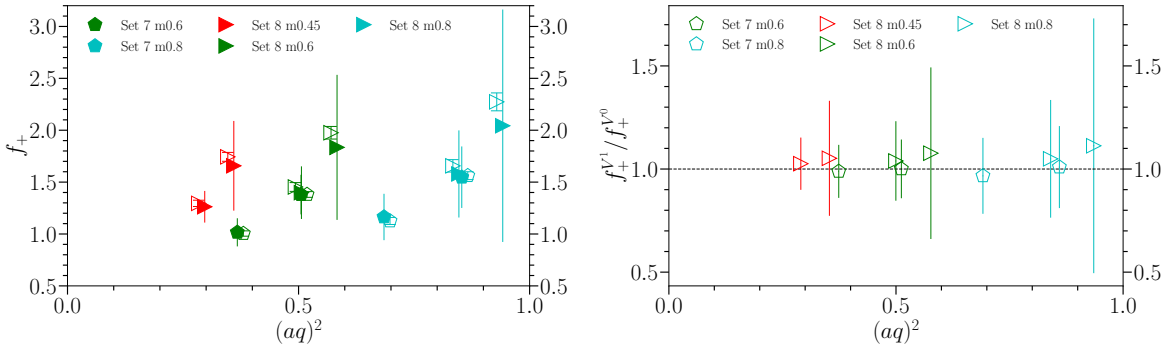

Figure 7.2: Left: a comparison of values and their statistical errors for the vector form factor derived from matrix elements for the spatial and temporal vector currents on ensembles where both are available. The filled symbols are the temporal vector results and the open symbols the spatial vector results. We have offset spatial vector results slightly on the q^2-axis for clarity. Right: the ratio of the f_+ values for the spatial and temporal vector cases. We see no evidence of any differences between them (within our uncertainties) that would indicate discretisation effects.

agree very well, indicating that discretisation effects separating them are very small.

To include the V^1 elements, we simply use a current insertion with spin-taste $\gamma_1 \otimes \xi_1$, which is combined with the $\hat{K}$ and H defined above.

7.2.4 Correlator fits

As usual, two- and three-point correlator fits are performed using the recipe in Section 4.3, using the standard goodness of fit metrics of the χ^2 and log(GBF) value described there. Priors are estimated in the same way as the chapters above, making use of the relativistic dispersion relation of the kaon (Equations (4.47) and (4.48)), as was done in Chapter 5. Details of priors and other fit parameters, as well as correlator fit stability, are all presented in [3]. This is all very similar to the previous chapters, so we do not present it here in the interests of succinctness. There are, however, some subtleties which are unique to this calculation which we will address. For numerical results from correlator fits, see [3].

Unbinning ultrafine timesources

On all ensembles, with the exception of the very finest, set 8, averaging over t_0 takes place before the covariance matrix is calculated (as per Section 4.3.4). That is to say we have n_{cfg} samples,

each of which is an average of the data from the n_{src} source times used on that ensemble. This averaging takes place before the covariance matrix is calculated, as the gvar [87] function which generates the covariance matrix requires that the samples are uncorrelated in order to calculate the covariance matrix, and this is not necessarily the case for different source times on a single configuration. On set 8, however, we have a very large $N_t = 192$, and small $n_{\text{src}} = 4$, meaning that t_0 values are very widely spaced. Testing on this ensemble confirms that data from different t_0 values is not correlated, so we choose not to average over time sources. Evidence of this is presented in Figure 7.3, which shows the first 20 data points of four two-point correlators (see caption for details), spanning the range of masses and twists. The correlators are normalised by e^{-Mt} (for fit result M) and have had their error bars inflated 500 times for clarity. We see in each case that the uncertainties with and without binning over timesources are nearly identical. Quantitatively, averaging the ratio of the unbinned to binned uncertainties over all t values for all two-point correlators gives 0.994. This indicates that the unbinning procedure has not neglected any correlations, that is to say that the different t_0 values are not correlated. As an aside, this plot also nicely demonstrates the prevalence of oscillations in $\hat{H}$ and the increasing error for K at large twist.

This procedure is equivalent, in the context of the other ensembles, to saying that we have $n_{\text{cfg}} \times n_{\text{src}}$ configurations each with one source. In terms of the actual numbers going into the fit, this makes no difference to anything - we're still ultimately averaging over the same exact data, we're just calculating the covariance matrix with more uncorrelated samples. The main benefit is that by increasing the sample size, we are able to reduce the SVD cut, and so improve the uncertainties on the final fit. I repeat that this is only possible because these time sources are uncorrelated. This approach is particularly required here, where we have relatively poor statistics. For $D \to K$, the situation was not so bad, but for $B \to K$, where there are many more fit parameters on account of the variable heavy mass, this unbinning is essential. Without it, the fit behaves well enough, but some parameters have very large uncertainties. This is because the SVD cut has affected these eigenvalues particularly. Unbinning the data, and the accompanying reduction in SVD cut, gives these fit parameters much more sensible values, whilst the fit parameters which were accurately determined in the binned fit are unaffected.

This is demonstrated in Figure 7.4, where we plot the form factor data points (with the pole removed) in z space. The data points shown in red (labelled 'ub') result from correlator fits to the unbinned data on set 8, as described above. This is the data we use in our calculation. The points in black ('b'), meanwhile, are from identical fits, but where the data is binned over

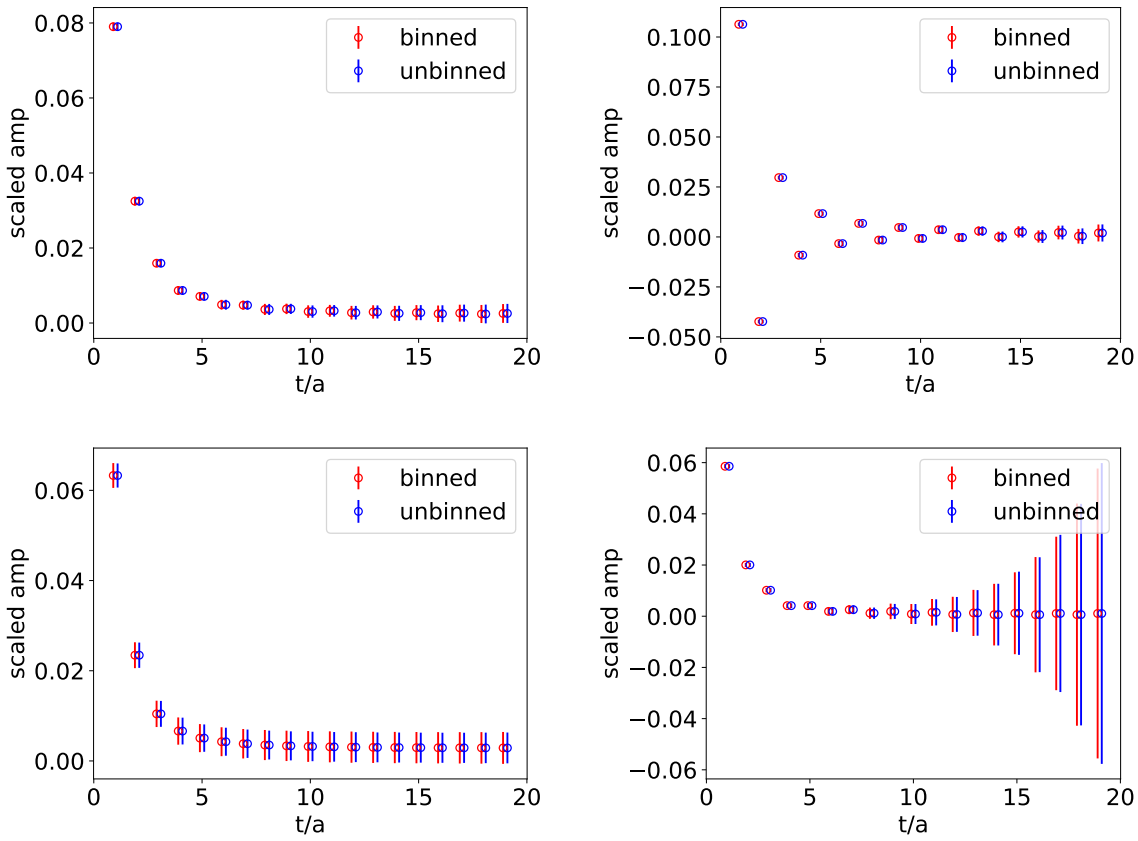

Figure 7.3: Comparison of raw correlation function data from binned and unbinned data on set 8. In each case, data is normalised by e^{-Mt}, offset from integer t/a values and has errors inflated 500 times for clarity. Clockwise from top left, we have plots for $H(am_h = 0.194)$, $\hat{H}(am_h = 0.8)$, $\hat{K}(\theta = 4.705)$, $K(\theta = 0)$.

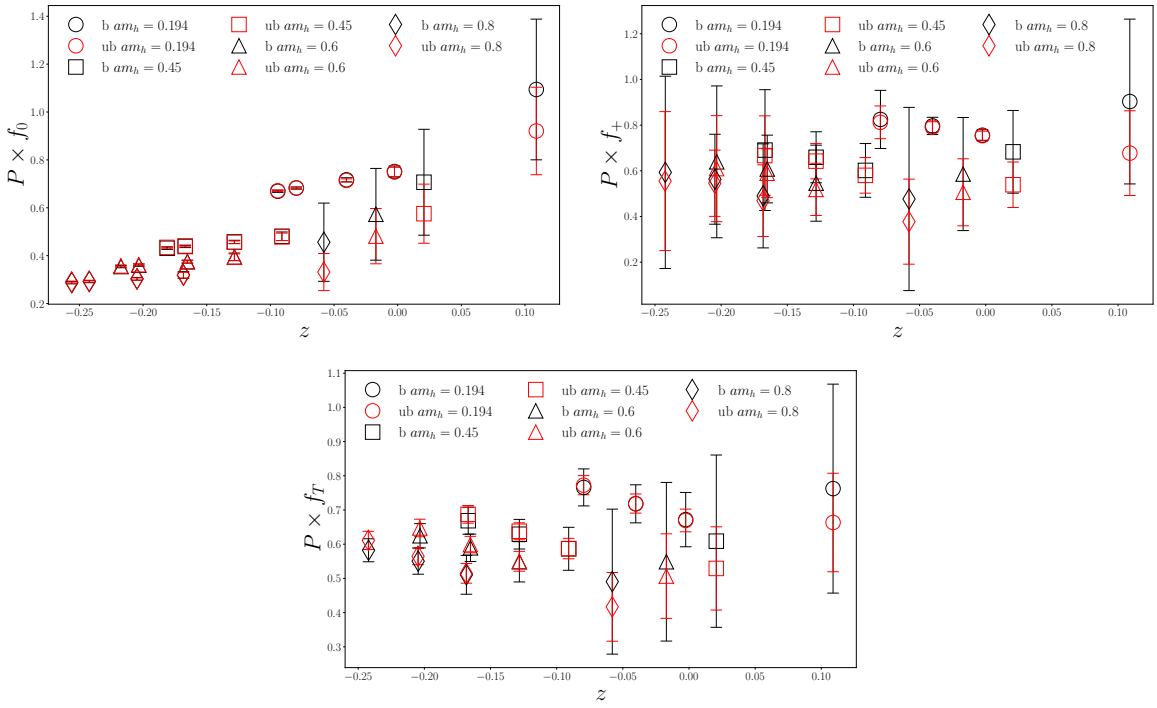

Figure 7.4: f_0,f_+ and f_T results in z space for $B \to K$ on set 8. In red are the results we use, from the unbinned ('ub') data. For comparison in black are the results from the binned ('b') data.

the $n_{\rm src} = 4$ time sources on each configuration before the covariance matrix is calculated[3], and hence we have fewer samples, and so require a larger SVD cut. As is clear from these plots, whilst both fits agree to within error, the large SVD cut required for the binned data results in large uncertainties, particularly at large z values.

$B_s \to \eta_s$ data

As it stands, we have data with two different light masses. On sets 1, 2 and 3, $m_l = m_l^{\rm phys}$ and on the others $m_l = m_s/5$. These values will both be used in the chiral extrapolation below (as for $D \to K$) to tune the light quark to its physical value. This extrapolation could be improved with a third light mass, and it just so happens that we have one. Our data for $B_s \to \eta_s$ in Chapter 6 is just the same as $B \to K$ data with $m_l = m_s$. Indeed, this is no great secret, being exactly why we used that unphysical decay as a test for $B \to K$ in the first place. In order to include this data faithfully, we must ensure it is properly correlated with our $m_l = m_s/5$ data on the same ensembles. This means using exactly the same configurations, and performing simultaneous correlator fits. For this reason, we only use $B_s \to \eta_s$ data for sets 6 and 7, as this had all the same configurations available. We left out the Set 8 data as many fewer configurations were used for $B_s \to \eta_s$, which resulted in large uncertainties anyway. For clarity, in some figures the set 6 and set 7 $B_s \to \eta_s$ data will be labelled as set 9 and 10 respectively. The $B_s \to \eta_s$ and $B \to K$ data is fitted simultaneously, but the enormous size of the fits require a form of chaining, which we outline below.

Separated mass fits

On sets 6, 7 and 8, we have very large fits, owing to the number of twists and masses, the number of t values on the ensembles, and the presence of 'extra' spatial vector and $B_s \to \eta_s$ data. This makes these fits intractable in any reasonable time, and so they must be split up, but in such a way that correlations are respected. Firstly, we run the fits without any 'extra' data and at a smaller number of exponentials (such that is is possible to run a fit), to generate control data. We then plot a distribution of the correlations between matrix elements of different mass and twist. This gives us an idea of the size of correlations, and we can compare this with the correlations in a comparable fit which is split up.

The left hand side of Figure 7.5 shows these correlations for set 8. The correlations between

[3] As is the case for all other ensembles in this calculation

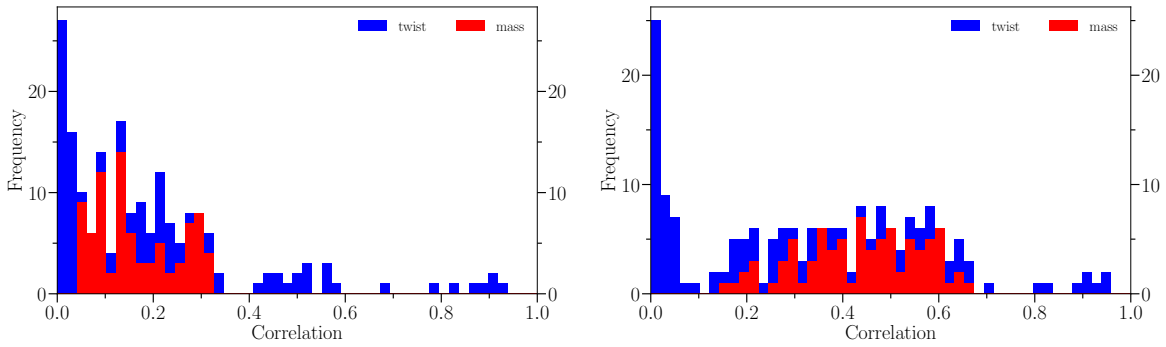

Figure 7.5: Correlations between matrix elements for a simultaneous fit (left) and one which is split up according to heavy mass (right) on set 8. Correlations between different twists at fixed mass are shown in blue, and different masses at fixed twist in red. The bars are additive (i.e. blue sits above red, not behind it). See text for more details.

$J_{00}^{\rm nn}$ matrix elements for different currents, masses and twists are shown. For example, 'twist' (in blue) includes correlations between currents of the same mass and different twists, such as $S_{00}^{\rm nn}(m_h = 0.194, \theta = 0)$, $S_{00}^{\rm nn}(m_h = 0.194, \theta = 0.706)$, $S_{00}^{\rm nn}(m_h = 0.194, \theta = 1.529)$, e.t.c. and similarly for $V_{00}^{0,{\rm nn}}$ and $T_{00}^{\rm nn}$ (but not correlations between the different currents S,V^0,T). Similarly, correlations between masses at fixed twist are shown in red. We do not show correlations between different matrix elements, or those with different masses and twists, as these are small, and mask the important correlations.

From this plot, it is clear that correlations between different masses are in general smaller than those between different twists (and in this case all less than 0.4), so we split our fits up according to heavy mass. We fit each heavy mass sequentially, and then average any shared parameters (e.g. the kaon). This process is almost identical to the chaining described in the previous chapter, except that posteriors never appear as priors in the fit. It greatly speeds up the fits and makes them tractable.

To check that this approach preserves correlations, we repeat the above procedure to produce a plot for the case where the fit is split up according to heavy mass. This is shown on the right of Figure 7.5. We see that the correlations between different twists are well preserved, whilst the correlations between different masses have been increased somewhat, though they are still quite small[4]. As increasing correlations is generally a conservative move, this means that doing the fit

[4] A good rule of thumb is that correlations below about 0.8 do not affect the form factors too much.

in this way will not invalidate our results, though it may increase our uncertainties. As with the SVD cuts described earlier, sometimes it is necessary to take a small increase in uncertainty as a price to make a large, correlated fit possible.

7.2.5 z expansion

As usual, our z expansion is based on Section 4.4.2, with $t_0 = 0$, and we use the basic format of Equations (4.61), (4.62) and (4.63), once again finding $N = 3$ to be sufficient. This time however, we need to include both chiral logarithm terms, as in Chapter 5, and account for the heavy mass dependence, as in Chapter 6.

$$\begin{aligned} f_0(q^2) &= \frac{\mathcal{L}}{1 - \frac{q^2}{M^2_{H^*_{s0}}}} \sum_{n=0}^{N-1} a_n^0 z^n \\ f_+(q^2) &= \frac{\mathcal{L}}{1 - \frac{q^2}{M^2_{H^*_s}}} \sum_{n=0}^{N-1} a_n^+ \Big(z^n - \frac{n}{N}(-1)^{n-N} z^N \Big) \\ f_T(q^2) &= \frac{\mathcal{L}}{1 - \frac{q^2}{M^2_{H^*_s}}} \sum_{n=0}^{N-1} a_n^T \Big(z^n - \frac{n}{N}(-1)^{n-N} z^N \Big). \end{aligned} \tag{7.2}$$

We take $M_{H^*_{s0}}$ to be $M_H + \Delta$ with $\Delta = 0.45$GeV. As discussed in [2], the exact value used here is unimportant, as the pole is very far from $q^2_{\rm max}$, so any change is easily absorbed into the shape of the z expansion. The value of Δ is taken from experimental results for the D system [29]; there are no experimental results for the B system but we expect the splitting to be largely independent of m_h. The vector mass $M_{H^*_s}$ can be estimated, as in [2, 147], with the PDG [29] values $M^{\rm phys}_{D^*_s} = 2.1122(4)$GeV, $M^{\rm phys}_{B^*_s} = 5.4158(15)$GeV. For this we use a mass dependent expression like that in Chapter 6,

$$M_{H^*_s} = M_H + \frac{M^{\rm phys}_D}{M_H}\Delta(D) + \frac{M^{\rm phys}_B}{M_H}\Big(\frac{M_H - M^{\rm phys}_D}{M^{\rm phys}_B - M^{\rm phys}_D}\Big(\Delta(B) - \frac{M^{\rm phys}_D}{M^{\rm phys}_B}\Delta(D)\Big)\Big), \tag{7.3}$$

where $\Delta(H) = M^{\rm phys}_{H^*_s} - M^{\rm phys}_H$[5].

Within Equation (7.2) our a_n expressions again take a slightly different form to that seen previ-

[5]The physical masses used are $(K^0 + K^\pm)/2$, $(B^0 + B^\pm)/2$ and $(D^0 + D^\pm)/2$ (all from [29]), as our lattice results have $m_u = m_d = m_l$.

ously.

$$\begin{aligned} a_n^{0,+,T} =& \Big(\frac{M_D}{M_H}\Big)^{\zeta_n}\Big(1+\rho_n^{0,+,T}\log\Big(\frac{M_H}{M_D}\Big)\Big)(1+\mathcal{N}_n^{0,+,T}) \\ &\times \sum_{i,j,k,l=0}^{N_{ijkl}-1} d_{ijkln}^{0,+,T}\Big(\frac{\Lambda_{\rm QCD}}{M_H}\Big)^i\Big(\frac{am_h^{\rm val}}{\pi}\Big)^{2j}\Big(\frac{a\Lambda_{\rm QCD}}{\pi}\Big)^{2k}(x_\pi - x_\pi^{\rm phys})^l, \end{aligned} \tag{7.4}$$

where $N_{ijkl} \equiv (N_i, N_j, N_k, N_l) = (3,2,2,3)$. We will address the different elements of this separately below.

Discretisation Effects

Discretisation effects are accounted for in two ways in Equation (7.4). We allow for heavy quark mass dependent effects through the terms in am_h with power $2j$. The size of these terms will vary between results for different m_h on a given ensemble, and of course between different form factors. Discretisation effects that do not vary with heavy quark mass but instead are set by some other scale (for example associated with the K mesons) are allowed for in the powers of $a\Lambda_{\rm QCD}$ and do not carry heavy mass dependence.

Dependence on heavy quark mass

We also include several terms in Equation (7.4) to model the physical dependence of the form factors on heavy quark mass, as we did with $B_s \to \eta_s$ in Chapter 6. This time we will be connecting the form factors for $D \to K$ to those for $B \to K$, and we include the same polynomial terms in $\frac{\Lambda_{\rm QCD}}{M_H}$, as well as the logarithmic term to account for the matching of HQET to QCD.

The main difference is the addition of a multiplicative prefactor $(M_D/M_H)^{\zeta_n}$, with fitted power ζ_n. This $(M_D/M_H)^{\zeta_n}$ term models behaviour predicted by Large Energy Effective Theory (LEET) [205], and the extra prior knowledge is especially helpful here because it links the tensor form factor, which has the largest uncertainties, to the better determined scalar and vector form factors.

The LEET expectation is for all form factors for a specific heavy to light transition to exhibit common $\backsim M_H^{-3/2}$ behaviour in the region of $q^2 = 0$ (where the light meson energy is close to $M_H/2$). This behaviour was in fact observed for $B_s \to \eta_s$ in Figure 6.7, with an M_{H_s} power between -1.5 and -1 towards $q^2 = 0$. In that case, this dependence was absorbed in the fit by the logarithmic term, but here we fit it explicitly.

Because we have taken $t_0 = 0$, $q^2 = 0$ maps to $z = 0$. We therefore include the LEET expectation at $q^2 = 0$ via a prior $P[\zeta_0] = 1.5(5)$ as a common factor for the a_0 coefficients and set $\zeta_{n\neq 0} = 0$ for $a_{n\neq 0}$. M_K/M_H corrections to LEET can be accounted for in the form factor dependent $(\Lambda_{\rm QCD}/M_H)^i$ terms in our fit, as $M_K \approx \Lambda_{\rm QCD}$. We find that including this term in our fit increases log(GBF), reduces uncertainty at $q^2 = 0$, particularly for f_T, and returns a posterior of $\zeta_0 = 1.43(12)$. Allowing a broader prior $P[\zeta_0] = 1.0(1.0)$ returns an almost identical posterior consistent with 1.5 (1.42(12)) and does not change the form factor result. Allowing ζ to vary between form factors simply increases the uncertainty on $f_T(0)$, whilst leaving the central values unchanged. These tests of the form factors, along with others, will be shown below (Figure 7.6).

Dependence on spectator quark mass

As in $D \to K$ (Chapter 5), we here must perform an extrapolation in spectator quark mass, to allow for the the fact that some of our data does not have physical light quarks. The additional change here is that we not only have $m_l = m_l^{\rm phys}$ and $m_l = m_s/5$ data, but also some $m_l = m_s$ data ($B_s \to \eta_s$), as discussed in Section 7.2.4.

As with $D \to K$, we include this dependence via a chiral logarithm, as outlined in Section 4.4.2. The difference is that, owing to the larger range of light masses to include the $B_s \to \eta_s$ data, we include more terms. We also adopt the notation $\mathcal{L}$ to be consistent with [3].

$$\mathcal{L} = 1 - \frac{9g^2}{8} x_\pi \Big(\log x_\pi + \delta_{FV} \Big) - \Big(\frac{1}{2} + \frac{3g^2}{4} \Big) x_K \log x_K - \Big(\frac{1}{6} + \frac{g^2}{8} \Big) x_\eta \log x_\eta, \tag{7.5}$$

where $x_{\mathcal{M}} = \frac{M_{\mathcal{M}}^2}{(4\pi f_\pi)^2}$ as usual, and g is the coupling between H, H^* and the light mesons. This is now mass dependent. The form of $\mathcal{L}$ is appropriate for the vector and scalar form factors and, as in [104], we make use of the fact that f_T and f_+ in HQET are the same up to $\mathcal{O}(1/M_H)$ terms to use the same $\mathcal{L}$ for the tensor form factor. Any corrections to this are easily absorbed by our HQET expansion. We take $x_\pi = 2\frac{m_{\rm spectator}}{5.63 m_s}$ where the factor of 2 accounts for the definition of f_π^2 in [103] and we construct x_K and x_η similarly. $\delta_{\rm FV}$ is discussed in more detail in Section 4.4.2. In order to capture the heavy mass dependence of g, we take

$$g(M_H) = g_\infty + C_1 \frac{\Lambda_{\rm QCD}}{M_H} + C_2 \frac{\Lambda_{\rm QCD}^2}{M_H^2}, \tag{7.6}$$

including this dependence in the z expansion fit, with data points $g_\infty = 0.48(11)$ [206], $g(M_D) =$

0.570(6) [148] and $g(M_B) = 0.500(33)$, an average of the values in [177, 207, 208]. Priors $\mathcal{P}[C_1] = 0.5(1.0)$ and $\mathcal{P}[C_2] = 0.0(3.0)$ are broad and based on a trial fit to just the g data points given above. Our final fit has a tighter value for g_∞, giving posterior $g_\infty = 0.457(56)$ with coefficients $C_1 = 0.73(62)$ and $C_2 = -1.2(1.7)$. $g(M_D)$ and $g(M_B)$ are not significantly changed. For more details on all priors, see Section 7.2.5.

As well as the chiral logarithm term $\mathcal{L}$ that is common to all terms in the z-expansion, we include analytic terms in the spectator quark mass that can vary for different form factors and with the power of z, n. These appear through powers of $(x_\pi - x_\pi^{\text{phys}})$ in Equation (7.4), where $x_\pi^{\text{phys}} = 2\frac{m_l^{\text{phys}}}{5.63 m_s^{\text{phys}}}$.

Mistuning effects for other quark masses

The pattern for quark mistuning here should by now be very familiar, and we will simply state the form used in this case, noting that the mass of the strange daughter quark is always the valence s quark mass, listed in Table 3.1. We take

$$\begin{aligned}\mathcal{N}_n^{0,+,T} = c_{s,n}^{\text{val},0,+,T}\delta_s^{\text{val}} + c_{s,n}^{\text{sea},0,+,T}\delta_s^{\text{sea}} + 2c_{l,n}^{\text{sea},0,+,T}\delta_l \\ + c_{c,n}^{\text{sea},0,+,T}\delta_c^{\text{sea}}.\end{aligned} \tag{7.7}$$

Everything has been defined before (Equations (5.13) and (5.14)), save for the sea charm quarks, where we define

$$\delta_c^{\text{sea}} = \frac{m_c^{\text{sea}} - m_c^{\text{tuned}}}{m_c^{\text{tuned}}}, \tag{7.8}$$

with $m_c^{\text{tuned}} = m_c^{\text{val}}$, which is fixed in [80].

z expansion priors and posteriors

Tables 7.2-7.6 provide prior and posterior values for our final z expansion fit. We use our standard method for determining such priors, introduced in Chapter 5: the z expansion is constructed such that we expect a_n values to be of order unity. Thus, we take most priors as 0(1). In some cases we tighten up priors, for example when they relate to terms containing discretisation effect which go like a^2. We know such effect are highly suppressed in the HISQ action, so we give them narrower priors. After such adjustments, we apply the empirical Bayes method discussed in Section 4.3.5 to broad groups of priors, for example c_l^{sea}. We then adjust the widths if the study suggests they are overly conservative, in this case down to 0.0(5). Generally, we do not

Table 7.2: Priors and their posterior values for the z expansion fit.

Parameter	Prior	Posterior	Parameter	Prior	Posterior
g_∞	0.48(11)	0.457(56)	d^0_{01112}	0.0(1.0)	0.0005(1.0000)
C_1	0.5(1.0)	0.73(62)	d^0_{01120}	0.0(1.0)	0.009(1.000)
C_2	0.0(3.0)	-1.2(1.7)	d^0_{01121}	0.0(1.0)	-0.002(1.000)
ρ^0_0	0.0(1.0)	0.87(29)	d^0_{01122}	0.0(1.0)	0.0002(1.0000)
ρ^0_1	0.0(1.0)	-0.62(13)	d^0_{10000}	0.0(1.0)	-0.29(57)
ρ^0_2	0.0(1.0)	-0.92(26)	d^0_{10001}	0.0(1.0)	0.50(59)
ρ^+_0	0.0(1.0)	0.87(29)	d^0_{10002}	0.0(1.0)	0.004(932)
ρ^+_1	0.0(1.0)	-0.01(27)	d^0_{10010}	0.0(1.0)	0.32(49)
ρ^+_2	0.0(1.0)	-0.22(92)	d^0_{10011}	0.0(1.0)	-0.08(83)
ρ^T_0	0.0(1.0)	0.99(53)	d^0_{10012}	0.0(1.0)	-0.03(99)
ρ^T_1	0.0(1.0)	-0.03(33)	d^0_{10020}	0.0(1.0)	1.05(78)
ρ^T_2	0.0(1.0)	0.06(98)	d^0_{10021}	0.0(1.0)	-0.14(97)
d^0_{00000}	0.0(1.0)	0.61(15)	d^0_{10022}	0.0(1.0)	-0.006(0.999)
d^0_{00001}	0.0(1.0)	0.55(17)	d^0_{10100}	0.00(30)	-0.06(29)
d^0_{00002}	0.0(1.0)	0.63(35)	d^0_{10101}	0.00(30)	0.01(30)
d^0_{00010}	0.0(1.0)	-0.53(12)	d^0_{10102}	0.00(30)	-0.0008(3000)
d^0_{00011}	0.0(1.0)	-0.60(38)	d^0_{10110}	0.00(30)	-0.02(30)
d^0_{00012}	0.0(1.0)	0.04(78)	d^0_{10111}	0.00(30)	0.001(300)
d^0_{00020}	0.0(1.0)	0.54(23)	d^0_{10112}	0.00(30)	-0.0001(3000)
d^0_{00021}	0.0(1.0)	-0.25(79)	d^0_{10120}	0.00(30)	-0.004(300)
d^0_{00022}	0.0(1.0)	-0.02(97)	d^0_{10121}	0.00(30)	0.0001(3000)
d^0_{00100}	0.00(30)	-0.21(23)	d^0_{10122}	0.00(30)	-7e-06 +- 0.3
d^0_{00101}	0.00(30)	0.03(30)	d^0_{11000}	0.00(30)	-0.12(27)
d^0_{00102}	0.00(30)	-0.002(300)	d^0_{11001}	0.00(30)	0.01(30)
d^0_{00110}	0.00(30)	-0.04(30)	d^0_{11002}	0.00(30)	0.0008(2999)
d^0_{00111}	0.00(30)	0.002(300)	d^0_{11010}	0.00(30)	0.01(30)
d^0_{00112}	0.00(30)	-0.00003(30000)	d^0_{11011}	0.00(30)	-0.006(300)
d^0_{00120}	0.00(30)	-0.01(30)	d^0_{11012}	0.00(30)	0.001(300)
d^0_{00121}	0.00(30)	-0.0004(3000)	d^0_{11020}	0.00(30)	0.003(299)
d^0_{00122}	0.00(30)	0.00007(30000)	d^0_{11021}	0.00(30)	-0.003(300)
d^0_{01000}	0.00(30)	0.123(73)	d^0_{11022}	0.00(30)	0.0004(3000)
d^0_{01001}	0.00(30)	-0.03(25)	d^0_{11100}	0.0(1.0)	-0.09(99)
d^0_{01002}	0.00(30)	0.01(30)	d^0_{11101}	0.0(1.0)	0.009(1.000)
d^0_{01010}	0.00(30)	0.08(17)	d^0_{11102}	0.0(1.0)	-0.0009(1.0000)
d^0_{01011}	0.00(30)	-0.03(30)	d^0_{11110}	0.0(1.0)	0.0004(1.0000)
d^0_{01012}	0.00(30)	0.005(300)	d^0_{11111}	0.0(1.0)	-0.0001(1.0000)
d^0_{01020}	0.00(30)	0.002(281)	d^0_{11112}	0.0(1.0)	5e-05 +- 1
d^0_{01021}	0.00(30)	-0.009(299)	d^0_{11120}	0.0(1.0)	0.0009(1.0000)
d^0_{01022}	0.00(30)	0.002(300)	d^0_{11121}	0.0(1.0)	-0.0003(1.0000)
d^0_{01100}	0.0(1.0)	-0.23(92)	d^0_{11122}	0.0(1.0)	4e-05 +- 1
d^0_{01101}	0.0(1.0)	0.02(1.00)	d^0_{20000}	0.0(1.0)	0.35(87)
d^0_{01102}	0.0(1.0)	-0.002(1.000)	d^0_{20001}	0.0(1.0)	0.21(93)
d^0_{01110}	0.0(1.0)	0.02(1.00)	d^0_{20002}	0.0(1.0)	-0.03(99)
d^0_{01111}	0.0(1.0)	-0.003(1.000)	d^0_{20010}	0.0(1.0)	-0.53(83)

Table 7.3: Priors and their posterior values for the z expansion fit.

Parameter	Prior	Posterior
d^0_{20011}	0.0(1.0)	0.12(98)
d^0_{20012}	0.0(1.0)	-0.03(1.00)
d^0_{20020}	0.0(1.0)	0.16(95)
d^0_{20021}	0.0(1.0)	-0.004(0.997)
d^0_{20022}	0.0(1.0)	-0.006(1.000)
d^0_{20100}	0.00(30)	-0.02(30)
d^0_{20101}	0.00(30)	0.003(300)
d^0_{20102}	0.00(30)	-0.0003(3000)
d^0_{20110}	0.00(30)	-0.005(300)
d^0_{20111}	0.00(30)	0.0005(3000)
d^0_{20112}	0.00(30)	-0.00004(30000)
d^0_{20120}	0.00(30)	-0.001(300)
d^0_{20121}	0.00(30)	0.00007(30000)
d^0_{20122}	0.00(30)	-7e-06 +- 0.3
d^0_{21000}	0.00(30)	-0.07(30)
d^0_{21001}	0.00(30)	0.008(300)
d^0_{21002}	0.00(30)	-0.0005(3000)
d^0_{21010}	0.00(30)	-0.004(300)
d^0_{21011}	0.00(30)	-0.0004(3000)
d^0_{21012}	0.00(30)	0.0001(3000)
d^0_{21020}	0.00(30)	-0.0007(3000)
d^0_{21021}	0.00(30)	-0.0005(3000)
d^0_{21022}	0.00(30)	0.00007(30000)
d^0_{21100}	0.0(1.0)	-0.03(1.00)
d^0_{21101}	0.0(1.0)	0.003(1.000)
d^0_{21102}	0.0(1.0)	-0.0003(1.0000)
d^0_{21110}	0.0(1.0)	-0.0009(1.0000)
d^0_{21111}	0.0(1.0)	0.0001(1.0000)
d^0_{21112}	0.0(1.0)	-5e-06 +- 1
d^0_{21120}	0.0(1.0)	-5e-05 +- 1
d^0_{21121}	0.0(1.0)	-3e-05 +- 1
d^0_{21122}	0.0(1.0)	5e-06 +- 1
$c^{\text{sea},0}_{s,0}$	0.00(50)	-0.07(44)
$c^{\text{sea},0}_{s,1}$	0.00(50)	0.005(497)
$c^{\text{sea},0}_{s,2}$	0.00(50)	-0.004(500)
$c^{\text{sea},0}_{l,0}$	0.00(50)	0.015(81)
$c^{\text{sea},0}_{l,1}$	0.00(50)	-0.15(47)
$c^{\text{sea},0}_{l,2}$	0.00(50)	0.03(50)
$c^{\text{sea},0}_{c,0}$	0.00(10)	0.049(51)
$c^{\text{sea},0}_{c,1}$	0.00(10)	0.003(98)
$c^{\text{sea},0}_{c,2}$	0.00(10)	-0.002(100)
$c^{\text{val},0}_{s,0}$	0.0(1.0)	0.62(90)
$c^{\text{val},0}_{s,1}$	0.0(1.0)	-0.03(1.00)
$c^{\text{val},0}_{s,2}$	0.0(1.0)	0.0008(1.0000)

Parameter	Prior	Posterior
d^T_{00000}	0.0(1.0)	0.59(23)
d^T_{00001}	0.0(1.0)	-0.68(27)
d^T_{00002}	0.0(1.0)	0.34(76)
d^T_{00010}	0.0(1.0)	-0.06(32)
d^T_{00011}	0.0(1.0)	-0.15(94)
d^T_{00012}	0.0(1.0)	0.03(1.00)
d^T_{00020}	0.0(1.0)	0.02(1.00)
d^T_{00021}	0.0(1.0)	-0.01(1.00)
d^T_{00022}	0.0(1.0)	0.002(1.000)
d^T_{00100}	0.00(30)	-0.06(30)
d^T_{00101}	0.00(30)	0.004(300)
d^T_{00102}	0.00(30)	-0.0004(3000)
d^T_{00110}	0.00(30)	-0.003(300)
d^T_{00111}	0.00(30)	0.0002(3000)
d^T_{00112}	0.00(30)	-2e-05 +- 0.3
d^T_{00120}	0.00(30)	-0.0002(3000)
d^T_{00121}	0.00(30)	7e-06 +- 0.3
d^T_{00122}	0.00(30)	-7e-07 +- 0.3
d^T_{01000}	0.00(30)	-0.16(15)
d^T_{01001}	0.00(30)	-0.001(296)
d^T_{01002}	0.00(30)	0.001(300)
d^T_{01010}	0.00(30)	-0.01(30)
d^T_{01011}	0.00(30)	-1e-06 +- 0.3
d^T_{01012}	0.00(30)	0.00006(30000)
d^T_{01020}	0.00(30)	-0.0006(3000)
d^T_{01021}	0.00(30)	-1e-05 +- 0.3
d^T_{01022}	0.00(30)	5e-06 +- 0.3
d^T_{01100}	0.0(1.0)	-0.05(1.00)
d^T_{01101}	0.0(1.0)	0.004(1.000)
d^T_{01102}	0.0(1.0)	-0.0004(1.0000)
d^T_{01110}	0.0(1.0)	-0.003(1.000)
d^T_{01111}	0.0(1.0)	0.0002(1.0000)
d^T_{01112}	0.0(1.0)	-2e-05 +- 1
d^T_{01120}	0.0(1.0)	-0.0001(1.0000)
d^T_{01121}	0.0(1.0)	1e-05 +- 1
d^T_{01122}	0.0(1.0)	-1e-06 +- 1
d^T_{10000}	0.0(1.0)	-0.31(84)
d^T_{10001}	0.0(1.0)	-0.04(93)
d^T_{10002}	0.0(1.0)	0.04(99)
d^T_{10010}	0.0(1.0)	0.07(95)
d^T_{10011}	0.0(1.0)	-0.04(1.00)
d^T_{10012}	0.0(1.0)	0.005(1.000)
d^T_{10020}	0.0(1.0)	0.01(1.00)
d^T_{10021}	0.0(1.0)	-0.003(1.000)
d^T_{10022}	0.0(1.0)	0.0004(1.0000)

Table 7.4: Priors and their posterior values for the z expansion fit.

Parameter	Prior	Posterior
d^T_{10100}	0.00(30)	-0.02(30)
d^T_{10101}	0.00(30)	0.001(300)
d^T_{10102}	0.00(30)	-0.0001(3000)
d^T_{10110}	0.00(30)	-0.0008(3000)
d^T_{10111}	0.00(30)	0.00006(30000)
d^T_{10112}	0.00(30)	-6e-06 +- 0.3
d^T_{10120}	0.00(30)	-0.00004(30000)
d^T_{10121}	0.00(30)	3e-06 +- 0.3
d^T_{10122}	0.00(30)	-3e-07 +- 0.3
d^T_{11000}	0.00(30)	-0.06(29)
d^T_{11001}	0.00(30)	0.002(300)
d^T_{11002}	0.00(30)	-0.00008(30000)
d^T_{11010}	0.00(30)	-0.004(300)
d^T_{11011}	0.00(30)	0.00009(30000)
d^T_{11012}	0.00(30)	-2e-06 +- 0.3
d^T_{11020}	0.00(30)	-0.0002(3000)
d^T_{11021}	0.00(30)	3e-06 +- 0.3
d^T_{11022}	0.00(30)	2e-07 +- 0.3
d^T_{11100}	0.0(1.0)	-0.01(1.00)
d^T_{11101}	0.0(1.0)	0.001(1.000)
d^T_{11102}	0.0(1.0)	-0.0001(1.0000)
d^T_{11110}	0.0(1.0)	-0.0008(1.0000)
d^T_{11111}	0.0(1.0)	7e-05 +- 1
d^T_{11112}	0.0(1.0)	-7e-06 +- 1
d^T_{11120}	0.0(1.0)	-4e-05 +- 1
d^T_{11121}	0.0(1.0)	3e-06 +- 1
d^T_{11122}	0.0(1.0)	-3e-07 +- 1
d^T_{20000}	0.0(1.0)	-0.30(94)
d^T_{20001}	0.0(1.0)	0.05(99)
d^T_{20002}	0.0(1.0)	-0.001(1.000)
d^T_{20010}	0.0(1.0)	0.03(99)
d^T_{20011}	0.0(1.0)	-0.008(1.000)
d^T_{20012}	0.0(1.0)	0.0008(1.0000)
d^T_{20020}	0.0(1.0)	0.003(1.000)
d^T_{20021}	0.0(1.0)	-0.0005(1.0000)
d^T_{20022}	0.0(1.0)	6e-05 +- 1
d^T_{20100}	0.00(30)	-0.004(300)
d^T_{20101}	0.00(30)	0.0004(3000)
d^T_{20102}	0.00(30)	-0.00005(30000)
d^T_{20110}	0.00(30)	-0.0002(3000)
d^T_{20111}	0.00(30)	2e-05 +- 0.3
d^T_{20112}	0.00(30)	-2e-06 +- 0.3
d^T_{20120}	0.00(30)	-1e-05 +- 0.3
d^T_{20121}	0.00(30)	9e-07 +- 0.3

Parameter	Prior	Posterior
d^T_{20122}	0.00(30)	-1e-07 +- 0.3
d^T_{21000}	0.00(30)	-0.02(30)
d^T_{21001}	0.00(30)	0.001(300)
d^T_{21002}	0.00(30)	-0.0001(3000)
d^T_{21010}	0.00(30)	-0.001(300)
d^T_{21011}	0.00(30)	0.00005(30000)
d^T_{21012}	0.00(30)	-5e-06 +- 0.3
d^T_{21020}	0.00(30)	-0.00005(30000)
d^T_{21021}	0.00(30)	2e-06 +- 0.3
d^T_{21022}	0.00(30)	-2e-07 +- 0.3
d^T_{21100}	0.0(1.0)	-0.004(1.000)
d^T_{21101}	0.0(1.0)	0.0004(1.0000)
d^T_{21102}	0.0(1.0)	-4e-05 +- 1
d^T_{21110}	0.0(1.0)	-0.0002(1.0000)
d^T_{21111}	0.0(1.0)	2e-05 +- 1
d^T_{21112}	0.0(1.0)	-2e-06 +- 1
d^T_{21120}	0.0(1.0)	-1e-05 +- 1
d^T_{21121}	0.0(1.0)	1e-06 +- 1
d^T_{21122}	0.0(1.0)	-1e-07 +- 1
$c^{\mathrm{sea},T}_{s,0}$	0.00(50)	0.08(50)
$c^{\mathrm{sea},T}_{s,1}$	0.00(50)	0.01(50)
$c^{\mathrm{sea},T}_{s,2}$	0.00(50)	0.0005(5000)
$c^{\mathrm{sea},T}_{l,0}$	0.00(50)	0.44(36)
$c^{\mathrm{sea},T}_{l,1}$	0.00(50)	0.02(50)
$c^{\mathrm{sea},T}_{l,2}$	0.00(50)	0.001(500)
$c^{\mathrm{sea},T}_{c,0}$	0.00(10)	0.021(97)
$c^{\mathrm{sea},T}_{c,1}$	0.00(10)	0.003(100)
$c^{\mathrm{sea},T}_{c,2}$	0.00(10)	0.0002(1000)
d^+_{00000}	0.0(1.0)	0.61(15)
d^+_{00001}	0.0(1.0)	-0.72(24)
d^+_{00002}	0.0(1.0)	0.39(65)
d^+_{00010}	0.0(1.0)	-0.51(12)
d^+_{00011}	0.0(1.0)	-0.36(44)
d^+_{00012}	0.0(1.0)	0.10(96)
d^+_{00020}	0.0(1.0)	0.66(26)
d^+_{00021}	0.0(1.0)	-0.07(92)
d^+_{00022}	0.0(1.0)	0.03(99)
d^+_{00100}	0.00(30)	-0.09(24)
d^+_{00101}	0.00(30)	0.002(300)
d^+_{00102}	0.00(30)	0.0003(3000)
d^+_{00110}	0.00(30)	-0.02(30)
d^+_{00111}	0.00(30)	0.0002(3000)
d^+_{00112}	0.00(30)	0.00005(30000)

Table 7.5: Priors and their posterior values for the z expansion fit.

Parameter	Prior	Posterior
d^+_{00120}	0.00(30)	-0.001(300)
d^+_{00121}	0.00(30)	0.0001(3000)
d^+_{00122}	0.00(30)	1e-05 +- 0.3
d^+_{01000}	0.00(30)	0.063(93)
d^+_{01001}	0.00(30)	0.01(29)
d^+_{01002}	0.00(30)	-0.0007(2999)
d^+_{01010}	0.00(30)	-0.12(24)
d^+_{01011}	0.00(30)	-0.003(300)
d^+_{01012}	0.00(30)	0.0004(3000)
d^+_{01020}	0.00(30)	-0.01(29)
d^+_{01021}	0.00(30)	-0.001(300)
d^+_{01022}	0.00(30)	0.0001(3000)
d^+_{01100}	0.0(1.0)	0.06(93)
d^+_{01101}	0.0(1.0)	0.002(1.000)
d^+_{01102}	0.0(1.0)	0.0002(1.0000)
d^+_{01110}	0.0(1.0)	-0.02(1.00)
d^+_{01111}	0.0(1.0)	-8e-05 +- 1
d^+_{01112}	0.0(1.0)	3e-05 +- 1
d^+_{01120}	0.0(1.0)	-0.001(1.000)
d^+_{01121}	0.0(1.0)	-1e-05 +- 1
d^+_{01122}	0.0(1.0)	7e-06 +- 1
d^+_{10000}	0.0(1.0)	-0.29(57)
d^+_{10001}	0.0(1.0)	0.002(901)
d^+_{10002}	0.0(1.0)	0.14(98)
d^+_{10010}	0.0(1.0)	0.18(51)
d^+_{10011}	0.0(1.0)	0.02(96)
d^+_{10012}	0.0(1.0)	0.02(1.00)
d^+_{10020}	0.0(1.0)	0.38(90)
d^+_{10021}	0.0(1.0)	0.02(99)
d^+_{10022}	0.0(1.0)	0.005(1.000)
d^+_{10100}	0.00(30)	-0.03(30)
d^+_{10101}	0.00(30)	0.0009(3000)
d^+_{10102}	0.00(30)	0.00007(30000)
d^+_{10110}	0.00(30)	-0.006(300)
d^+_{10111}	0.00(30)	0.00008(30000)
d^+_{10112}	0.00(30)	1e-05 +- 0.3
d^+_{10120}	0.00(30)	-0.0003(3000)
d^+_{10121}	0.00(30)	0.00004(30000)
d^+_{10122}	0.00(30)	2e-06 +- 0.3
d^+_{11000}	0.00(30)	-0.01(29)
d^+_{11001}	0.00(30)	0.001(300)
d^+_{11002}	0.00(30)	0.0004(3000)
d^+_{11010}	0.00(30)	-0.03(30)
d^+_{11011}	0.00(30)	-0.0004(3000)
d^+_{11012}	0.00(30)	0.00008(30000)
d^+_{11020}	0.00(30)	-0.003(300)
d^+_{11021}	0.00(30)	-0.00008(30000)

Parameter	Prior	Posterior
d^+_{11022}	0.00(30)	2e-05 +- 0.3
d^+_{11100}	0.0(1.0)	0.01(99)
d^+_{11101}	0.0(1.0)	0.0007(1.0000)
d^+_{11102}	0.0(1.0)	5e-05 +- 1
d^+_{11110}	0.0(1.0)	-0.006(1.000)
d^+_{11111}	0.0(1.0)	5e-06 +- 1
d^+_{11112}	0.0(1.0)	6e-06 +- 1
d^+_{11120}	0.0(1.0)	-0.0004(1.0000)
d^+_{11121}	0.0(1.0)	4e-06 +- 1
d^+_{11122}	0.0(1.0)	1e-06 +- 1
d^+_{20000}	0.0(1.0)	0.35(87)
d^+_{20001}	0.0(1.0)	0.10(97)
d^+_{20002}	0.0(1.0)	0.03(1.00)
d^+_{20010}	0.0(1.0)	0.10(90)
d^+_{20011}	0.0(1.0)	0.02(99)
d^+_{20012}	0.0(1.0)	0.004(1.000)
d^+_{20020}	0.0(1.0)	0.14(98)
d^+_{20021}	0.0(1.0)	0.01(1.00)
d^+_{20022}	0.0(1.0)	0.0009(1.0000)
d^+_{20100}	0.00(30)	-0.009(300)
d^+_{20101}	0.00(30)	0.0003(3000)
d^+_{20102}	0.00(30)	1e-05 +- 0.3
d^+_{20110}	0.00(30)	-0.002(300)
d^+_{20111}	0.00(30)	3e-05 +- 0.3
d^+_{20112}	0.00(30)	2e-06 +- 0.3
d^+_{20120}	0.00(30)	-0.00007(30000)
d^+_{20121}	0.00(30)	1e-05 +- 0.3
d^+_{20122}	0.00(30)	3e-07 +- 0.3
d^+_{21000}	0.00(30)	-0.01(30)
d^+_{21001}	0.00(30)	0.0006(3000)
d^+_{21002}	0.00(30)	0.0001(3000)
d^+_{21010}	0.00(30)	-0.009(300)
d^+_{21011}	0.00(30)	-2e-05 +- 0.3
d^+_{21012}	0.00(30)	2e-05 +- 0.3
d^+_{21020}	0.00(30)	-0.0006(3000)
d^+_{21021}	0.00(30)	9e-06 +- 0.3
d^+_{21022}	0.00(30)	4e-06 +- 0.3
d^+_{21100}	0.0(1.0)	0.002(1.000)
d^+_{21101}	0.0(1.0)	0.0002(1.0000)
d^+_{21102}	0.0(1.0)	9e-06 +- 1
d^+_{21110}	0.0(1.0)	-0.002(1.000)
d^+_{21111}	0.0(1.0)	6e-06 +- 1
d^+_{21112}	0.0(1.0)	1e-06 +- 1
d^+_{21120}	0.0(1.0)	-0.0001 +- 1
d^+_{21121}	0.0(1.0)	2e-06 +- 1
d^+_{21122}	0.0(1.0)	2e-07 +- 1

Table 7.6: Priors and their posterior values for the z expansion fit.

Parameter	Prior	Posterior
$c_{s,0}^{\text{sea},+}$	0.00(50)	0.14(44)
$c_{s,1}^{\text{sea},+}$	0.00(50)	0.0002(4993)
$c_{s,2}^{\text{sea},+}$	0.00(50)	-4e-05 +- 0.5
$c_{l,0}^{\text{sea},+}$	0.00(50)	-0.08(12)
$c_{l,1}^{\text{sea},+}$	0.00(50)	-0.03(50)
$c_{l,2}^{\text{sea},+}$	0.00(50)	0.0004(5000)
$c_{c,0}^{\text{sea},+}$	0.00(10)	0.027(63)
$c_{c,1}^{\text{sea},+}$	0.00(10)	0.0008(993)
$c_{c,2}^{\text{sea},+}$	0.00(10)	5e-06 +- 0.1
$c_{s,0}^{\text{val},+}$	0.0(1.0)	-0.11(96)
$c_{s,1}^{\text{val},+}$	0.0(1.0)	-0.002(1.000)
$c_{s,2}^{\text{val},+}$	0.0(1.0)	-0.0002(1.0000)

narrow priors unless the width favoured by log(GBF), is less than half of what we are using, which in this case leaves most priors at 0(1). The effect of broadening and narrowing priors on the final form factors will be discussed in the next section.

All but one of our priors and posteriors (a^0_{10020} is slightly over 1σ from its prior) agree to within 1 standard deviation, indicating that our priors are indeed conservative. We do not include other parameters which do technically appear as priors in the fit, but have very well informed values. For example the physical meson masses, such as M_B^{phys}. Because such priors are known to a very high degree of accuracy, the fit tends not to move them at - all such posteriors agree with their priors to within 1σ, and most are not changed from their priors by the fit.

Our preferred fit has a χ^2/d.o.f. of 0.193, for 333 degrees of freedom, corresponding to a Q value of 1.00. log(GBF) = 791.36. The continuum form factors resulting from this fit, as well as tests of the fit stability against reasonable variations in the fit form and parameters will be given below.

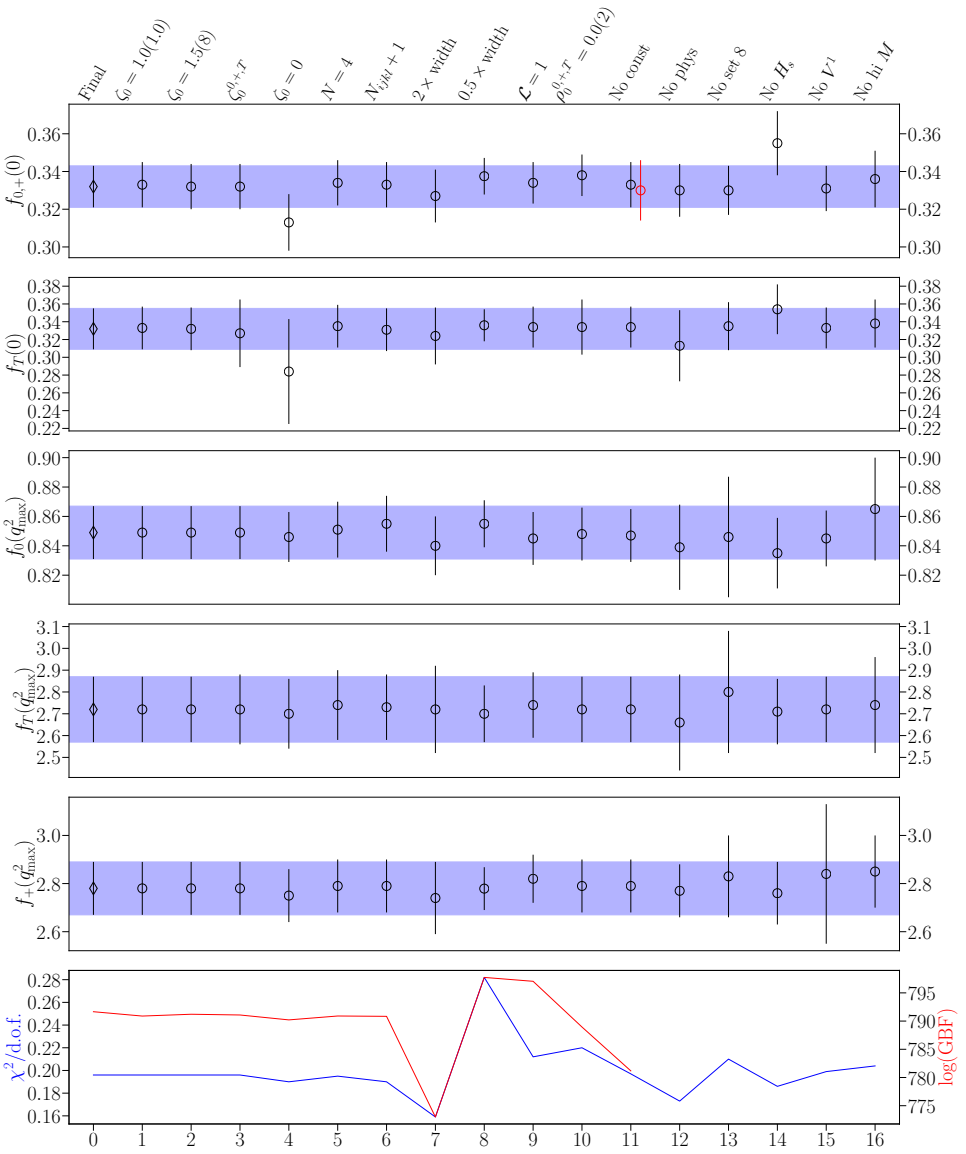

Figure 7.6: Stability tests for the z expansion fit evaluated at the physical B mass. Test 0 is the final result, 1 and 2 take different priors for ζ_0, test 3 allows ζ_0 to vary between the form factors and test 4 drops the term containing ζ entirely. Test 5 increases N by 1 (to 4) and test 6 does the same for each component of N_{ijkl}. Test 7 doubles the width of ζ_n and all d and ρ priors, and 8 halves them. Test 9 removes the chiral logarithm term by setting $\mathcal{L} = 1$, and 10 tightens the prior on the ρ coefficients considerably (from 0(1)). Test 11 removes the $f_0(0) = f_+(0)$ constraint; here the black point is $f_0(0)$ and the red is $f_+(0)$. Tests 12, 13 and 14 remove all the lattices with physical light masses, all of set 8 data, and results with $m_l = m_s$ respectively. Test 15 removes the spatial vector data, and 16 removes the largest mass from all ensembles with multiple masses. The χ^2 per degree of freedom and log(GBF) value for each test are shown in the bottom pane in blue and red respectively. For the latter tests, data is removed from the fit, resulting in a log(GBF) which is not comparable with others and so not displayed. As discussed, χ^2 values are only meaningful relatively. χ^2 values for tests 7 and 8, which change widths on many priors, are not directly comparable with other tests.

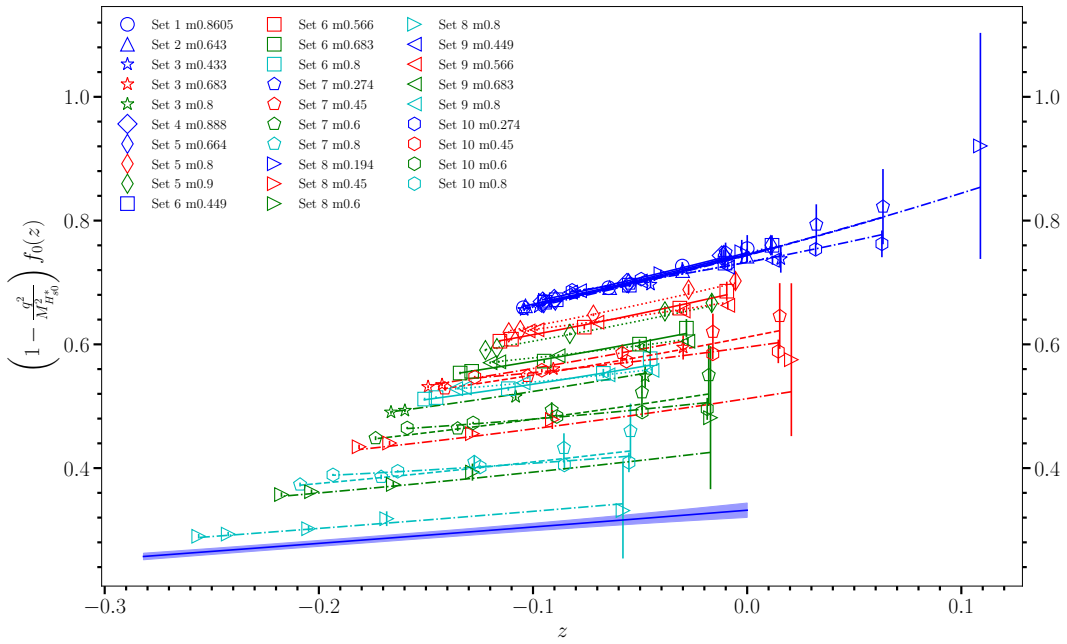

Figure 7.7: $\left(1 - \frac{q^2}{M^2_{H^*_{s0}}}\right) f_0(z)$ data points and final result at the physical point (blue band). Data points are labelled by heavy quark mass, where e.g. m0.8 indicates $am_h = 0.8$ on that ensemble. Lines between data points of a given heavy mass are the result of the fit evaluated on this ensemble and mass with all lattice artefacts present. Sets 9 and 10 are the $H_s \to \eta_s$ data from sets 1 and 2 in [2], which were fitted simultaneously with sets 6 and 7 respectively.

7.3 Form Factor results

7.3.1 $B \to K$ form factors

The stability of our z expansion against a variety of tests is shown in Figure 7.6. These tests are described in detail in the caption. We see that the fit is stable against all tests. Of particular interest are tests 1-3, which show that the fit is very stable to changes in ζ_0, the largest change being in test 3, where we give each form factor an independent ζ_0. This removes the tie between $f_T(0)$ and $f_{0/+}(0)$ and increases uncertainty, though still less than in test 4, which removes the ζ_0 term completely. Indeed, test 4 highlights the fact that it is principally f_T which benefits from the extra prior information contained within ζ_0. Figures 7.7, 7.8 and 7.9 show the form factors

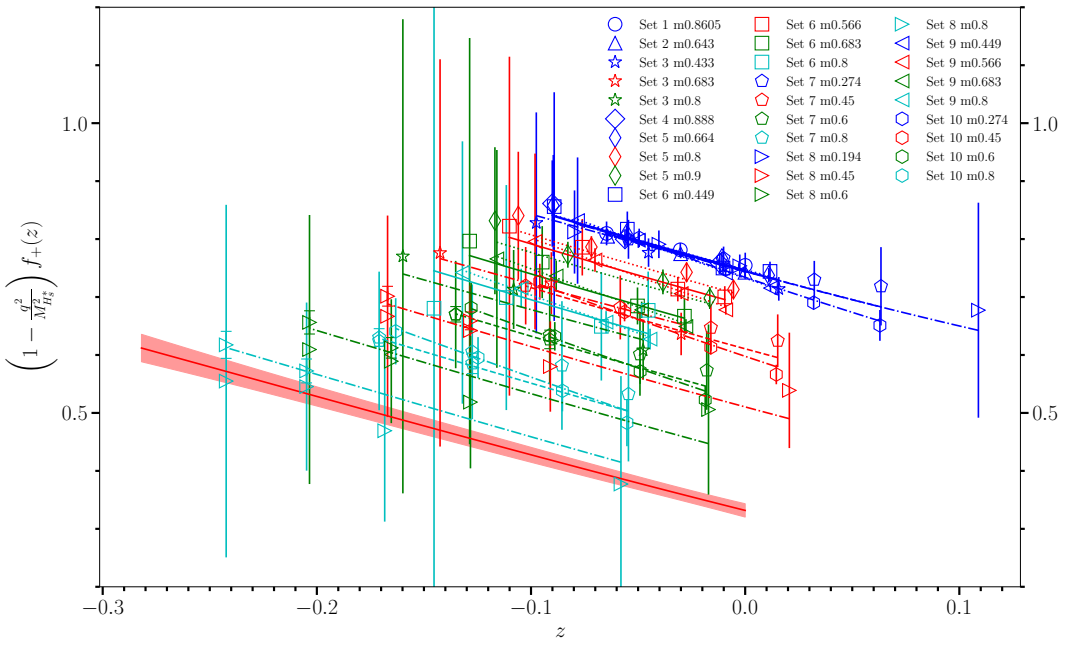

Figure 7.8: $\left(1-\frac{q^2}{M^2_{H_s^*}}\right) f_+(z)$ data points and final result at the physical point (red band). Data points are labelled by heavy quark mass, where e.g. m0.8 indicates $am_h = 0.8$ on that ensemble. Lines between data points of a given heavy mass are the result of the fit evaluated on this ensemble and mass with all lattice artefacts present. Sets 9 and 10 are the $H_s \to \eta_s$ data from sets 1 and 2 in [2], which were fitted simultaneously with sets 6 and 7 respectively. At large $|z|$ (large q^2), data obtained from both temporal and spatial components of V^μ are shown, the latter with end caps specifying the associated uncertainty. As discussed in Section 7.2.3, errors for f_+ at large q^2 are significantly smaller when obtained from spatial vector components.

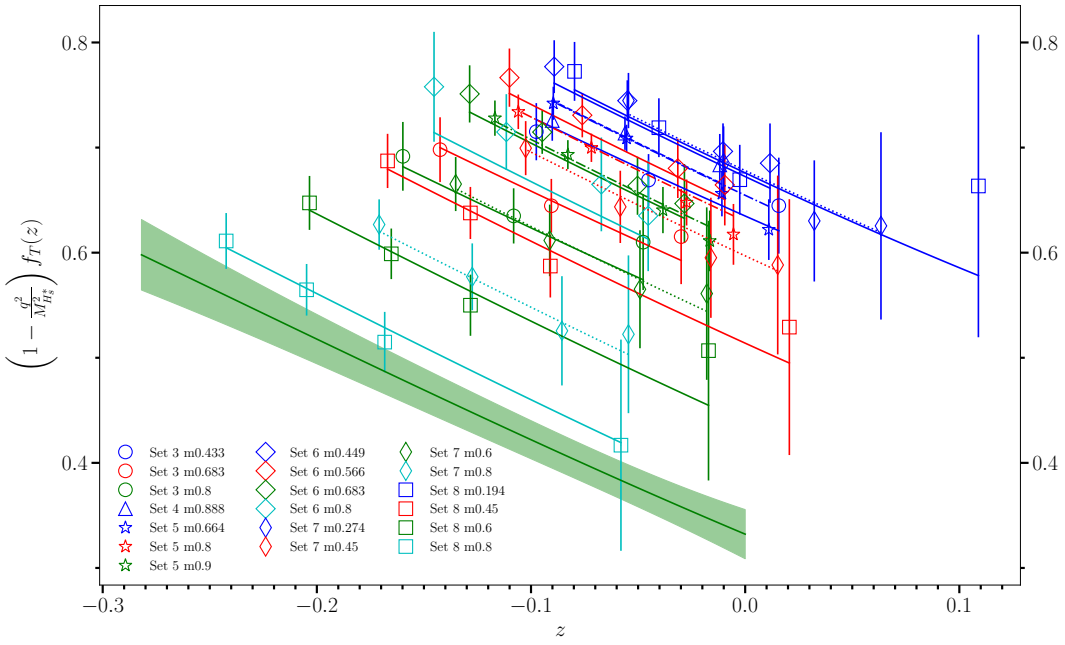

Figure 7.9: $\left(1-\frac{q^2}{M_{H_s^*}^2}\right) f_T(z)$ data points and final result at the physical point (green band). Data points are labelled by heavy quark mass, where e.g. m0.8 indicates $am_h = 0.8$ on that ensemble. Lines between data points of a given heavy mass are the result of the fit evaluated on this ensemble and mass with all lattice artefacts present.

Table 7.7: Values of fit coefficients $a_n^{0,+,T}$, pole masses, and the $\mathcal{L}$ term with correlation matrix below, evaluated at the physical point and the B mass. Note that $a_0^+ = a_0^0$. Masses are in GeV. The pole masses and $\mathcal{L}$ are very slightly correlated due to the way the fit function is constructed. These correlations are too small to have any meaningful effect on the fit, but we include them for completeness.

$a_0^{0/+}$	a_1^0	a_2^0	a_1^+	a_2^+	a_0^T	a_1^T	a_2^T	$M_{B_{s0}^*}^{\rm phys}$	$M_{B_s^*}^{\rm phys}$	$\mathcal{L}$
0.2545(90)	0.210(76)	0.02(17)	-0.71(14)	0.32(59)	0.255(18)	-0.66(23)	0.36(84)	5.729495(85)	5.4158(15)	1.304(10)
1.00000	0.80619	0.56441	0.30543	0.04776	0.42939	0.19136	0.06240	-0.00032	-0.00197	-0.19815
	1.00000	0.91180	0.35256	0.06186	0.31091	0.16899	0.05677	0.00006	-0.00250	0.02839
		1.00000	0.28531	0.08655	0.18297	0.09938	0.04827	0.00005	-0.00181	0.03245
			1.00000	0.84649	0.06813	0.09633	0.05829	0.00074	-0.01316	0.09126
				1.00000	-0.02470	0.02366	0.04442	-0.00054	0.00963	0.00353
					1.00000	0.59841	0.32316	-0.00030	0.00167	-0.11487
						1.00000	0.85349	0.00032	-0.00574	0.04788
							1.00000	-0.00046	0.00825	0.00184
								1.00000	0.00003	-0.00003
									1.00000	0.00052
										1.00000

in z space, with the poles removed, along with the data on all ensembles[6]. As was observed in previous chapters, the z expansion here is almost linear, justifying our $N = 3$ choice. In Figure 7.8, the benefit of the spatial vector data points (which have error bar caps to distinguish them) is extremely apparent. The contributions to the total uncertainty in each of the $B \to K$ form factors, across the full q^2 range, is shown in Figure 7.10. One of the important findings from the $B_s \to \eta_s$ calculation, which provided reassurance that a heavy-HISQ $B \to K$ calculation was viable, was that the uncertainty was statistics dominated (see Figure 6.4). Here we observe a similar pattern, with statistics the largest contribution, followed by the heavy mass extrapolation. This suggests that our heavy-HISQ method is not a limit to the precision at this level of statistics. We also note that discretisation effects are tiny. The $B \to K$ form factors are shown together in Figure 7.11. The similarity of f_T and f_+ is striking.

We provide values for the continuum z expansion coefficients, pole masses, and $\mathcal{L}$, along with correlations, in Table 7.7. Using these, and Equation (7.2), one can reconstruct our form factors.

7.3.2 Form factors and uncertainties across the M_H range

Because of the sliding heavy mass in our heavy-HISQ approach, we are able to look at the form factors across the whole heavy mass range from $B \to K$ down to $D \to K$. Whilst the scalar and vector $D \to K$ form factors we produce here are highly correlated to those in Chapter 5, sharing

[6]Note that the $B_s \to \eta_s$ data on sets 6 and 7 is labelled sets 9 and 10.

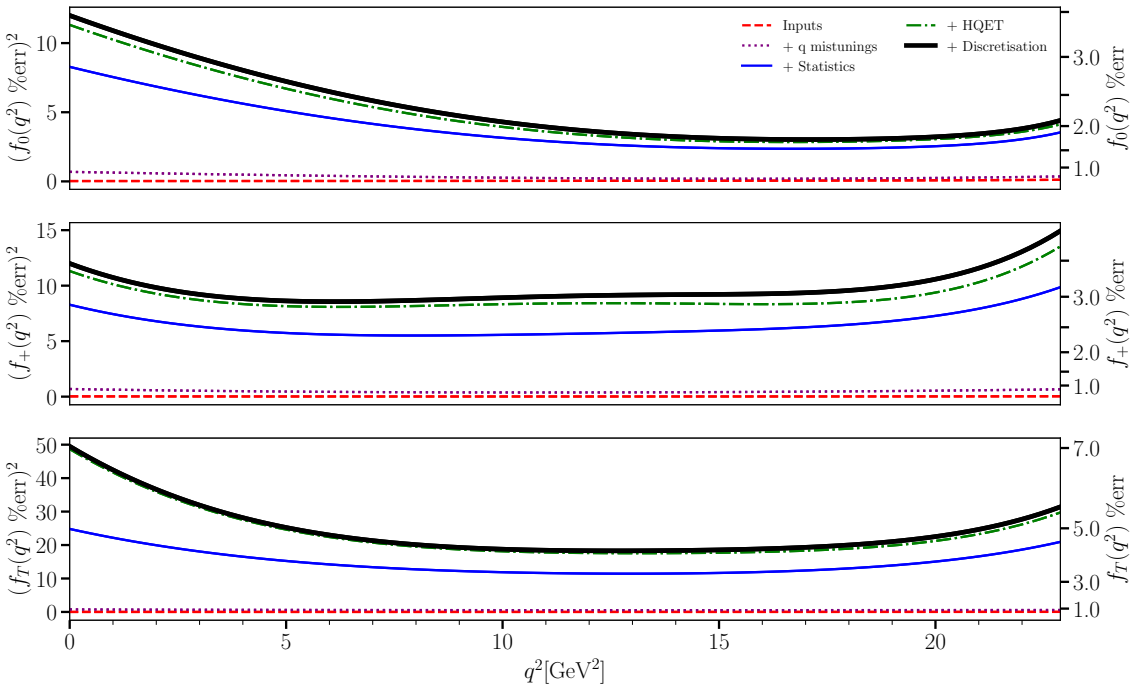

Figure 7.10: The contributions to the total percentage error (black line) of $B \to K$ form factors $f_0(q^2)$ (top) and $f_+(q^2)$ (middle) and $f_T(q^2)$ (bottom) from different sources, shown as an accumulating error. The red dashed line ('inputs') includes values for parameters, such as masses, taken from the PDG [29] and used in the fit as described above. The purple dotted line ('q mistunings') adds, negligibly, to the inputs the error contribution from the quark mistunings associated with c fit parameters and errors from the light quark chiral extrapolation, whilst the solid blue line ('statistics') further adds the error from our correlator fits. The green dot-dash line ('HQET') includes the contribution from the expansion in the heavy quark mass, and, finally, the thick black line ('Discretisation'), the total error on the form factor, also includes the discretisation errors. In the case of the tensor form factor, the difference here is so small as to obscure the HQET line. The percentage variance adds linearly and the scale for this is given on the left hand axis. The percentage standard deviation, the square root of this, can be read from the scale on the right-hand side.

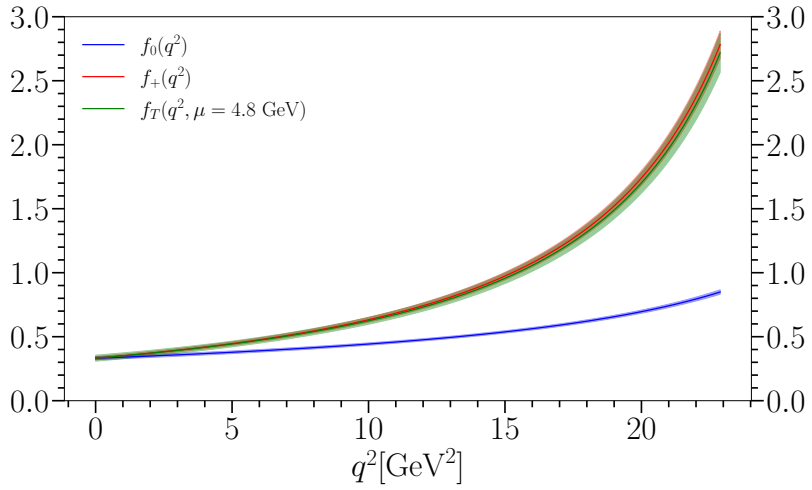

Figure 7.11: Final $B \to K$ form factor results at the physical point across the full q^2 range.

all the $m_h = m_c$ data, the tensor form factor is presented for the first time here.

Running $f_T(q^2, \mu)$

One subtlety in the straightforward extrapolation in M_H is the scale dependence of $f_T(q^2, \mu)$. In our calculation of the tensor form factor, we used $Z_T(\mu = 4.8\ \text{GeV})$, calculated in [99] in the RI-SMOM scheme [209] and matched to the $\overline{\text{MS}}$ scheme [210, 211] through $\mathcal{O}(\alpha_s^3)$ in the continuum, with uncertainties from $\mathcal{O}(\alpha_s^4)$ accounted for via a floating fit parameter (see Equation (16) of [99]). This process involves the removal of contamination from nonperturbative condensates, which would otherwise give a systematic error of 1.5%.

For our calculation, the scale μ is taken to be approximately equal to the b quark pole mass. Whilst this is appropriate for the $B \to K$ results, we use a smaller scale, $\mu = 2$ GeV, for $D \to K$ for consistency with previous results. In order to produce results at arbitrary M_H, we use a linear interpolation of μ between these two values,

$$\mu(M_H)[\text{GeV}] = 2 + \frac{2.8}{M_B^{\text{phys}} - M_D^{\text{phys}}}(M_H - M_D^{\text{phys}}). \tag{7.9}$$

Following [99], we then run from $\mu(M_B) = 4.8$ GeV to our desired μ scale. The maximal extent of this running is down to 2 GeV (i.e. for $M_H = M_D$), and this results in a factor of 1.0773(17) multiplying $f_T(q^2, 4.8\ \text{GeV})$.

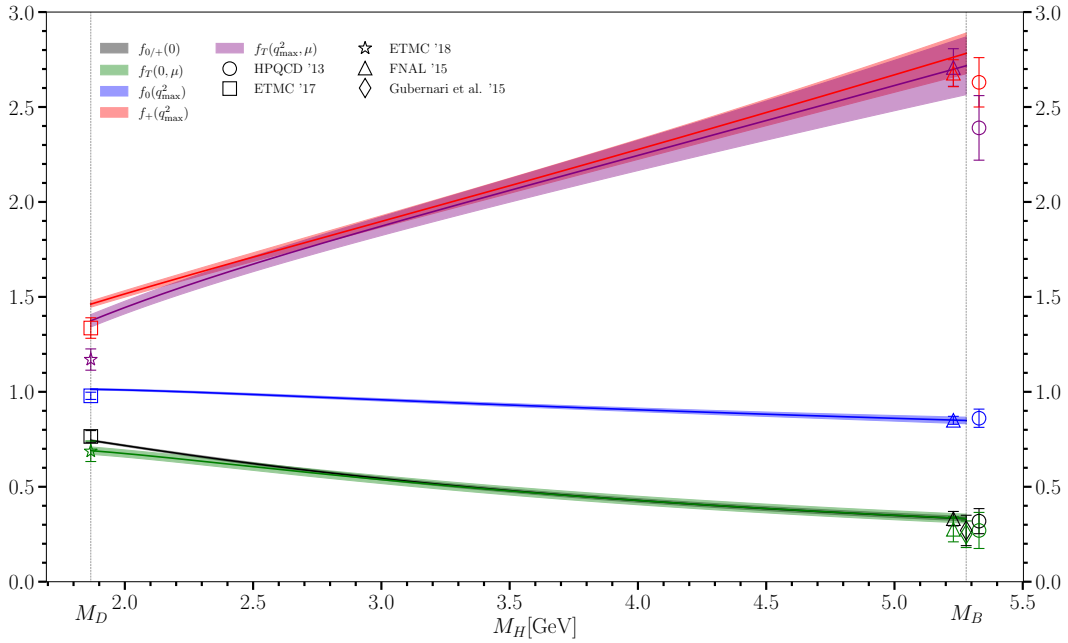

Figure 7.12: The form factors at $q^2_{\rm max}$ and $q^2 = 0$ evaluated across the range of physical heavy masses from the D to the B. Other lattice studies [140, 173, 182, 212] of both $D \to K$ and $B \to K$ are shown for comparison. We also include some $B \to K$ results at $q^2 = 0$ from Gubernari et al. [213], a calculation using light cone sum rules. We do not include Chapter 5's $D \to K$ results that share data with our calculation here. At the B end, data points are offset from M_B for clarity. Note that we have run Z_T to scale μ in this plot, where μ is defined linearly between 2 GeV and $m_b = 4.8$ GeV, according to Equation (7.9).

Comparisons

Figure 7.12 plots the forms factors at extremal q^2 values against M_H, giving a smooth transition between $D \to K$ and $B \to K$ results. We see that our form factors are in good agreement with previous work, except for $f_T^{D\to K}(q^2_{\rm max})$, which is addressed below. In most cases, our result represents a reduction in uncertainty over previous work, with the exception being $f_+^{B\to K}(q^2_{\rm max})$ and $f_T^{B\to K}(q^2_{\rm max})$, where our results are slightly less precise than those of [182]. Table 7.8 gives our $B \to K$ and $D \to K$ form factor results at extremal q^2 values. Also included are the $D \to K$ results described in Chapter 5. As discussed, these results share data and so are not wholly independent. However, as should be clear from this Chapter and Chapter 5, the correlator fits and z expansions used in each case are very different. As such, the agreement between the two

Table 7.8: Form factor results at the q^2 extremes. As described in the text, the $f_0^{D\to K}$ and $f_+^{D\to K}$ share data with the results in [1] (included for comparison) so should not be viewed as an independent calculation.

	$q^2 = 0$	$q^2 = q^2_{\rm max}$
	This work	
$f_0^{B\to K}(q^2)$	0.332(12)	0.849(17)
$f_+^{B\to K}(q^2)$	0.332(12)	2.78(11)
$f_T^{B\to K}(q^2, \mu = 4.8\text{ GeV})$	0.332(24)	2.72(15)
$f_0^{D\to K}(q^2)$	0.7441(40)	1.0136(36)
$f_+^{D\to K}(q^2)$	0.7441(40)	1.462(16)
$f_T^{D\to K}(q^2, \mu = 2\text{ GeV})$	0.690(20)	1.374(33)
	c.f. $D \to K$ [1]	
$f_0^{D\to K}(q^2)$	0.7380(44)	1.0158(41)
$f_+^{D\to K}(q^2)$	0.7380(44)	1.465(20)

Table 7.9: $B_s \to \eta_s$ form factor results at the q^2 extremes. As described in the text, the f_0 and f_+ share data with the results in [2] (included for comparison) so should not be viewed as an independent calculation.

	$q^2 = 0$	$q^2 = q^2_{\rm max}$
	This work	
$f_0^{B_s\to\eta_s}(q^2)$	0.3191(85)	0.819(17)
$f_+^{B_s\to\eta_s}(q^2)$	0.3191(85)	2.45(19)
$f_T^{B_s\to\eta_s}(q^2, \mu = 4.8\text{ GeV})$	0.370(78)	2.32(56)
	c.f. $B_s \to \eta_s$ [2]	
$f_0^{B_s\to\eta_s}(q^2)$	0.296(25)	0.808(15)
$f_+^{B_s\to\eta_s}(q^2)$	0.296(25)	2.58(208)
	c.f. $B_s \to \eta_s$ [171]	
$f_0^{B_s\to\eta_s}(q^2)$	-	0.811(17)

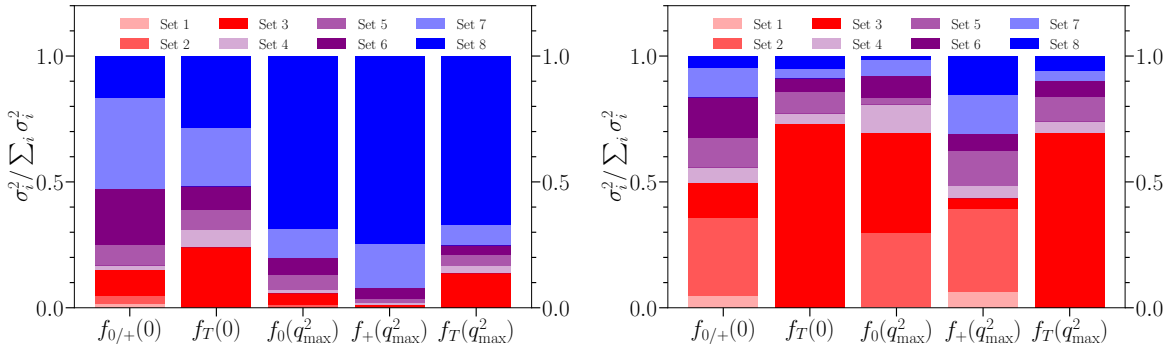

Figure 7.13: Breakdown of the contributions to the statistical uncertainty of the form factors at their extremes from data on each ensemble. Uncertainty from each ensemble σ_i is added in quadrature, normalised by the total uncertainty squared $\sum_i \sigma_i^2$. Sets 6 and 7 include contributions from $H_s \to \eta_s$ data. Left: evaluated at $M_H = M_B$ ($B \to K$). Right: evaluated at $M_H = M_D$ ($D \to K$).

results is a strong affirmation of both calculations.

Another interesting test to consider is to evaluate our $B \to K$ results at a spectator quark mass $m_l = m_s$. This is made possible by the chiral extrapolation (Equation (7.5)), which enabled us to include the $H_s \to \eta_s$ data in the first place (Section 7.2.4). By setting $m_l = m_s$ and adjusting the physical meson masses appropriately, we can produce $B_s \to \eta_s$ results to compare with Chapter 6[7]. This is what is given in Table 7.9. Again, there is some data overlap between these results, however, they differ markedly in their correlator and z expansion fits. Furthermore, the continuum data point (from [171]) included in [2] (Chapter 6) is not included here. Without this data point, our result agrees well with that in [2], and indeed with [171], which is also given in Table 7.9. This is another strong affirmation of our results.

Whilst Figure 7.10 tells us that the overall uncertainty is statistics dominated, we can go further, analysing the contributions to this statistical uncertainty from each of our eight ensembles. This breakdown is shown in Figure 7.13, which shows the contributions from each ensemble to the form factors at extremal q^2 values for both $B \to K$ (left) and $D \to K$ (right). We can see a clear difference between the two. Whilst the uncertainty in the $B \to K$ case is dominated by the finest ensembles, which reach masses closest to the physical $m_h = m_b$, the $D \to K$ end has $m_h = m_c$ masses available on all ensembles, and so the three ensembles with physical light masses (sets

[7] We also give the tensor form factor, which was not calculated there.

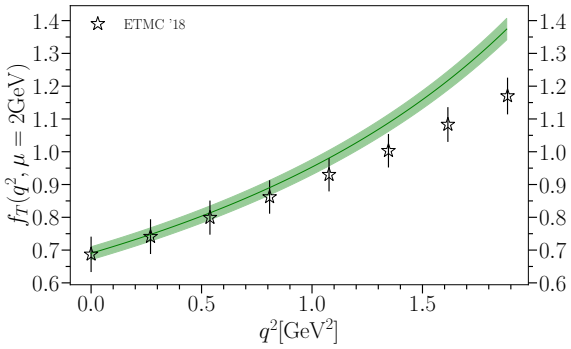

Figure 7.14: The green band gives our $D \to K$ tensor form factor at $\mu = 2$ GeV, across the physical q^2 range. Results from [212] are included for comparison.

1, 2 and 3) contribute most to the uncertainty. If we compare this with the equivalent plot for the $D \to K$ calculation in Chapter 5 (Figure 5.8), we note that in this case the uncertainty is much more evenly spread across the ensembles, owing to the more constrained chiral logarithm adopted there.

$D \to K$ tensor form factor

Our result for the $D \to K$ tensor form factor, $f_T(q^2, \mu = 2 \text{ GeV})$ which was not calculated in Chapter 5 is presented in Figure 7.14 (and numerical results are given in Table 7.8). Results from [212] are included for comparison, and we find our form factor to be roughly twice as precise, and in good agreement at $q^2 = 0$, but in considerable tension at larger q^2 values (up to 3.1σ at q^2_{max}). Additionally, we report the ratio $f_T^{D\to K}(0, \mu = 2 \text{ GeV})/f_+^{D\to K}(0) = 0.928(27)$, which agrees with the 0.898(50) given in [212].

7.3.3 Spectator quark dependence of form factors

We can compare our form factors with others sharing the same $b \to s$ current, but with different spectator quarks. In the case of $B \to K$, we compare our form factors with our own $B_s \to \eta_s$ form factors from Chapter 6 ([2]), as well as $B_c \to D_s$ [214]. This comparison is shown for f_0 and f_+, with and without the pole factors, in Figure 7.15. We see very mild spectator quark dependence for the light and strange quarks, at most a deviation of $\approx 1\sigma$, which is roughly consistent with the modest effect of setting $\mathcal{L} = 1$ in Figure 7.6. The transition to a heavy c spectator leads to a much larger change. The heavier spectator gives a smaller form factor at

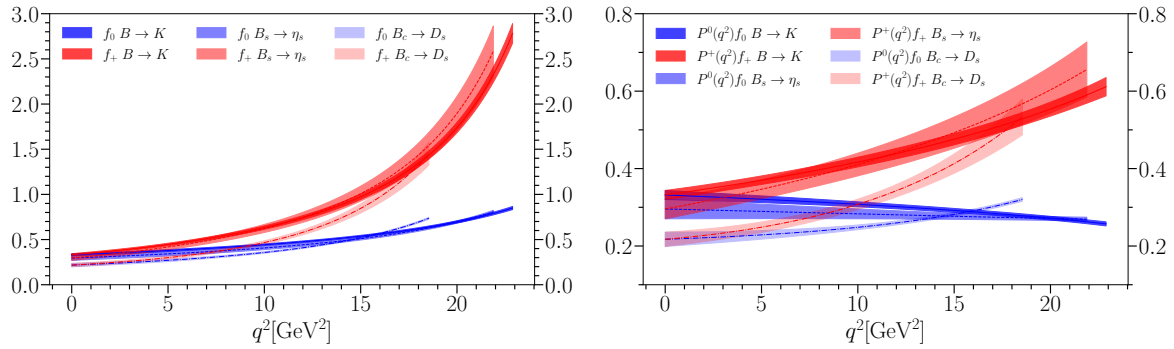

Figure 7.15: Comparison of our $B \to K$ scalar and vector form factors with those of $B_s \to \eta_s$ [2] and $B_c \to D_s$ [214] to show the impact of changing the spectator quark mass. In the right pane, we have multiplied the form factors by their common pole factors for clarity. We take $P^0(q^2) = 1 - \frac{q^2}{M^2_{B^*_{s0}}}$, $P^+(q^2) = 1 - \frac{q^2}{M^2_{B^*_s}}$, using PDG masses [29].

$q^2 = 0$ that rises more steeply to q^2_{max} (which has a considerably smaller value). The behaviour of f_T shown on the left of Figure 7.16 is similar, but with a smaller shift at $q^2 = 0$. The right hand side of Figure 7.16 shows an equivalent plot for the scalar and vector $D \to K$ form factors, compared with the form factors for $B_c \to B_s$ [98], as well as our $D_s \to \eta_s$ form factors from Chapter 6. As for $B \to K$, we see a very gentle dependence when we change the spectator from light to strange. This agrees with the conclusions of [138] but is much more compelling here because of the high precision of both sets of form factors. The biggest deviation is for f_0, at the maximum q^2 for $D_s \to \eta_s$, where $D_s \to \eta_s$ is larger than $D \to K$ by $\approx 2\sigma$ (or about 2%). When we take the spectator quark mass all the way up to b, however, we unsurprisingly see a much larger change. The form factors for $B_c \to B_s$ have smaller values at $q^2 = 0$ (but only by $\sim$20%) and rise much more steeply with q^2 than is the case with lighter spectator quarks. This trend is exactly the same, but magnified by the larger quark mass change, as that seen for the $B \to K$ case as we change from a light to a charm spectator (compare Figure 7.15).

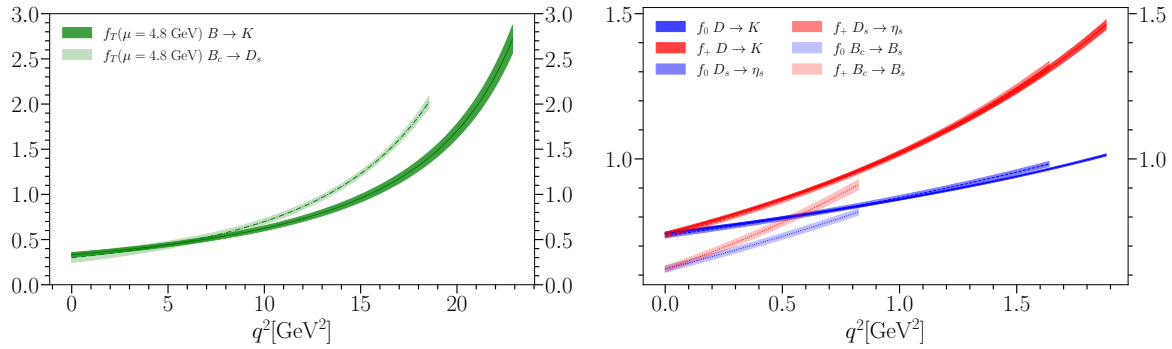

Figure 7.16: Left: comparison of our $B \to K$ tensor form factor (at $\mu = 4.8$ GeV) with those of $B_c \to D_s$ [214] to show the impact of changing the spectator quark mass. Right: a similar comparison of our $D \to K$ scalar and vector form factors with those for $D_s \to \eta_s$ [2] and $B_c \to B_s$ [98].

7.4 Key phenomenology

The form factors calculated above, and presented in [3] can be used to produce SM values for experimentally measurable quantities pertaining to $B \to K\bar{\ell}\ell$ decays, most obviously the branching fraction $\mathcal{B}$. In [4], our results are given in exhaustive detail, and numerical values are tabulated extensively. Here, we will report the key results from that paper, relevant to the search for new physics, and tabulate numerical results, but we direct the reader there for further detail.

7.4.1 $B \to K\ell^+\ell^-$

Differential decay rate

The SM differential decay rate for $B \to K\ell^+\ell^-$ for lepton ℓ is constructed as follows, where we follow the notation in [109, 215]:

$$\frac{d\Gamma_\ell}{dq^2} = 2a_\ell + \frac{2}{3}c_\ell, \tag{7.10}$$

where a_ℓ and c_ℓ are given by

$$\begin{aligned} a_\ell &= \mathcal{C}\Big[q^2|F_P|^2 + \frac{\lambda(q, M_B, M_K)}{4}(|F_A|^2 + |F_V|^2) \\ &\quad + 4m_\ell^2 M_B^2 |F_A|^2 + 2m_\ell(M_B^2 - M_K^2 + q^2)\mathrm{Re}(F_P F_A^*)\Big], \\ c_\ell &= -\frac{\mathcal{C}\lambda(q, M_B, M_K)\beta_\ell^2}{4}(|F_A|^2 + |F_V|^2), \end{aligned} \tag{7.11}$$

with

$$\begin{aligned} \mathcal{C} &= \frac{(\eta_{\mathrm{EW}} G_F)^2 \alpha_{EW}^2 |V_{tb}V_{ts}^*|^2}{2^9\pi^5 M_B^3}\beta_\ell\sqrt{\lambda(q, M_B, M_K)}, \\ \beta_\ell &= \sqrt{1 - 4m_\ell^2/q^2}, \\ \lambda(a, b, c) &= a^4 + b^4 + c^4 - 2(a^2b^2 + a^2c^2 + b^2c^2). \end{aligned} \tag{7.12}$$

$F_{P,V,A}$ are constructed from the scalar, vector and tensor form factors f_0, f_+ and f_T respectively, by

$$\begin{aligned} F_P &= -m_\ell C_{10}\Big[f_+ - \frac{M_B^2 - M_K^2}{q^2}(f_0 - f_+)\Big], \\ F_V &= C_9^{\mathrm{eff},1} f_+ + \frac{2m_b^{\overline{\mathrm{MS}}}(\mu_b)}{M_B + M_K} C_7^{\mathrm{eff},1} f_T(\mu_b), \\ F_A &= C_{10} f_+, \end{aligned} \tag{7.13}$$

where the Wilson coefficient $C_9^{\mathrm{eff},1} = C_9^{\mathrm{eff},0} + \Delta C_9^{\mathrm{eff}} + \delta C_9^{\mathrm{eff}}$ includes nonfactorisable corrections in $\Delta C_9^{\mathrm{eff}}$, as well as $\mathcal{O}(\alpha_s)$ and more heavily suppressed corrections in $\delta C_9^{\mathrm{eff}}$. Similarly, $C_7^{\mathrm{eff},1} = C_7^{\mathrm{eff},0} + \delta C_7^{\mathrm{eff}}$ contains $\mathcal{O}(\alpha_s)$ corrections in $\delta C_7^{\mathrm{eff}}$. The size and effect of these corrections has not been addressed in previous work, and we go into further detail below.

$C_9^{\text{eff},0} = C_9 + Y(q^2)$ is a function of q^2 through

$$\begin{aligned} Y(q^2) = &\frac{4}{3}C_3 + \frac{64}{9}C_5 + \frac{64}{27}C_6 \\ &- \frac{1}{2}h(q^2, 0)\left(C_3 + \frac{4}{3}C_4 + 16C_5 + \frac{64}{3}C_6\right) \\ &+ h(q^2, m_c)\left(\frac{4}{3}C_1 + C_2 + 6C_3 + 60C_5\right) \\ &- \frac{1}{2}h(q^2, m_b)\left(7C_3 + \frac{4}{3}C_4 + 76C_5 + \frac{64}{3}C_6\right), \end{aligned} \tag{7.14}$$

where

$$\begin{aligned} h(q^2, m) = &-\frac{4}{9}\left(\ln\frac{m^2}{\mu^2} - \frac{2}{3} - x\right) - \frac{4}{9}(2+x) \\ &\times \begin{cases} \sqrt{x-1}\arctan\frac{1}{\sqrt{x-1}} & , x > 1 \\ \sqrt{1-x}\left(\ln\frac{1+\sqrt{1-x}}{\sqrt{x}} - \frac{i\pi}{2}\right), & x \leq 1, \end{cases} \end{aligned} \tag{7.15}$$

with $x = 4m^2/q^2$.

The Wilson coefficients used here are given in Table 7.10 and discussed in Section 2.3.2. The numerical values for other inputs used here are also given in Table 7.10. Below we discuss a couple of subtleties which have not been well described previously.

Additionally, it is worth noting that the expressions above do not describe the somewhat baroque behaviour of the differential decay rate in certain q^2 regions. $u\bar{u}$ resonances appear below $q^2 = 1$ GeV^2, and $c\bar{c}$ resonances at higher values of q^2. For this reason, experimentalists typically designate two vetoed regions, which we shall take as $8.68 \leq q^2/\text{GeV}^2 \leq 10.11$ (labelled J/Ψ) and $12.86 \leq q^2/\text{GeV}^2 \leq 14.18$ (labelled $\Psi(2S)$) [220]. In fact, where relevant, we linearly interpolate our differential decay rate across the whole of this region $8.68 \leq q^2/\text{GeV}^2 \leq 14.18$. Further more, when quoting integrated results for the decay rate or branching fraction, it's conventional to focus on two reliable regions of q^2, far from these resonances: $1.1 \leq q^2/\text{GeV}^2 \leq 6$ and $15 \leq q^2/\text{GeV}^2 \leq 22$. This is where we shall give our main results.

Parameter	Value	Reference
$\eta_{\rm EW} G_F$	$1.1745(23) \times 10^{-5}\,\text{GeV}^{-2}$	[216], Eq. (5.20)
$m_c^{\overline{\rm MS}}(m_c^{\overline{\rm MS}})$	1.2719(78) GeV	See caption
$m_b^{\overline{\rm MS}}(\mu_b)$	4.209(21) GeV	[217]
m_c	1.68(20) GeV	-
m_b	4.87(20) GeV	-
f_{K^+}	0.1557(3) GeV	[75, 132, 135, 218]
f_{B^+}	0.1894(14) GeV	[79]
τ_{B^0}	1.519(4) ps	[129]
$\tau_{B^\pm}$	1.638(4) ps	[129]
$1/\alpha_{\rm EW}(M_Z)$	127.952(9)	[216]
$\sin^2\theta_W$	0.23124(4)	[216]
$\lvert V_{tb}V_{ts}^*\rvert$	0.04185(93)	[90]
$C_1(\mu_b)$	-0.294(9)	[40]
$C_2(\mu_b)$	1.017(1)	[40]
$C_3(\mu_b)$	-0.0059(2)	[40]
$C_4(\mu_b)$	-0.087(1)	[40]
$C_5(\mu_b)$	0.0004	[40]
$C_6(\mu_b)$	0.0011(1)	[40]
$C_7^{\text{eff},0}(\mu_b)$	-0.2957(5)	[40]
$C_8^{\text{eff}}(\mu_b)$	-0.1630(6)	[40]
$C_9(\mu_b)$	4.114(14)	[40]
$C_9^{\text{eff},0}(\mu_b)$	$C_9(\mu_b) + Y(q^2)$	-
$C_{10}(\mu_b)$	-4.193(33)	[40]

Table 7.10: Input parameters used to calculate SM observables. Details of the error on the Wilson coefficients are given in [40], and these are quoted at $\mu_b = 4.2\,\text{GeV}$. For more details on Wilson coefficients, see Section 2.3.2. $m_c^{\overline{\rm MS}}(m_c^{\overline{\rm MS}})$ is obtained using $m_c^{\overline{\rm MS}}(3\text{ GeV}) = 0.9841(51)$ [219] and running the scale to its own mass. m_c and m_b are the c and b quark pole masses obtained from the masses in the $\overline{\rm MS}$ scheme at three loops (see Section 7.4.1 for details).

Calculating pole masses

The first subtlety above is the values which we take for masses m_c and m_b. These are pole masses. The three loop relation between quark masses in the $\overline{\text{MS}}$ scheme and the pole mass scheme is [221]

$$\begin{aligned}\frac{\overline{m}(m)}{m} &= 1 + A\Big(\frac{\alpha_s}{\pi}\Big) + B\Big(\frac{\alpha_s}{\pi}\Big)^2 + C\Big(\frac{\alpha_s}{\pi}\Big)^3, \\ A &= -\frac{4}{3}, \\ B &= 1.0414 N_L - 14.3323, \\ C &= -0.65269 N_L^2 + 26.9239 N_L - 198.7068,\end{aligned} \tag{7.16}$$

where $\overline{m}$ is the $\overline{\text{MS}}$ mass, m is the pole mass and N_L is the number of active light quarks. We evaluate α_s at scale m and use $N_L = 3$ for c and $N_L = 4$ for b.

Inverting (7.16) gives

$$\begin{aligned}m =& \overline{m}(m)\Big[1 - A\Big(\frac{\alpha_s}{\pi}\Big) + (A^2 - B)\Big(\frac{\alpha_s}{\pi}\Big)^2 \\ &+ (-A^3 + 2AB - C)\Big(\frac{\alpha_s}{\pi}\Big)^3\Big].\end{aligned} \tag{7.17}$$

We solve Equation (7.17) iteratively, making an initial guess for m then evaluating $\alpha_s(m)$ by running from $\alpha_s(5.0\,\text{GeV}) = 0.2128(25)$ [80] and $\overline{m}(m)$ by running from the $m_c^{\overline{\text{MS}}}(\mu = m_c^{\overline{\text{MS}}})$ or $m_b^{\overline{\text{MS}}}(\mu = m_b^{\overline{\text{MS}}})$ (values in [4]). Plugging these results into Equation (7.17) results in an updated value for m. The initial guess for m is adjusted to reduce the difference between it and the value obtained from Equation (7.17). This process is repeated until the values of m converge.

Using this method, we obtain the pole masses $m_c = 1.684(22)\,\text{GeV}$ and $m_b = 4.874(32)\,\text{GeV}$. The perturbation series in this expression suffers from the presence of a renormalon in the pole mass [222], so we take a 200 MeV uncertainty on both numbers. We note that $4m_c^2 = 11.34\,\text{GeV}^2$ falls within the vetoed resonance region.

Investigation of corrections to $C_7^{\text{eff},0}$ and $C_9^{\text{eff},0}$

The second subtlety mentioned above is that corrections need to be applied to $C_7^{\text{eff},0}$ and $C_9^{\text{eff},0}$ to obtain the values of $C_7^{\text{eff},1}$ and $C_9^{\text{eff},1}$ that enter Equation (7.13). The corrections are defined

by

$$C_7^{\text{eff},1} = C_7^{\text{eff},0} + \delta C_7^{\text{eff}}, \tag{7.18}$$
$$C_9^{\text{eff},1} = C_9^{\text{eff},0} + \Delta C_9^{\text{eff}} + \delta C_9^{\text{eff}}, \tag{7.19}$$

where $C_7^{\text{eff},0}$ and $C_9 = C_9^{\text{eff},0} - Y(q^2)$ are constants, given in [4]. The corrections are discussed in Appendix B of [181], which compiles results from [184, 185, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234]. We direct the reader there for more detail. Below we outline the form of the corrections, plot them against q^2 and discuss their sizes relative to $C_7^{\text{eff},0}$ and $C_9^{\text{eff},0}$. All numerical inputs not explicitly stated below can be found in [4].

The leading contribution to the correction δC_7^{eff} is from $\mathcal{O}(\alpha_s)$ effects,

$$\delta C_7^{\text{eff}} = -\frac{\alpha_s}{4\pi}\Big((C_1 - 6C_2)F_{1,c}^{(7)} + C_8 F_8^{(7)}\Big). \tag{7.20}$$

The expression for $F_{1,c}^{(7)}$ is lengthy and provided in the C++ header files of [224], whilst $F_8^{(7)}$ is given in Appendix B of [181]. We use $\alpha_s(4.2\,\text{GeV}) = 0.2253(28)$, which is run from $\alpha_s(5.0\,\text{GeV}) = 0.2128(25)$ [80]. The next higher order contribution, which we neglect, is suppressed by a factor of $\lambda_u^{(s)} = \frac{V_{us}^* V_{ub}}{V_{ts}^* V_{tb}} = 0.01980(62)$ and is $\mathcal{O}(\alpha_s \lambda_u^{(s)})$.

The leading order contributions to the correction δC_9^{eff} are given by

$$\begin{aligned}\delta C_9^{\text{eff}} = &-\frac{\alpha_s}{4\pi}\left(C_1 F_{1,c}^{(9)} + C_2 F_{2,c}^{(9)} + C_8 F_8^{(9)}\right) \\ &+\lambda_u^{(s)}\left(h(q^2, m_c) - h(q^2, 0)\right)\left(\frac{4}{3}C_1 + C_2\right).\end{aligned} \tag{7.21}$$

We neglect the $\mathcal{O}(\alpha_s \lambda_u^{(s)})$ term, which is even smaller (see Equation (B11) of [181] for more details). The function $h(q^2, m)$ is defined in Equation (7.15) and $F_8^{(9)}$ is given in Appendix B of [181]. Expressions for $F_{1,c}^{(9)}$ and $F_{2,c}^{(9)}$ are also provided in the C++ header files of [224]. The corrections δC_7^{eff} and δC_9^{eff} are applicable across the full q^2 range.

We plot the real and imaginary parts of δC_7^{eff} and δC_9^{eff} in Figure 7.17, showing separately the $\mathcal{O}(\alpha_s)$ and $\mathcal{O}(\lambda_u^{(s)})$ contributions to δC_9^{eff}. The correction $\text{Re}[\delta C_7^{\text{eff}}]$ is approximately 20% of $\text{Re}[C_7^{\text{eff},0}] \approx -0.3$ [40]. The magnitude of $\text{Im}[\delta C_7^{\text{eff}}]$ is small in comparison to $|C_7^{\text{eff},1}|$. The $\mathcal{O}(\lambda_u^{(s)})$ contributions to δC_9^{eff} are negligible in comparison to $C_9^{\text{eff},0}$, which is approximately 4 across all q^2.

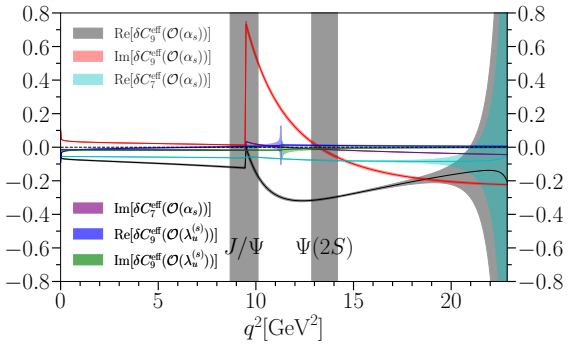

Figure 7.17: The real and imaginary parts of $\delta C_7^{\mathrm{eff}}$ and of the $\mathcal{O}(\alpha_s)$ and $\mathcal{O}(\lambda_u^{(s)})$ contributions to $\delta C_9^{\mathrm{eff}}$, as defined in Equations (7.21) and (7.20).

The $\mathcal{O}(\alpha_s)$ contributions to $\delta C_9^{\mathrm{eff}}$ peak at $q^2 \approx 10\,\mathrm{GeV}^2$, owing to the behaviour of the functions $F_{1,c}^{(7)}$, $F_{2,c}^{(7)}$, $F_{1,c}^{(9)}$ and $F_{2,c}^{(9)}$. This peak occurs within the experimentally vetoed J/Ψ resonance region and is largely contained within the region of q^2 between the J/Ψ and $\Psi(2S)$ resonances, outside of which, the contributions are modest. As a result, it has minimal impact on results in the well behaved regions of q^2 below the J/Ψ and above the $\Psi(2S)$. Uncertainty in the $\mathcal{O}(\alpha_s)$ contribution to $\mathrm{Re}[\delta C_7^{\mathrm{eff}}]$ and $\mathrm{Re}[\delta C_9^{\mathrm{eff}}]$ grows rapidly towards q^2_{max}, an effect that is suppressed in observables by the fact that the differential decay rate vanishes at q^2_{max}. This effect is not noticeable in plots of observables versus q^2 or in uncertainties of observables in the largest q^2 bins (see [4] and results below).

Non-factorisable corrections are accounted for via $\Delta C_9^{\mathrm{eff}}$. Following the notation of [181],

$$\begin{aligned}\Delta C_9^{\mathrm{eff}} &= \frac{2\pi^2 m_b f_B f_K}{3 M_B^2 f_+} \times \\ &\sum_{\pm} \int_0^\infty \frac{d\omega}{\omega} \Phi_{B,\pm}(\omega) \int_0^1 du \Phi_K(u) \left[T_{K,\pm}^{(0)} + \frac{\alpha_s}{4\pi} C_F T_{K,\pm}^{(\mathrm{nf})} \right],\end{aligned} \tag{7.22}$$

where $C_F = 4/3$, and $T_{K,+}^{(0)} = 0$. The functions $\Phi_{B,\pm}(\omega)$, Φ_K, and $T_{K,\pm}^{(0/\mathrm{nf})}$ are given in [223] and Equation (7.22) is discussed in detail in Appendix B of [181]. We evaluate the expressions for $\Phi_{B,\pm}(\omega)$, Φ_K, and $T_{K,\pm}^{(0/\mathrm{nf})}$ using the following inputs: $\omega_0^{-1} = 3(1)\,\mathrm{GeV}^{-1}$ [223], $a_1^K = 0.0453(30)$, and $a_2^K = 0.175(50)$ [235] (we take Φ_K to second order). The non-factorisable corrections are valid for small q^2, and we turn them off at $q^2 = 8.68\,\mathrm{GeV}^2$, the start of the vetoed J/Ψ resonance region. We do not calculate through this region. Instead, we linearly interpolate, beginning from

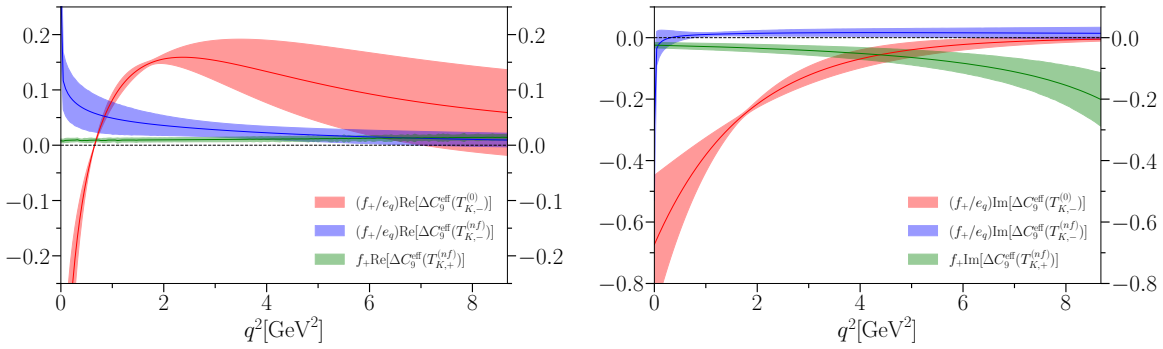

Figure 7.18: The real (left) and imaginary (right) parts of the contributions to $f_+\Delta C_9^{\text{eff}}$ from each of the three non-zero terms, $T_{K,-}^{(0)}$ and $T_{K,\pm}^{(\text{nf})}$ (Equation (7.22)). A decay channel-specific factor of $e_q \in \{2/3, -1/3\}$ is removed from $T_{K,-}^{(0)}$ and $T_{K,-}^{(\text{nf})}$.

the point where the corrections are turned off through the $\Psi(2S)$ resonance, so the differential branching fraction is a smooth function of q^2.

The contribution from ΔC_9^{eff} to observables is via the term $f_+\Delta C_9^{\text{eff},1}$ in Equation (7.13). This has the effect of cancelling the dependence of $f_+\Delta C_9^{\text{eff}}$ on the form factor f_+. The real and imaginary parts of the three nonzero contributions to $f_+\Delta C_9^{\text{eff}}$ in Equation (7.22), corresponding to $T_{K,-}^{(0)}$ and $T_{K,\pm}^{(0/\text{nf})}$, are plotted in Figure 7.18. In these plots we remove a decay channel-specific factor of the light quark charge, e_q, which is $2/3$ for $B^+ \to K^+$ and $-1/3$ for $B^0 \to K^0$. Among these terms, the $T_{K,-}^{(0)}$ contribution is dominant, especially for $q^2 \lesssim 1\,\text{GeV}^2$.

The combined effect of these terms to $f_+\Delta C_9^{\text{eff}}$ is shown in Figure 7.19, where the real and imaginary parts are plotted separately for both the B^0 and B^+ cases. Both the real and imaginary parts are smooth functions of q^2 in the region below the J/Ψ resonance where they are considered ($4m_\ell^2 \le q^2 \le 8.68\ \text{GeV}^2$), and are small for $q^2 > 1\,\text{GeV}^2$. The differential decay rate is suppressed by at least $\beta = \sqrt{1 - 4m_\ell^2/q^2}$ at low q^2 where the nonfactorisable corrections are largest. As a result, the corrections do not make a significant contribution, as can be seen in the plots of the differential branching fraction.

The overall modest contributions of the above corrections are shown for C_7^{eff} in Figure 7.19 and C_9^{eff} in Figure 7.20. In each plot, both corrected and uncorrected values are plotted. Cusps in C_7^{eff} and C_9^{eff} either occur within the vetoed J/Ψ resonance region near $q^2 \approx 10\,\text{GeV}^2$ or at $q^2 = 4m_c^2 \approx 11.3\,\text{GeV}^2$, between the J/Ψ and $\Psi(2S)$ resonances. In this region between

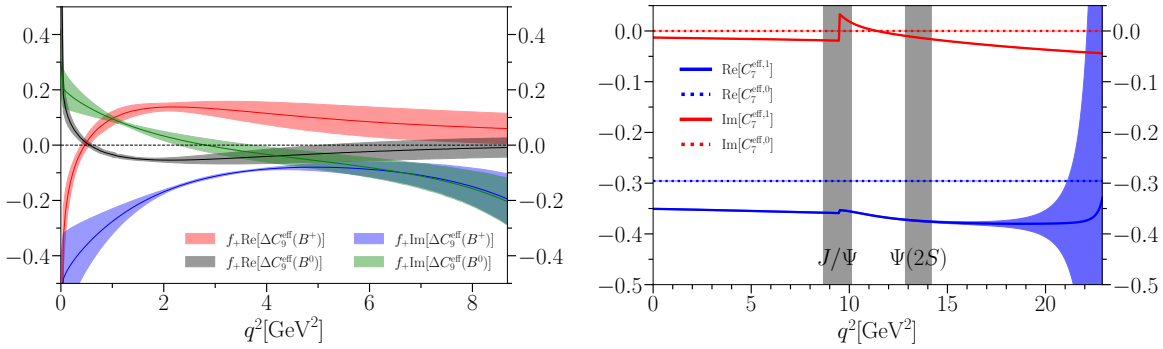

Figure 7.19: Left: the real and imaginary parts of $f_+\Delta C_9^{\mathrm{eff}}$ for both the B^0 and B^+ cases. Right: the real and imaginary parts of $C_7^{\mathrm{eff},0}$ and $C_7^{\mathrm{eff},1}$ (see Equation (7.18)), showing the combined effect of the nonfactorisable and $\mathcal{O}(\alpha_s)$ corrections. Corrected values are shown with solid lines and dark fill colour, while uncorrected values are shown with dotted lines and light fill colour.

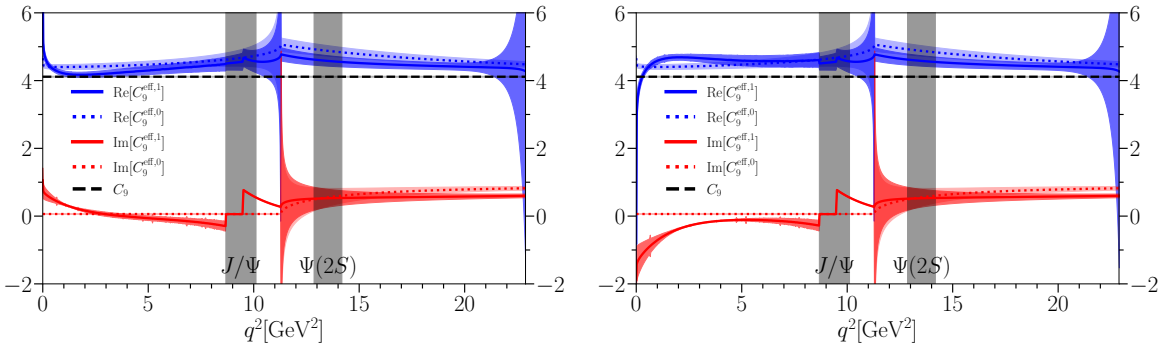

Figure 7.20: The real and imaginary parts of $C_9^{\mathrm{eff},0}$ and $C_9^{\mathrm{eff},1}$ (see Equation (7.19)), for $B^0 \to K^0\ell^+\ell^-$ (left) and $B^+ \to K^+\ell^+\ell^-$ (right), showing the combined effect of the nonfactorisable and $\mathcal{O}(\alpha_s)$ corrections. Corrected values are shown with solid lines and dark fill colour, while uncorrected values are shown with dotted lines and light fill colour. $C_9 = C_9^{\mathrm{eff},0} - Y(q^2) = 4.114(14)$ [40] is included for comparison.

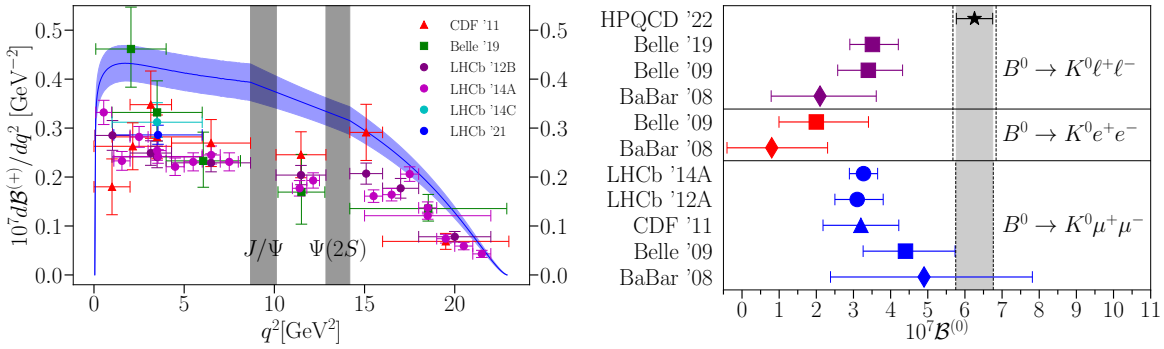

Figure 7.21: Left: differential branching fraction for $B^+ \to K^+\ell^+\ell^-$, with our result in blue, compared with experimental results [9, 193, 194, 200, 202, 204]. Note that Belle '19, and LHCb '14C and '21 have $\ell = e$, whilst otherwise $\ell = \mu$. Horizontal error bars indicate bin widths. Right: the total branching fraction for $B^0 \to K^0\ell^+\ell^-$. Our result (HPQCD '22) is given by the black band, as compared with experimental results [193, 194, 196, 199, 200, 201]. Dashed lines indicate the effect of adding QED uncertainty (see [4]) to our result.

the resonances, we linearly interpolate the differential decay rate and are therefore unaffected by the cusps. The most significant effect of the corrections is an approximately 20% shift to $\mathrm{Re}[C_7^{\mathrm{eff},1}]$ arising from $\mathrm{Re}[\delta C_7^{\mathrm{eff}}]$. The growth of corrections and in their uncertainties at high and low q^2 is suppressed by kinematic factors in the decay rate, resulting in minimal impact in the well-behaved regions of q^2 where we will give our main results.

We also note in Figure 7.20, that there is a slight channel dependence to $C_9^{\mathrm{eff},1}$, particularly at low q^2. This causes a slight difference in the shape of the differential branching fractions between $B^0 \to K^0\ell^+\ell^-$ and $B^+ \to K^+\ell^+\ell^-$.

Results

Comprehensive results for different combinations of meson charges and final state leptons are presented in [4] and tabulated below in Section 7.5. Here we shall plot the most important of these results.

Figure 7.21 shows on the left the differential branching fraction for $B^+ \to K^+\ell^+\ell^-$ (where $\ell = e, \mu$) across the full q^2 range, as well as binned experimental results. The vetoed regions discussed above are shown by black bands, and we linearly interpolate through and between them. We see that our result, the blue band, is considerably higher than experiment across

Channel	Result	q^2/GeV^2 range	$\mathcal{B} \times 10^7$	Tension
$B^+ \to K^+e^+e^-$	LHCb '21	$(1.1, 6)$	$1.401^{+0.074}_{-0.069} \pm 0.064$	-3.3σ
$B^+ \to K^+e^+e^-$	HPQCD '22	$(1.1, 6)$	2.07 ± 0.17	-
$B^0 \to K^0\mu^+\mu^-$	LHCb '14A	$(1.1, 6)$	$0.92^{+0.17}_{-0.15} \pm 0.044$	-3.6σ
$B^0 \to K^0\mu^+\mu^-$	HPQCD '22	$(1.1, 6)$	1.74 ± 0.15	-
$B^0 \to K^0\mu^+\mu^-$	LHCb '14A	$(15, 22)$	$0.67^{+0.11}_{-0.11} \pm 0.035$	-3.2σ
$B^0 \to K^0\mu^+\mu^-$	HPQCD '22	$(15, 22)$	1.16 ± 0.10	-
$B^+ \to K^+\mu^+\mu^-$	LHCb '14A	$(1.1, 6)$	$1.186 \pm 0.034 \pm 0.059$	-4.7σ
$B^+ \to K^+\mu^+\mu^-$	HPQCD '22	$(1.1, 6)$	2.07 ± 0.17	-
$B^+ \to K^+\mu^+\mu^-$	LHCb '14A	$(15, 22)$	$0.847 \pm 0.028 \pm 0.042$	-3.4σ
$B^+ \to K^+\mu^+\mu^-$	HPQCD '22	$(15, 22)$	1.26 ± 0.11	-

Table 7.11: Comparison of branching fractions with recent experimental results [9, 194] in well behaved regions of q^2. Note that the $B^+ \to K^+e^+e^-$ result quoted here from LHCb '21 is obtained using the $B^+ \to K^+\mu^+\mu^-$ result from LHCb '14A, combined with the ratio determined in LHCb '21. In the fifth column, the tension with our result is given by mean(Experiment - HPQCD)/sdev(Experiment - HPQCD).

most of the q^2 range. On the right we see the total branching fraction (i.e. the integral of the plot on the left), for $B^0 \to K^0\ell^+\ell^-$. Here, dotted lines indicate an additional QED uncertainty from final-state interactions of the charged leptons of 5%(2)% for the e (μ) cases. Including this uncertainty is a conservative move, as it is well accounted for [236, 237] in experimental results using PHOTOS [238], but we shall see that it has minimal effect on our results. ℓ is the average of the μ and e cases (see [4]). We see considerable tension with experimental results.

In Table 7.11, we compare our branching fractions with the most precise experimental results available for $B \to Ke^+e^-$ and $B \to K\mu^+\mu^-$, in both cases looking only at the well behaved regions below and above the $c\bar{c}$ resonances. We find our branching factions to be significantly higher than experiment, even just considering these theoretically clean regions. These same results, plotted in the form of a ratio of the branching fractions, are shown in Figure 7.22, with error bars indicating 1, 3 and 5 σ. It is worth noting that the corrections to $C_7^{\text{eff},0}$ and $C_9^{\text{eff},0}$ discussed above have made a difference of between 0.3 and 1 σ to these results.

In [4], we also calculated a new theoretical value for $R_e^\mu = \mathcal{B}^\mu/\mathcal{B}^e = 0.99954(26)_{\text{lat}}(1000)_{\text{QED}}$. Whilst this represents an improvement in the theoretical determination, it is dominated by QED uncertainty [239] and irrelevant for the tension with the recent LHCb result $R_e^\mu = 0.846(44)$ [9], which is almost entirely driven by experimental error.

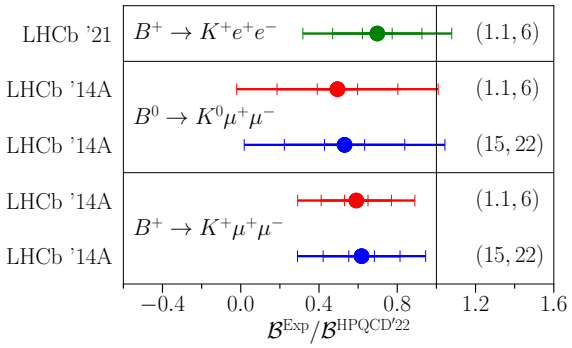

Figure 7.22: Comparison of branching fractions with recent experimental results [9, 194] in well behaved regions of q^2. Here we show the ratio of the experimental branching fraction to our results, meaning our results correspond to the line at 1. The error bars are 5σ long, with markers at 1, 3 and 5 σ. Note that the σ here are different from those calculated in Table 7.11. On the right, labels indicate q^2 bins in units of GeV^2.

7.4.2 $B \to K\nu\bar{\nu}$

We can also study the rare decays $B \to K\nu\bar{\nu}$, which are of phenomenological interest [240]. There is limited experimental data at the present time, but we anticipate more will be collected in future [241]. We can calculate the short distance contribution to the differential branching fraction (summed over neutrino flavours) [181, 242, 243], which depends only on the vector form factor.

$$\frac{d\mathcal{B}(B \to K\nu\bar{\nu})_{\text{SD}}}{dq^2} = \tau_B |V_{tb}V_{ts}^*|^2 \frac{(\eta_{\text{EW}} G_F)^2 \alpha_e^2 X_t^2}{32\pi^5 \sin^4\theta_W} |\vec{p}_K|^3 f_+^2(q^2), \tag{7.23}$$

with $X_t = 1.468(17)$ [244], and other quantities given in [4]. This short distance expression is correct in the case of $B^0 \to K^0\nu\bar{\nu}$ decays and dominates in $B^+ \to K^+\nu_\ell\bar{\nu}_\ell$ decays, except in the case $\ell = \tau$. In this case, long distance effects can come into play. These effects are calculated in Equation (2.27) of [181]. We repeat this calculation using some updated values. $|V_{ub}| = 0.00370(10)(12)$ is given from exclusive $B \to \pi\ell\nu$ decays in [129, 216], based on experimental data and lattice QCD form factors from [172, 245]. Other values, $|V_{us}|f_{K^+} = 0.03509(4)(4)\,\text{GeV}$ [128, 216, 246], $f_{B^+} = 0.1894(8)(11)(3)(1)\,\text{GeV}$ [79] and $\tau_\tau = 0.2903(5)\,\text{ps}$ [216] are used to obtain $\mathcal{B}(B^+ \to K^+\nu_\tau\bar{\nu_\tau})_{\text{LD}} = 6.26(55) \times 10^{-7}$, roughly 10% ($\approx 1\sigma$) of the short distance contribution.

We plot the short distance contribution (Equation 7.23) against q^2 on the left hand side of Figure 7.23. Our total branching fraction results have $\approx 10\%$ uncertainty. Current experimental

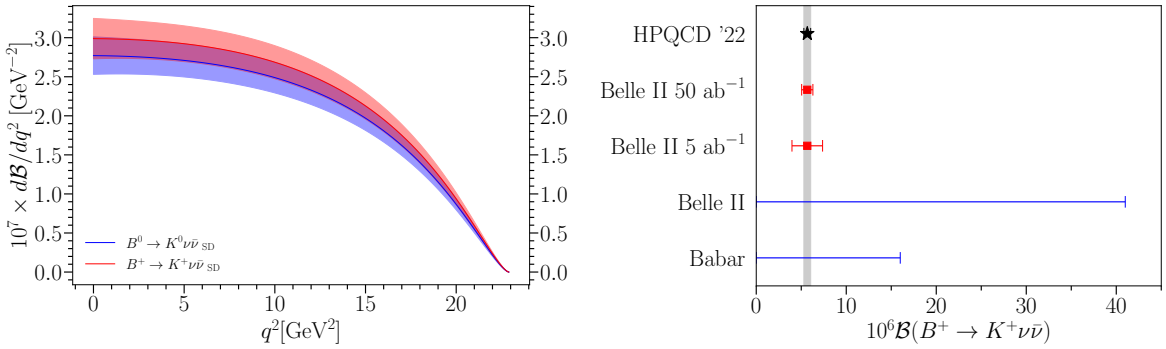

Figure 7.23: Left: the short distance contribution to $d\mathcal{B}(B \to K\nu\bar{\nu})/dq^2$, for both the charged and neutral cases. Right: our result for $B^+ \to K^+\nu\bar{\nu}$, including long distance effects, as compared with current 90% confidence limits from BaBar [247] and Belle II [248]. Theoretical uncertainties achieved by Belle II with 5 ab^{-1} and 50 ab^{-1} [241], assuming our central value, are also included.

bounds are roughly an order of magnitude larger, but our uncertainty complements the $\approx 10\%$ error expected from Belle II with 50 ab^{-1} [241], allowing for a more stringent test of this quantity in future. This is demonstrated on the right hand side of Figure 7.23, which compares our result with the current 90% confidence limits set by BaBar [247] and Belle II [248], as well as the theoretical future precision achieved by Belle II with 5 ab^{-1} and 50 ab^{-1} [241], assuming a result centred on our value.

7.5 Numerical results

q^2 bin	$(0.001, 4)$	$(4, 8)$	$(8, 12)$	$(12, 16)$	$(16, 20)$	$(20, q^2_{\text{max}})$
$10^6\mathcal{B}(B^+ \to K^+\nu\bar{\nu}$ SD$)$	1.189(97)	1.155(90)	1.071(84)	0.905(72)	0.597(48)	0.127(11)
$10^6\mathcal{B}(B^0 \to K^0\nu\bar{\nu}$ SD$)$	1.102(90)	1.071(83)	0.992(78)	0.837(66)	0.550(44)	0.1149(97)

Table 7.12: Short distance contributions to branching fractions $\mathcal{B}^{(0/+)} = \mathcal{B}(B^{0/+} \to K^{0/+}\nu\bar{\nu})_{\text{SD}}$ for the rare decay $B \to K\nu\bar{\nu}$ (Equation (7.23)) integrated over various q^2 bins. Numerical values for q^2 bins are in units of GeV^2.

q^2 bin	$(4m_\ell^2, q^2_{\rm max})$	$(0.05, 2)$	$(1, 6)$	$(2, 4.3)$	$(4.3, 8.68)$	$(14.18, 16)$	$(16, 18)$	$(18, 22)$
$10^7\mathcal{B}(B^+ \to K^+e^+e^-)$	7.04(55) [6.06(47)]	0.815(71)	2.11(18)	0.978(82)	1.77(16)	0.533(48)	0.478(43)	0.496(45)
$10^7\mathcal{B}(B^0 \to K^0e^+e^-)$	6.28(49) [5.39(42)]	0.710(62)	1.78(15)	0.814(70)	1.56(14)	0.493(44)	0.442(39)	0.454(41)
$10^7\mathcal{B}(B \to Ke^+e^-)$	6.66(52) [5.73(44)]	0.763(66)	1.94(16)	0.896(75)	1.67(14)	0.513(46)	0.460(41)	0.475(43)
$10^7\mathcal{B}(B^+ \to K^+\mu^+\mu^-)$	7.03(55) [6.06(47)]	0.809(70)	2.11(18)	0.978(82)	1.77(16)	0.533(48)	0.479(43)	0.498(45)
$10^7\mathcal{B}(B^0 \to K^0\mu^+\mu^-)$	6.26(49) [5.36(42)]	0.703(61)	1.78(15)	0.815(70)	1.56(14)	0.493(44)	0.442(39)	0.456(41)
$10^7\mathcal{B}(B \to K\mu^+\mu^-)$	6.64(51) [5.71(44)]	0.756(65)	1.94(16)	0.896(75)	1.67(14)	0.513(46)	0.461(41)	0.477(43)
$10^7\mathcal{B}(B^+ \to K^+\ell^+\ell^-)$	7.04(55) [6.06(47)]	0.812(70)	2.11(18)	0.978(82)	1.77(16)	0.533(48)	0.479(43)	0.497(45)
$10^7\mathcal{B}(B^0 \to K^0\ell^+\ell^-)$	6.27(49) [5.38(42)]	0.707(62)	1.78(15)	0.815(70)	1.56(14)	0.493(44)	0.442(39)	0.455(41)
$10^7\mathcal{B}(B \to K\ell^+\ell^-)$	6.65(51) [5.72(44)]	0.759(66)	1.94(16)	0.896(75)	1.67(14)	0.513(46)	0.460(41)	0.476(43)

q^2 bin	$(0.1, 2)$	$(2, 4)$	$(4, 6)$	$(6, 8)$	$(15, 17)$	$(17, 19)$	$(19, 22)$	$(1.1, 6)$	$(15, 22)$
$10^7\mathcal{B}(B^+ \to K^+e^+e^-)$	0.798(69)	0.852(71)	0.825(72)	0.803(72)	0.537(48)	0.412(37)	0.308(28)	2.07(17)	1.26(11)
$10^7\mathcal{B}(B^0 \to K^0e^+e^-)$	0.690(60)	0.708(60)	0.713(61)	0.713(63)	0.496(44)	0.380(34)	0.281(26)	1.74(15)	1.16(10)
$10^7\mathcal{B}(B \to Ke^+e^-)$	0.744(64)	0.780(65)	0.769(65)	0.758(66)	0.517(46)	0.396(35)	0.294(27)	1.90(16)	1.21(11)
$10^7\mathcal{B}(B^+ \to K^+\mu^+\mu^-)$	0.795(69)	0.852(71)	0.826(72)	0.803(72)	0.538(48)	0.413(37)	0.309(28)	2.07(17)	1.26(11)
$10^7\mathcal{B}(B^0 \to K^0\mu^+\mu^-)$	0.687(60)	0.708(60)	0.714(61)	0.714(63)	0.497(44)	0.381(34)	0.282(26)	1.74(15)	1.16(10)
$10^7\mathcal{B}(B \to K\mu^+\mu^-)$	0.741(64)	0.780(65)	0.770(65)	0.759(66)	0.518(46)	0.397(35)	0.296(27)	1.90(16)	1.21(11)
$10^7\mathcal{B}(B^+ \to K^+\ell^+\ell^-)$	0.796(69)	0.852(71)	0.825(72)	0.803(72)	0.538(48)	0.412(37)	0.308(28)	2.07(17)	1.26(11)
$10^7\mathcal{B}(B^0 \to K^0\ell^+\ell^-)$	0.688(60)	0.708(60)	0.713(61)	0.713(63)	0.497(44)	0.380(34)	0.282(26)	1.74(15)	1.16(10)
$10^7\mathcal{B}(B \to K\ell^+\ell^-)$	0.742(64)	0.780(65)	0.769(65)	0.758(66)	0.517(46)	0.396(35)	0.295(27)	1.90(16)	1.21(11)

q^2 bin	$(0.1, 4)$	$(4, 8.12)$	$(10.2, 12.8)$	$(14.18, q^2_{\rm max})$	$(10.09, 12.86)$	$(16, q^2_{\rm max})$
$10^7\mathcal{B}(B^+ \to K^+e^+e^-)$	1.65(14)	1.68(15)	0.918(76)	1.52(13)	0.979(81)	0.987(88)
$10^7\mathcal{B}(B^0 \to K^0e^+e^-)$	1.40(12)	1.47(13)	0.837(69)	1.40(12)	0.893(74)	0.906(81)
$10^7\mathcal{B}(B \to Ke^+e^-)$	1.52(13)	1.57(13)	0.878(72)	1.46(13)	0.936(77)	0.946(85)
$10^7\mathcal{B}(B^+ \to K^+\mu^+\mu^-)$	1.65(14)	1.68(15)	0.919(76)	1.52(13)	0.980(81)	0.990(88)
$10^7\mathcal{B}(B^0 \to K^0\mu^+\mu^-)$	1.39(12)	1.47(13)	0.838(69)	1.40(12)	0.894(74)	0.909(81)
$10^7\mathcal{B}(B \to K\mu^+\mu^-)$	1.52(13)	1.57(13)	0.878(72)	1.46(13)	0.937(77)	0.949(85)
$10^7\mathcal{B}(B^+ \to K^+\ell^+\ell^-)$	1.65(14)	1.68(15)	0.918(76)	1.52(13)	0.979(81)	0.988(88)
$10^7\mathcal{B}(B^0 \to K^0\ell^+\ell^-)$	1.40(12)	1.47(13)	0.838(69)	1.40(12)	0.893(74)	0.908(81)
$10^7\mathcal{B}(B \to K\ell^+\ell^-)$	1.52(13)	1.57(13)	0.878(72)	1.46(13)	0.936(77)	0.948(85)

Table 7.13: Branching fractions integrated over some commonly used q^2 binning schemes for the electron, muon and ℓ, which is the average of the two. In the first bin of the top panel, giving the branching fraction integrated over the full q^2 range, the first number is the whole integral, whilst the one which follows in square brackets is the result when the ranges 8.68-10.11 GeV^2 and 12.86-14.18 GeV^2 are excluded. Numerical values for q^2 bins are in units of GeV^2. QED corrections are not included here.

q^2 bin	$(4m_\mu^2, q^2_{\rm max})$	$(0.05, 2)$	$(1, 6)$	$(2, 4.3)$	$(4.3, 8.68)$	$(14.18, 16)$	$(16, 18)$	$(18, 22)$
$10^3(R_e^{\mu(+)} - 1)$	-0.08(28)	-7.84(95)	0.39(30)	0.41(30)	0.62(25)	1.43(23)	1.91(25)	3.81(37)
$10^3(R_e^{\mu(0)} - 1)$	-0.81(29)	-10.30(96)	0.44(34)	0.47(33)	0.65(27)	1.43(23)	1.92(25)	3.85(37)
$10^3(R_e^{\mu} - 1)$	-0.44(26)	-9.03(90)	0.41(32)	0.44(31)	0.63(26)	1.43(23)	1.92(25)	3.83(37)

q^2 bin	$(0.1, 2)$	$(2, 4)$	$(4, 6)$	$(6, 8)$	$(15, 17)$	$(17, 19)$	$(19, 22)$	$(1.1, 6)$	$(15, 22)$
$10^3(R_e^{\mu(+)} - 1)$	-3.29(76)	0.40(30)	0.54(27)	0.65(25)	1.63(24)	2.31(27)	4.57(42)	0.40(30)	2.57(29)
$10^3(R_e^{\mu(0)} - 1)$	-4.25(82)	0.46(34)	0.58(29)	0.68(26)	1.64(24)	2.32(27)	4.63(43)	0.45(33)	2.59(29)
$10^3(R_e^{\mu} - 1)$	-3.75(79)	0.43(32)	0.56(28)	0.66(25)	1.63(24)	2.32(27)	4.60(42)	0.43(32)	2.58(29)

q^2 bin	$(0.1, 4)$	$(4, 8.12)$	$(10.2, 12.8)$	$(14.18, q^2_{\rm max})$	$(10.09, 12.86)$	$(16, q^2_{\rm max})$
$10^3(R_e^{\mu(+)} - 1)$	-1.35(50)	0.60(26)	0.98(23)	2.50(29)	0.97(23)	3.08(32)
$10^3(R_e^{\mu(0)} - 1)$	-2.30(54)	0.63(27)	0.99(23)	2.51(29)	0.99(23)	3.10(32)
$10^3(R_e^{\mu} - 1)$	-1.80(52)	0.61(26)	0.98(23)	2.50(29)	0.98(23)	3.09(32)

Table 7.14: The ratio for $B \to K\mu^+\mu^-$ and $B \to Ke^+e^-$ branching fractions, R_e^μ, integrated over some commonly used q^2 binning schemes. We give results for the charged meson case, the neutral meson case and the charge-averaged case (defined as the ratio of the two charge-averaged integrals). Numerical values for q^2 bins are in units of GeV^2. A 1% uncertainty from QED effects [239] is not included in these numbers.

q^2 bin	$(4m_\tau^2, q^2_{\rm max})$	$(14.18, q^2_{\rm max})$	$(14.18, 16)$	$(16, 18)$	$(18, 22)$	$(15, 17)$	$(17, 19)$	$(19, 22)$	$(15, 22)$
$10^7\mathcal{B}(B^+ \to K^+\tau^+\tau^-)$	1.83(13)	1.62(11)	0.388(28)	0.462(33)	0.712(50)	0.454(32)	0.449(32)	0.492(34)	1.396(97)
$10^7\mathcal{B}(B^0 \to K^0\tau^+\tau^-)$	1.68(12)	1.49(10)	0.359(26)	0.427(30)	0.655(46)	0.420(30)	0.415(29)	0.452(31)	1.288(90)
$10^7\mathcal{B}(B \to K\tau^+\tau^-)$	1.75(12)	1.55(11)	0.374(27)	0.445(31)	0.683(48)	0.437(31)	0.432(30)	0.472(33)	1.342(94)

Table 7.15: Branching fractions for decays to τ leptons, integrated over some commonly used q^2 bins. Numerical values for q^2 bins are in units of GeV^2. No uncertainty to allow for QED is included in these numbers.

q^2 bin	$(4m_\tau^2, q^2_{\rm max})$	$(14.18, q^2_{\rm max})$	$(14.18, 16)$	$(16, 18)$	$(18, 22)$	$(15, 17)$	$(17, 19)$	$(19, 22)$	$(15, 22)$
$R_e^{\tau(+)}$	0.902(34)	1.065(42)	0.729(26)	0.966(36)	1.434(62)	0.846(31)	1.091(42)	1.600(74)	1.111(44)
$R_e^{\tau(0)}$	0.902(35)	1.066(42)	0.730(26)	0.967(36)	1.441(63)	0.847(31)	1.094(42)	1.610(74)	1.113(44)
R_e^{τ}	0.902(35)	1.065(42)	0.729(26)	0.967(36)	1.438(63)	0.847(31)	1.092(42)	1.605(74)	1.112(44)
$R_\mu^{\tau(+)}$	0.900(34)	1.063(42)	0.728(26)	0.964(36)	1.429(62)	0.845(31)	1.089(42)	1.592(73)	1.108(44)
$R_\mu^{\tau(0)}$	0.900(34)	1.063(42)	0.729(26)	0.966(36)	1.436(62)	0.846(31)	1.091(42)	1.603(73)	1.110(44)
R_μ^{τ}	0.900(34)	1.063(42)	0.728(26)	0.965(36)	1.432(62)	0.845(31)	1.090(42)	1.597(73)	1.109(44)
$R_\ell^{\tau(+)}$	0.901(34)	1.064(42)	0.729(26)	0.965(36)	1.432(62)	0.845(31)	1.090(42)	1.596(73)	1.110(44)
$R_\ell^{\tau(0)}$	0.901(34)	1.064(42)	0.729(26)	0.967(36)	1.438(63)	0.846(31)	1.092(42)	1.606(74)	1.112(44)
R_ℓ^{τ}	0.901(34)	1.064(42)	0.729(26)	0.966(36)	1.435(62)	0.846(31)	1.091(42)	1.601(74)	1.111(44)

Table 7.16: The ratio of branching fractions, R for the τ case to the electron, muon and ℓ, (which is the average of the two) cases, integrated over some commonly used q^2 bins. Numerical values for q^2 bins are in units of GeV^2. No uncertainty to allow for QED effects is included in these numbers.

7.6 Conclusions

We have performed the first fully relativistic calculation of the semileptonic $B \to K$ decay, using the heavy-HISQ technique on MILC $N_f = 2 + 1 + 1$ ensembles. We have obtained scalar, vector and tensor form factors which are roughly three times as precise as previous work at $q^2 = 0$ and used our them to illustrate spectator quark mass dependence in $b \to s$ and $c \to s$ semileptonic decays.

With an eye to phenomenology, we have calculated the differential branching fractions for $B \to K\ell^+\ell^-$ decays, paying special attention to corrections to the Wilson coefficients C_7 and C_9 which we have calculated, plotted and shown to be relatively small. We have further confirmed previously measured tensions between theory and experiment in the q^2 regions which are free of resonances, finding a tensions of $3 - 5\sigma$ for different channels. This provides a strong indication of new physics, or at least of a lack of understanding of the systematic errors present in this system. The latter option is somewhat disfavoured because of the exisiting tension in R_e^μ, which is theoretically cleaner.

We have provided updated differential braching fractions for $B \to K\nu\bar{\nu}$ decays, which have similar levels of uncertainty to those expected for Belle II at 50 ab^{-1}.

Chapter 8

Conclusions

The Standard Model must surely break eventually, but from which direction will the fatal stroke come? In this thesis we have presented lattice QCD updates on two areas within flavour physics where the SM can be tested.

Firstly, in Chapter 5, we presented a precise new determination of the scalar and vector form factors for the semileptonic decay $D \to K\ell\nu$ using the HISQ action on MILC $N_f = 2 + 1 + 1$ ensembles. Using these form factors, we determined the CKM matrix element V_{cs} using three semi-independent methods, with our preferred result giving $|V_{cs}| = 0.9663(80)$, the first time this quantity has been shown to be significantly less than one, and with sub percentage uncertainty. This new result significantly reduces the uncertainty on the new physics sensitive unitarity constraints it is involved in, and they remain consistent with unity. This new value has now been appropriately averaged with other semileptonic determinations, and with their leptonic counterparts, by the PDG review [29], halving the uncertainty on the global average value of $|V_{cs}|$, which will in turn feed into CKM fits and other work, tightening the screw on the SM.

In Chapter 6, we sought to explore the viability of a heavy-HISQ $B \to K$ form factor calculation, via a heavier spectator mass proxy, giving rise to the unphysical $B_s \to \eta_s$ calculation presented there. We found heavy-HISQ to be effective for $b \to s$ decays with light spectators, with our resulting scalar and vector form factors' errors dominated by statistical uncertainty. We showed that the spectator quark mass dependence of the form factors was mild for light quarks, exhibiting only slight breaking of SU(3) flavour symmetry, and we also used our form factors to test some predictions from heavy quark effective theory, with mixed results.

Building on the heavy-HISQ work of Chapter 6, we presented the first fully relativistic determination of the $B \to K$ semileptonic form factors in Chapter 7, providing results free from the matching uncertainties of previous effective theory calculations. Again, we used the HISQ action on MILC $N_f = 2+1+1$ ensembles, this time calculating scalar, vector and tensor form factors. We made use of the $B_s \to \eta_s$ data from Chapter 6 by folding this in to the chiral logarithm expansion in the spectator quark mass. We used our new form factors to update the theoretical determinations of the branching fraction for $B \to K\ell^+\ell^-$ and other quantities, which are good places to look for new physics. We found tensions of between 3 and 5 σ with experimental results from LHCb, in resonance free parts of the q^2 range; a strong indication of new physics, also hinted at by the tension in the theoretically cleaner ratio R_e^μ.

Bibliography

[1] B. Chakraborty, W. G. Parrott, C. Bouchard, C. T. H. Davies, J. Koponen, and G. P. Lepage ((HPQCD Collaboration)§, HPQCD), Phys. Rev. D **104**, 034505 (2021), arXiv:2104.09883 [hep-lat] .

[2] W. G. Parrott, C. Bouchard, C. T. H. Davies, and D. Hatton, Phys. Rev. D **103**, 094506 (2021), arXiv:2010.07980 [hep-lat] .

[3] W. G. Parrott, C. Bouchard, and C. T. H. Davies, (2022), arXiv:2207.12468 [hep-lat] .

[4] W. G. Parrott, C. Bouchard, and C. T. H. Davies, (2022), arXiv:2207.13371 [hep-ph] .

[5] B. Abi *et al.* (Muon g-2), Phys. Rev. Lett. **126**, 141801 (2021), arXiv:2104.03281 [hep-ex] .

[6] E. Graverini (ATLAS, CMS, LHCb), J. Phys. Conf. Ser. **1137**, 012025 (2019), arXiv:1807.11373 [hep-ex] .

[7] P. A. Boyle *et al.*, in *2022 Snowmass Summer Study* (2022) arXiv:2205.15373 [hep-lat] .

[8] D. Loose, *Flavorful new physics models in the light of the B decay anomalies*, Ph.D. thesis, Dortmund U. (2021).

[9] R. Aaij *et al.* (LHCb), Nature Phys. **18**, 277 (2022), arXiv:2103.11769 [hep-ex] .

[10] E. Follana, Q. Mason, C. Davies, K. Hornbostel, G. Lepage, J. Shigemitsu, H. Trottier, and K. Wong (HPQCD, UKQCD), Phys. Rev. D **75**, 054502 (2007), arXiv:hep-lat/0610092 .

[11] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory* (2016).

[12] M. Srednicki, *Quantum Field Theory* (2006).

[13] D. Tong, "Lectures on Quantum Field Theory," .

[14] B. Gripaios, "Gauge Field Theory," .

[15] H. A. Lorentz, "Electromagnetic phenomena in a system moving with any velocity smaller than that of light," in *Collected Papers: Volume V* (Springer Netherlands, Dordrecht, 1937) pp. 172–197.

[16] P. A. M. Dirac, Proc. Roy. Soc. Lond. A **117**, 610 (1928).

[17] M. Gell-Mann, Physics Letters **8**, 214 (1964).

[18] H. Fritzsch, M. Gell-Mann, and H. Leutwyler, Phys. Lett. B **47**, 365 (1973).

[19] A. Salam and J. C. Ward, Phys. Lett. **13**, 168 (1964).

[20] S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967).

[21] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson, Phys. Rev. **105**, 1413 (1957).

[22] P. W. Higgs, Phys. Rev. Lett. **13**, 508 (1964).

[23] F. Englert and R. Brout, Phys. Rev. Lett. **13**, 321 (1964).

[24] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, Phys. Rev. Lett. **13**, 585 (1964).

[25] H. Yukawa, Progress of Theoretical Physics Supplement **1**, 1 (1955), https://academic.oup.com/ptps/article-pdf/doi/10.1143/PTPS.1.1/5310694/1-1.pdf .

[26] N. Cabibbo, *Meeting of the Italian School of Physics and Weak Interactions Bologna, Italy, April 26-28, 1984*, Phys. Rev. Lett. **10**, 531 (1963), [,648(1963)].

[27] M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49**, 652 (1973).

[28] E. Kuflik, Y. Nir, and T. Volansky, Phys. Rev. Lett. **110**, 091801 (2013), arXiv:1204.1975 [hep-ph] .

[29] P. D. Group, Progress of Theoretical and Experimental Physics **2020** (2020), 10.1093/ptep/ptaa104, 083C01, https://academic.oup.com/ptep/article-pdf/2020/8/083C01/33653179/ptaa104.pdf .

[30] L. Wolfenstein, Phys. Rev. Lett. **51**, 1945 (1983).

[31] J. Charles, A. Hocker, H. Lacker, S. Laplace, F. R. Le Diberder, J. Malcles, J. Ocariz, M. Pivk, and L. Roos (CKMfitter Group), Eur. Phys. J. C **41**, 1 (2005), arXiv:hep-ph/0406184 .

[32] A. Crivellin, G. D'Ambrosio, and J. Heeck, Phys. Rev. Lett. **114**, 151801 (2015), arXiv:1501.00993 [hep-ph] .

[33] W. Altmannshofer, P. Stangl, and D. M. Straub, Phys. Rev. D **96**, 055008 (2017), arXiv:1704.05435 [hep-ph] .

[34] K. Abe *et al.* (Belle), Phys. Rev. Lett. **92**, 171802 (2004), arXiv:hep-ex/0310029 .

[35] A. M. Baldini *et al.* (MEG), Eur. Phys. J. C **76**, 434 (2016), arXiv:1605.05081 [hep-ex] .

[36] O. W. Greenberg and C. A. Nelson, Phys. Rev. D **10**, 2567 (1974).

[37] M. Bauer and M. Neubert, Phys. Rev. Lett. **116**, 141802 (2016), arXiv:1511.01900 [hep-ph] .

[38] I. Brivio and M. Trott, Phys. Rept. **793**, 1 (2019), arXiv:1706.08945 [hep-ph] .

[39] B. Grzadkowski, M. Iskrzynski, M. Misiak, and J. Rosiek, JHEP **10**, 085 (2010), arXiv:1008.4884 [hep-ph] .

[40] T. Blake, G. Lanfranchi, and D. M. Straub, Prog. Part. Nucl. Phys. **92**, 50 (2017), arXiv:1606.00916 [hep-ph] .

[41] G. C. Wick, Phys. Rev. **96**, 1124 (1954).

[42] K. G. Wilson, Phys. Rev. D **10**, 2445 (1974).

[43] G. P. Lepage, in *13th Annual HUGS AT CEBAF (HUGS 98)* (1998) pp. 49–90, arXiv:hep-lat/0506036 .

[44] S. Weinzierl, (2000), arXiv:hep-ph/0006269 .

[45] M. Creutz, L. Jacobs, and C. Rebbi, Phys. Rev. D **20**, 1915 (1979).

[46] S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth, Phys. Lett. B **195**, 216 (1987).

[47] D. J. E. Callaway and A. Rahman, Phys. Rev. Lett. **49**, 613 (1982).

[48] D. J. E. Callaway and A. Rahman, Phys. Rev. D **28**, 1506 (1983).

[49] S. Duane and J. B. Kogut, Nucl. Phys. B **275**, 398 (1986).

[50] S. Duane and J. B. Kogut, Phys. Rev. Lett. **55**, 2774 (1985).

[51] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).

[52] K. G. Wilson, NATO Sci. Ser. B **59**, 363 (1980).

[53] M. Luscher and P. Weisz, Commun. Math. Phys. **97**, 59 (1985), [Erratum: Commun.Math.Phys. 98, 433 (1985)].

[54] K. Symanzik, Nucl. Phys. B **226**, 187 (1983).

[55] G. P. Lepage and P. B. Mackenzie, Phys. Rev. D **48**, 2250 (1993), arXiv:hep-lat/9209022 .

[56] L. H. Karsten and J. Smit, Nucl. Phys. B **183**, 103 (1981).

[57] N. Kawamoto and J. Smit, Nucl. Phys. B **192**, 100 (1981).

[58] B. Sheikholeslami and R. Wohlert, Nucl. Phys. B **259**, 572 (1985).

[59] S. Naik, Nucl. Phys. B **316**, 238 (1989).

[60] G. P. Lepage, Phys. Rev. D **59**, 074502 (1999), arXiv:hep-lat/9809157 .

[61] G. P. Lepage, Physical Review D **59**, 074502 (1999).

[62] S. Naik, Nuclear Physics B **316**, 238 (1989).

[63] C. Monahan, J. Shigemitsu, and R. Horgan, Phys. Rev. D **87**, 034017 (2013), arXiv:1211.6966 [hep-lat] .

[64] J. B. Kogut and L. Susskind, Phys. Rev. D **11**, 395 (1975).

[65] C. T. Sachrajda and G. Villadoro, Phys. Lett. B **609**, 73 (2005), arXiv:hep-lat/0411033 .

[66] D. Guadagnoli, F. Mescia, and S. Simula, Phys. Rev. D **73**, 114504 (2006), arXiv:hep-lat/0512020 .

[67] P. F. Bedaque, Phys. Lett. B **593**, 82 (2004), arXiv:nucl-th/0402051 .

[68] G. M. de Divitiis, R. Petronzio, and N. Tantalo, Phys. Lett. B **595**, 408 (2004), arXiv:hep-lat/0405002 .

[69] J. M. Flynn, A. Juttner, and C. T. Sachrajda (UKQCD), Phys. Lett. B **632**, 313 (2006), arXiv:hep-lat/0506016 .

[70] M. R. Hestenes and E. Stiefel, Journal of research of the National Bureau of Standards **49**, 409 (1952).

[71] S. A. Gottlieb, W. Liu, D. Toussaint, R. L. Renken, and R. L. Sugar, Phys. Rev. D **35**, 2531 (1987).

[72] S. Borsanyi *et al.*, JHEP **09**, 010 (2012), arXiv:1203.4469 [hep-lat] .

[73] M. Lüscher, JHEP **08**, 071 (2010), [Erratum: JHEP 03, 092 (2014)], arXiv:1006.4518 [hep-lat] .

[74] S. Durr, Z. Fodor, C. Hoelbling, S. D. Katz, S. Krieg, T. Kurth, L. Lellouch, T. Lippert, K. K. Szabo, and G. Vulvert, JHEP **08**, 148 (2011), arXiv:1011.2711 [hep-lat] .

[75] R. J. Dowdall, C. T. H. Davies, G. P. Lepage, and C. McNeile, Phys. Rev. **D88**, 074504 (2013), arXiv:1303.1670 [hep-lat] .

[76] A. Bazavov *et al.* (MILC), Phys. Rev. D **82**, 074501 (2010), arXiv:1004.0342 [hep-lat] .

[77] A. Bazavov *et al.* (MILC), Phys. Rev. D **87**, 054505 (2013), arXiv:1212.4768 [hep-lat] .

[78] A. Bazavov *et al.* (MILC), Phys. Rev. D **93**, 094510 (2016), arXiv:1503.02769 [hep-lat] .

[79] A. Bazavov *et al.*, Phys. Rev. **D98**, 074512 (2018), arXiv:1712.09262 [hep-lat] .

[80] B. Chakraborty, C. T. H. Davies, B. Galloway, P. Knecht, J. Koponen, G. C. Donald, R. J. Dowdall, G. P. Lepage, and C. McNeile, Phys. Rev. **D91**, 054508 (2015), arXiv:1408.4169 [hep-lat] .

[81] A. Bazavov, C. Bernard, C. DeTar, W. Freeman, S. Gottlieb, U. M. Heller, J. E. Hetrick, J. Laiho, L. Levkova, M. Oktay, J. Osborn, R. L. Sugar, D. Toussaint, and R. S. Van de Water (MILC Collaboration), Phys. Rev. D **82**, 074501 (2010).

[82] A. Bazavov, C. Bernard, J. Komijani, C. DeTar, L. Levkova, W. Freeman, S. Gottlieb, R. Zhou, U. M. Heller, J. E. Hetrick, J. Laiho, J. Osborn, R. L. Sugar, D. Toussaint, and R. S. Van de Water (MILC Collaboration), Phys. Rev. D **87**, 054505 (2013).

[83] M. Foster and C. Michael (UKQCD), Phys. Rev. D **59**, 074503 (1999), arXiv:hep-lat/9810021 .

[84] C. McNeile and C. Michael (UKQCD), Phys. Rev. D **73**, 074506 (2006), arXiv:hep-lat/0603007 .

[85] P. Lepage and C. Gohlke, "gplepage/lsqfit: lsqfit version 12.0.1," (2021).

[86] P. Lepage, "gplepage/corrfitter: corrfitter version 8.2," (2021).

[87] P. Lepage, C. Gohlke, and D. Hackett, "gplepage/gvar: gvar version 11.9.4," (2021).

[88] B. P. Carlin and T. A. Louis, Statistics and Computing **7**, 153–154 (1997).

[89] G. P. Lepage, B. Clark, C. T. H. Davies, K. Hornbostel, P. B. Mackenzie, C. Morningstar, and H. Trottier, *Lattice field theory. Proceedings, 19th International Symposium, Lattice 2001, Berlin, Germany, August 19-24, 2001*, Nucl. Phys. Proc. Suppl. **106**, 12 (2002), arXiv:hep-lat/0110175 [hep-lat] .

[90] R. J. Dowdall, C. T. H. Davies, R. R. Horgan, G. P. Lepage, C. J. Monahan, J. Shigemitsu, and M. Wingate, Phys. Rev. D **100**, 094508 (2019), arXiv:1907.01025 [hep-lat] .

[91] C. Michael and A. McKerrell, Phys. Rev. D **51**, 3745 (1995), arXiv:hep-lat/9412087 .

[92] C. Michael, Phys. Rev. D **49**, 2616 (1994), arXiv:hep-lat/9310026 .

[93] E. McLean, C. T. H. Davies, A. T. Lytle, and J. Koponen, Phys. Rev. D **99**, 114512 (2019), arXiv:1904.02046 [hep-lat] .

[94] H. Na, C. T. H. Davies, E. Follana, G. P. Lepage, and J. Shigemitsu, Phys. Rev. **D82**, 114506 (2010), arXiv:1008.4562 [hep-lat] .

[95] D. Hatton, C. T. H. Davies, G. P. Lepage, and A. T. Lytle (HPQCD), Phys. Rev. D **100**, 114513 (2019), arXiv:1909.00756 [hep-lat] .

[96] G. C. Donald, C. T. H. Davies, J. Koponen, and G. P. Lepage (HPQCD), Phys. Rev. **D90**, 074506 (2014), arXiv:1311.6669 [hep-lat] .

[97] J. Harrison, C. Davies, and M. Wingate (HPQCD), Phys. Rev. D **97**, 054502 (2018), arXiv:1711.11013 [hep-lat] .

[98] L. J. Cooper, C. T. Davies, J. Harrison, J. Komijani, and M. Wingate (HPQCD), Phys. Rev. D **102**, 014513 (2020), arXiv:2003.00914 [hep-lat] .

[99] D. Hatton, C. T. H. Davies, G. P. Lepage, and A. T. Lytle (HPQCD), Phys. Rev. D **102**, 094509 (2020), arXiv:2008.02024 [hep-lat] .

[100] C. G. Boyd, B. Grinstein, and R. F. Lebed, Nucl. Phys. B **461**, 493 (1996), arXiv:hep-ph/9508211 .

[101] M. C. Arnesen, B. Grinstein, I. Z. Rothstein, and I. W. Stewart, Phys. Rev. Lett. **95**, 071802 (2005), arXiv:hep-ph/0504209 .

[102] C. Bourrely, I. Caprini, and L. Lellouch, Phys. Rev. **D79**, 013008 (2009), [Erratum: Phys. Rev.D82,099902(2010)], arXiv:0807.2722 [hep-ph] .

[103] J. Bijnens and I. Jemos, Nucl. Phys. B **846**, 145 (2011), arXiv:1011.6531 [hep-ph] .

[104] C. M. Bouchard, G. P. Lepage, C. Monahan, H. Na, and J. Shigemitsu, Phys. Rev. **D90**, 054506 (2014), arXiv:1406.2279 [hep-lat] .

[105] C. Bernard (MILC), Phys. Rev. D **65**, 054031 (2002), arXiv:hep-lat/0111051 .

[106] G. Lepage, L. Magnea, C. Nakhleh, U. Magnea, and K. Hornbostel, Phys. Rev. D **46**, 4052 (1992), arXiv:hep-lat/9205007 .

[107] R. J. Dowdall *et al.* (HPQCD), Phys. Rev. D **85**, 054509 (2012), arXiv:1110.6887 [hep-lat] .

[108] B. Colquhoun, R. J. Dowdall, C. T. H. Davies, K. Hornbostel, and G. P. Lepage, Phys. Rev. D **91**, 074514 (2015), arXiv:1408.5768 [hep-lat] .

[109] C. Bouchard, G. Lepage, C. Monahan, H. Na, and J. Shigemitsu (HPQCD), Phys. Rev. D **88**, 054509 (2013), [Erratum: Phys.Rev.D 88, 079901 (2013)], arXiv:1306.2384 [hep-lat] .

[110] H. Georgi, Phys. Lett. B **240**, 447 (1990).

[111] E. Eichten and B. R. Hill, Phys. Lett. B **234**, 511 (1990).

[112] M. Neubert, Subnucl. Ser. **34**, 98 (1997), arXiv:hep-ph/9610266 .

[113] P. Dimopoulos *et al.* (ETM), JHEP **01**, 046 (2012), arXiv:1107.1441 [hep-lat] .

[114] B. Blossier *et al.* (ETM), JHEP **04**, 049 (2010), arXiv:0909.3187 [hep-lat] .

[115] N. Carrasco *et al.* (ETM), JHEP **03**, 016 (2014), arXiv:1308.1851 [hep-lat] .

[116] A. Juttner, *Precision lattice computations in the heavy quark sector*, Other thesis (2004), arXiv:hep-lat/0503040 .

[117] A. Bussone *et al.* (ETM), Phys. Rev. D **93**, 114505 (2016), arXiv:1603.04306 [hep-lat] .

[118] P. A. Boyle, L. Del Debbio, N. Garron, A. Juttner, A. Soni, J. T. Tsang, and O. Witzel (RBC/UKQCD), (2018), arXiv:1812.08791 [hep-lat] .

[119] P. A. Boyle, L. Del Debbio, A. Jüttner, A. Khamseh, F. Sanfilippo, and J. T. Tsang, JHEP **12**, 008 (2017), arXiv:1701.02644 [hep-lat] .

[120] J. Rolf, M. Della Morte, S. Durr, J. Heitger, A. Juttner, H. Molke, A. Shindler, and R. Sommer (ALPHA), Nucl. Phys. B Proc. Suppl. **129**, 322 (2004), arXiv:hep-lat/0309072 .

[121] R. Sommer, Nucl. Phys. B Proc. Suppl. **119**, 185 (2003), arXiv:hep-lat/0209162 .

[122] J. Heitger and R. Sommer (ALPHA), Nucl. Phys. B Proc. Suppl. **106**, 358 (2002), arXiv:hep-lat/0110016 .

[123] M. Neubert, in *Theoretical Advanced Study Institute in Elementary Particle Physics: Physics in $D \geqq 4$* (2005) pp. 149–194, arXiv:hep-ph/0512222 .

[124] F. Bahr, D. Banerjee, F. Bernardoni, A. Joseph, M. Koren, H. Simma, and R. Sommer (ALPHA), Phys. Lett. **B757**, 473 (2016), arXiv:1601.04277 [hep-lat] .

[125] D. Banerjee (ALPHA), *Proceedings, 34th International Symposium on Lattice Field Theory (Lattice 2016): Southampton, UK, July 24-30, 2016*, PoS **LATTICE2016**, 292 (2016), arXiv:1701.03923 [hep-lat] .

[126] N. Cabibbo, Phys. Rev. Lett. **10**, 531 (1963).

[127] M. Kobayashi and T. Maskawa, Progress of Theoretical Physics **49**, 652 (1973), https://academic.oup.com/ptp/article-pdf/49/2/652/5257692/49-2-652.pdf .

[128] J. L. Rosner, S. Stone, and R. S. Van de Water, (2015), arXiv:1509.02220 [hep-ph] .

[129] Y. S. Amhis *et al.* (HFLAV), Eur. Phys. J. C **81**, 226 (2021), arXiv:1909.12524 [hep-ex] .

[130] B. Blossier *et al.* (ETM), JHEP **07**, 043 (2009), arXiv:0904.0954 [hep-lat] .

[131] C. Aubin *et al.*, Phys. Rev. Lett. **95**, 122002 (2005), arXiv:hep-lat/0506030 .

[132] Y. Aoki *et al.*, (2021), arXiv:2111.09849 [hep-lat] .

[133] E. Follana, C. T. H. Davies, G. P. Lepage, and J. Shigemitsu (HPQCD, UKQCD), Phys. Rev. Lett. **100**, 062002 (2008), arXiv:0706.1726 [hep-lat] .

[134] C. T. H. Davies, C. McNeile, E. Follana, G. P. Lepage, H. Na, and J. Shigemitsu, Phys. Rev. D **82**, 114504 (2010), arXiv:1008.4018 [hep-lat] .

[135] N. Carrasco *et al.*, Phys. Rev. D **91**, 054507 (2015), arXiv:1411.7908 [hep-lat] .

[136] P. Golonka and Z. Was, Eur. Phys. J. C **45**, 97 (2006), arXiv:hep-ph/0506026 .

[137] C. Aubin *et al.* (Fermilab Lattice, MILC, HPQCD), Phys. Rev. Lett. **94**, 011601 (2005), arXiv:hep-ph/0408306 .

[138] J. Koponen, C. T. H. Davies, G. C. Donald, E. Follana, G. P. Lepage, H. Na, and J. Shigemitsu, (2013), arXiv:1305.1462 [hep-lat] .

[139] V. Lubicz, L. Riggio, G. Salerno, S. Simula, and C. Tarantino (ETM), Phys. Rev. **D96**, 054514 (2017), [erratum: Phys. Rev.D99,no.9,099902(2019); Erratum: Phys. Rev.D100,no.7,079901(2019)], arXiv:1706.03017 [hep-lat] .

[140] L. Riggio, G. Salerno, and S. Simula, Eur. Phys. J. C **78**, 501 (2018), arXiv:1706.03657 [hep-lat] .

[141] R. Frezzotti, P. A. Grassi, S. Sint, and P. Weisz (Alpha), JHEP **08**, 058 (2001), arXiv:hep-lat/0101001 .

[142] R. Frezzotti and G. C. Rossi, Nucl. Phys. B Proc. Suppl. **128**, 193 (2004), arXiv:hep-lat/0311008 .

[143] R. Frezzotti and G. C. Rossi, JHEP **08**, 007 (2004), arXiv:hep-lat/0306014 .

[144] T. Kaneko, B. Fahy, H. Fukaya, and S. Hashimoto (JLQCD), PoS **LATTICE2016**, 297 (2017), arXiv:1701.00942 [hep-lat] .

[145] R. Li *et al.* (Fermilab Lattice, MILC), PoS **LATTICE2018**, 269 (2019), arXiv:1901.08989 [hep-lat] .

[146] S. Aoki *et al.* (Flavour Lattice Averaging Group), Eur. Phys. J. C **80**, 113 (2020), arXiv:1902.08191 [hep-lat] .

[147] E. McLean, C. Davies, J. Koponen, and A. Lytle, Phys. Rev. D **101**, 074513 (2020), arXiv:1906.00701 [hep-lat] .

[148] J. Lees *et al.* (BaBar), Phys. Rev. D **88**, 052003 (2013), [Erratum: Phys.Rev.D 88, 079902 (2013)], arXiv:1304.5009 [hep-ex] .

[149] M. Steffen, Astron. Astrophys. **239**, 443 (1990).

[150] D. Besson *et al.* (CLEO), Phys. Rev. D **80**, 032005 (2009), arXiv:0906.2983 [hep-ex] .

[151] B. Aubert *et al.* (BaBar), Phys. Rev. D **76**, 052005 (2007), arXiv:0704.0020 [hep-ex] .

[152] M. Ablikim *et al.* (BESIII), Phys. Rev. D **92**, 072012 (2015), arXiv:1508.07560 [hep-ex] .

[153] Y. Amhis *et al.* (HFLAV), Eur. Phys. J. C **77**, 895 (2017), arXiv:1612.07233 [hep-ex] .

[154] A. Sirlin, Nuclear Physics B **196**, 83 (1982).

[155] M. Antonelli *et al.* (FlaviaNet Working Group on Kaon Decays), Eur. Phys. J. C **69**, 399 (2010), arXiv:1005.2323 [hep-ph] .

[156] M. Ablikim *et al.* (BESIII), Phys. Rev. Lett. **122**, 011804 (2019), arXiv:1810.03127 [hep-ex] .

[157] S. Fajfer, I. Nisandzic, and U. Rojec, Phys. Rev. D **91**, 094009 (2015), arXiv:1502.07488 [hep-ph] .

[158] M. Ablikim *et al.* (BESIII), Phys. Rev. D **96**, 012002 (2017), arXiv:1703.09084 [hep-ex] .

[159] G. D'Agostini, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **346**, 306 (1994).

[160] M. Ablikim *et al.* (BESIII), Eur. Phys. J. C **76**, 369 (2016), arXiv:1605.00068 [hep-ex] .

[161] M. Ablikim *et al.* (BESIII), Phys. Rev. D **104**, 052008 (2021), arXiv:2104.08081 [hep-ex] .

[162] L. Widhalm *et al.* (Belle), Phys. Rev. Lett. **97**, 061804 (2006), arXiv:hep-ex/0604049 [hep-ex] .

[163] M. Ablikim *et al.* (BESIII), Phys. Rev. D **92**, 112008 (2015), arXiv:1510.00308 [hep-ex] .

[164] S. Aoki *et al.* (Flavour Lattice Averaging Group), Eur. Phys. J. **C80**, 113 (2020), arXiv:1902.08191 [hep-lat] .

[165] A. Bazavov *et al.*, Phys. Rev. Lett. **112**, 112001 (2014), arXiv:1312.1228 [hep-ph] .

[166] N. Carrasco, P. Lami, V. Lubicz, L. Riggio, S. Simula, and C. Tarantino, Phys. Rev. D **93**, 114512 (2016), arXiv:1602.04113 [hep-lat] .

[167] A. Bazavov *et al.* (Fermilab Lattice, MILC), Phys. Rev. D **99**, 114509 (2019), arXiv:1809.02827 [hep-lat] .

[168] J. Harrison, C. T. H. Davies, and A. Lytle (HPQCD), Phys. Rev. D **102**, 094518 (2020), arXiv:2007.06957 [hep-lat] .

[169] J. Harrison, C. T. H. Davies, and A. Lytle (LATTICE-HPQCD), Phys. Rev. Lett. **125**, 222003 (2020), arXiv:2007.06956 [hep-lat] .

[170] R. J. Dowdall, C. T. H. Davies, T. C. Hammant, and R. R. Horgan, Phys. Rev. **D86**, 094510 (2012), arXiv:1207.5149 [hep-lat] .

[171] B. Colquhoun, R. J. Dowdall, J. Koponen, C. T. H. Davies, and G. P. Lepage, Phys. Rev. **D93**, 034502 (2016), arXiv:1510.07446 [hep-lat] .

[172] J. A. Bailey *et al.* (Fermilab Lattice, MILC), Phys. Rev. **D92**, 014024 (2015), arXiv:1503.07839 [hep-lat] .

[173] C. Bouchard, G. P. Lepage, C. Monahan, H. Na, and J. Shigemitsu (HPQCD), Phys. Rev. **D88**, 054509 (2013), [Erratum: Phys. Rev.D88,no.7,079901(2013)], arXiv:1306.2384 [hep-lat] .

[174] R. J. Hill, Phys. Rev. **D73**, 014012 (2006), arXiv:hep-ph/0505129 [hep-ph] .

[175] R. J. Hill, *Proceedings, 4th Conference on Flavor Physics and CP Violation (FPCP 2006): Vancouver, British Columbia, Canada, April 9-12, 2006*, eConf **C060409**, 027 (2006), arXiv:hep-ph/0606023 [hep-ph] .

[176] G. Burdman, Z. Ligeti, M. Neubert, and Y. Nir, Phys. Rev. **D49**, 2331 (1994), arXiv:hep-ph/9309272 [hep-ph] .

[177] J. M. Flynn, P. Fritzsch, T. Kawanai, C. Lehner, B. Samways, C. T. Sachrajda, R. S. Van de Water, and O. Witzel (RBC, UKQCD), Phys. Rev. **D93**, 014510 (2016), arXiv:1506.06413 [hep-lat] .

[178] V. Chernyak and I. Zhitnitsky, Nucl. Phys. B **345**, 137 (1990).

[179] D. Becirevic and A. B. Kaidalov, Phys. Lett. B **478**, 417 (2000), arXiv:hep-ph/9904490 .

[180] C. Bouchard, G. P. Lepage, C. Monahan, H. Na, and J. Shigemitsu (HPQCD), Phys. Rev. Lett. **111**, 162002 (2013), [Erratum: Phys.Rev.Lett. 112, 149902 (2014)], arXiv:1306.0434 [hep-ph] .

[181] D. Du, A. El-Khadra, S. Gottlieb, A. Kronfeld, J. Laiho, E. Lunghi, R. Van de Water, and R. Zhou, Phys. Rev. D **93**, 034005 (2016), arXiv:1510.02349 [hep-ph] .

[182] J. A. Bailey *et al.*, Phys. Rev. **D93**, 025026 (2016), arXiv:1509.06235 [hep-lat] .

[183] C. Bobeth, G. Hiller, and G. Piranishvili, JHEP **12**, 040 (2007), arXiv:0709.4174 [hep-ph] .

[184] C. Bobeth, G. Hiller, D. van Dyk, and C. Wacker, JHEP **01**, 107 (2012), arXiv:1111.2558 [hep-ph] .

[185] C. Bobeth, G. Hiller, and D. van Dyk, Phys. Rev. D **87**, 034016 (2013), arXiv:1212.2321 [hep-ph] .

[186] A. Khodjamirian, T. Mannel, and Y. M. Wang, JHEP **02**, 010 (2013), arXiv:1211.0234 [hep-ph] .

[187] W. Altmannshofer and D. M. Straub, JHEP **08**, 121 (2012), arXiv:1206.0273 [hep-ph] .

[188] W.-F. Wang and Z.-J. Xiao, Phys. Rev. D **86**, 114025 (2012), arXiv:1207.0265 [hep-ph] .

[189] N. Gubernari, M. Reboud, D. van Dyk, and J. Virto, (2022), arXiv:2206.03797 [hep-ph] .

[190] P. Ball and R. Zwicky, Phys. Rev. D **71**, 014015 (2005), arXiv:hep-ph/0406232 .

[191] A. Khodjamirian, T. Mannel, A. A. Pivovarov, and Y. M. Wang, JHEP **09**, 089 (2010), arXiv:1006.4945 [hep-ph] .

[192] A. Bharucha, T. Feldmann, and M. Wick, JHEP **09**, 090 (2010), arXiv:1004.3249 [hep-ph] .

[193] S. Choudhury *et al.* (BELLE), JHEP **03**, 105 (2021), arXiv:1908.01848 [hep-ex] .

[194] R. Aaij *et al.* (LHCb), JHEP **06**, 133 (2014), arXiv:1403.8044 [hep-ex] .

[195] R. Aaij *et al.* (LHCb), Eur. Phys. J. C **77**, 161 (2017), arXiv:1612.06764 [hep-ex] .

[196] B. Aubert *et al.* (BaBar), Phys. Rev. Lett. **102**, 091803 (2009), arXiv:0807.4119 [hep-ex] .

[197] J. P. Lees *et al.* (BaBar), Phys. Rev. D **86**, 032012 (2012), arXiv:1204.3933 [hep-ex] .

[198] J. P. Lees *et al.* (BaBar), Phys. Rev. Lett. **118**, 031802 (2017), arXiv:1605.09637 [hep-ex] .

[199] J. T. Wei *et al.* (Belle), Phys. Rev. Lett. **103**, 171801 (2009), arXiv:0904.0770 [hep-ex] .

[200] T. Aaltonen *et al.* (CDF), Phys. Rev. Lett. **107**, 201802 (2011), arXiv:1107.3753 [hep-ex] .

[201] R. Aaij *et al.* (LHCb), JHEP **07**, 133 (2012), arXiv:1205.3422 [hep-ex] .

[202] R. Aaij *et al.* (LHCb), JHEP **02**, 105 (2013), arXiv:1209.4284 [hep-ex] .

[203] R. Aaij *et al.* (LHCb), JHEP **05**, 082 (2014), arXiv:1403.8045 [hep-ex] .

[204] R. Aaij *et al.* (LHCb), Phys. Rev. Lett. **113**, 151601 (2014), arXiv:1406.6482 [hep-ex] .

[205] J. Charles, A. Le Yaouanc, L. Oliver, O. Pene, and J. C. Raynal, Phys. Rev. D **60**, 014001 (1999), arXiv:hep-ph/9812358 .

[206] A. Abada, D. Becirevic, P. Boucaud, G. Herdoiza, J. Leroy, A. Le Yaouanc, and O. Pene, JHEP **02**, 016 (2004), arXiv:hep-lat/0310050 .

[207] W. Detmold, C.-J. Lin, and S. Meinel, Phys. Rev. Lett. **108**, 172003 (2012), arXiv:1109.2480 [hep-lat] .

[208] F. Bernardoni, J. Bulava, M. Donnellan, and R. Sommer (ALPHA), Phys. Lett. B **740**, 278 (2015), arXiv:1404.6951 [hep-lat] .

[209] G. Martinelli, C. Pittori, C. T. Sachrajda, M. Testa, and A. Vladikas, Nucl. Phys. B **445**, 81 (1995), arXiv:hep-lat/9411010 .

[210] G. 't Hooft, Nuclear Physics B **61**, 455 (1973).

[211] S. Weinberg, Phys. Rev. D **8**, 3497 (1973).

[212] V. Lubicz, L. Riggio, G. Salerno, S. Simula, and C. Tarantino (ETM), Phys. Rev. D **98**, 014516 (2018), arXiv:1803.04807 [hep-lat] .

[213] N. Gubernari, A. Kokulu, and D. van Dyk, JHEP **01**, 150 (2019), arXiv:1811.00983 [hep-ph] .

[214] L. J. Cooper, C. T. H. Davies, and M. Wingate (HPQCD), Phys. Rev. D **105**, 014503 (2022), arXiv:2108.11242 [hep-lat] .

[215] D. Becirevic, N. Kosnik, F. Mescia, and E. Schneider, Phys. Rev. D **86**, 034034 (2012), arXiv:1205.5811 [hep-ph] .

[216] P. Zyla *et al.* (Particle Data Group), PTEP **2020**, 083C01 (2020).

[217] D. Hatton, C. T. H. Davies, J. Koponen, G. P. Lepage, and A. T. Lytle, Phys. Rev. D **103**, 114508 (2021), arXiv:2102.09609 [hep-lat] .

[218] A. Bazavov *et al.* (Fermilab Lattice, MILC), Phys. Rev. D **90**, 074509 (2014), arXiv:1407.3772 [hep-lat] .

[219] D. Hatton, C. T. H. Davies, B. Galloway, J. Koponen, G. P. Lepage, and A. T. Lytle (HPQCD), Phys. Rev. D **102**, 054511 (2020), arXiv:2005.01845 [hep-lat] .

[220] M. Beneke, G. Buchalla, M. Neubert, and C. T. Sachrajda, Eur. Phys. J. C **61**, 439 (2009), arXiv:0902.4446 [hep-ph] .

[221] K. Melnikov and T. v. Ritbergen, Phys. Lett. B **482**, 99 (2000), arXiv:hep-ph/9912391 .

[222] M. Beneke and V. M. Braun, Nucl. Phys. B **426**, 301 (1994), arXiv:hep-ph/9402364 .

[223] M. Beneke, T. Feldmann, and D. Seidel, Nucl. Phys. B **612**, 25 (2001), arXiv:hep-ph/0106067 .

[224] C. Greub, V. Pilipp, and C. Schupbach, JHEP **12**, 040 (2008), arXiv:0810.4077 [hep-ph] .

[225] B. Grinstein and D. Pirjol, Phys. Rev. D **70**, 114005 (2004), arXiv:hep-ph/0404250 .

[226] C. Bobeth, G. Hiller, and D. van Dyk, JHEP **07**, 098 (2010), arXiv:1006.5013 [hep-ph] .

[227] M. Beylich, G. Buchalla, and T. Feldmann, Eur. Phys. J. C **71**, 1635 (2011), arXiv:1101.5118 [hep-ph] .

[228] C. Bobeth, G. Hiller, and D. van Dyk, JHEP **07**, 067 (2011), arXiv:1105.0376 [hep-ph] .

[229] M. Beneke and T. Feldmann, Nucl. Phys. B **592**, 3 (2001), arXiv:hep-ph/0008255 .

[230] H. H. Asatrian, H. M. Asatrian, C. Greub, and M. Walker, Phys. Lett. B **507**, 162 (2001), arXiv:hep-ph/0103087 .

[231] H. H. Asatryan, H. M. Asatrian, C. Greub, and M. Walker, Phys. Rev. D **65**, 074004 (2002), arXiv:hep-ph/0109140 .

[232] H. M. Asatrian, K. Bieri, C. Greub, and M. Walker, Phys. Rev. D **69**, 074007 (2004), arXiv:hep-ph/0312063 .

[233] D. Seidel, Phys. Rev. D **70**, 094038 (2004), arXiv:hep-ph/0403185 .

[234] M. Beneke, T. Feldmann, and D. Seidel, Eur. Phys. J. C **41**, 173 (2005), arXiv:hep-ph/0412400 .

[235] V. M. Braun *et al.*, Phys. Rev. D **74**, 074501 (2006), arXiv:hep-lat/0606012 .

[236] G. Isidori, S. Nabeebaccus, and R. Zwicky, JHEP **12**, 104 (2020), arXiv:2009.00929 [hep-ph] .

[237] G. Isidori, D. Lancierini, S. Nabeebaccus, and R. Zwicky, (2022), arXiv:2205.08635 [hep-ph] .

[238] N. Davidson, T. Przedzinski, and Z. Was, Comput. Phys. Commun. **199**, 86 (2016), arXiv:1011.0937 [hep-ph] .

[239] M. Bordone, G. Isidori, and A. Pattori, Eur. Phys. J. C **76**, 440 (2016), arXiv:1605.07633 [hep-ph] .

[240] S. Descotes-Genon, S. Fajfer, J. F. Kamenik, and M. Novoa-Brunet, Phys. Lett. B **809**, 135769 (2020), arXiv:2005.03734 [hep-ph] .

[241] S. Halder (Belle-II), (2021), arXiv:2101.11573 [hep-ex] .

[242] W. Altmannshofer, A. J. Buras, D. M. Straub, and M. Wick, JHEP **04**, 022 (2009), arXiv:0902.0160 [hep-ph] .

[243] A. J. Buras, J. Girrbach-Noe, C. Niehoff, and D. M. Straub, JHEP **02**, 184 (2015), arXiv:1409.4557 [hep-ph] .

[244] J. Brod, M. Gorbahn, and E. Stamou, Phys. Rev. D **83**, 034030 (2011), arXiv:1009.0947 [hep-ph] .

[245] J. M. Flynn, T. Izubuchi, T. Kawanai, C. Lehner, A. Soni, R. S. Van de Water, and O. Witzel, Phys. Rev. D **91**, 074510 (2015), arXiv:1501.05373 [hep-lat] .

[246] M. Moulson, in *8th International Workshop on the CKM Unitarity Triangle* (2014) arXiv:1411.5252 [hep-ex] .

[247] J. P. Lees *et al.* (BaBar), Phys. Rev. D **87**, 112005 (2013), arXiv:1303.7465 [hep-ex] .

[248] F. Abudinén *et al.* (Belle-II), Phys. Rev. Lett. **127**, 181802 (2021), arXiv:2104.12624 [hep-ex] .