Heavy Semileptonics from Lattice (NR)QCD

Euan McLean

February 5, 2018

Contents

1 Introduction

here is ma plan.

- *phenomenology (general overview of the virtual fronteer, semileptonic decays, CKM determination,lepton flavour violation).*
- *lattice qcd: build up picture of exactly how 2- and 3-point correlators are built with HISQ. then build NRQCD on that.*
- *NRQCD Axial Current Renormalization*
- *$B \to \pi$ finite volume effects*
- *$B_s - D_s$ heavy hisq*
- *$B_{(s)} - D_{(s)}$ NRQCD-HISQ*
- *$B_c \to \eta_c$, combination of all 3.*

2 Semileptonic Decays, specifically $B_s \to D_s l\nu$

The key motivations for studying decays like the $B_s \to D_s l\nu$ are:

- Determination of CKM matrix elements. Specifically, by combining the theoretical prediction and an observed brancing fraction of $B_s \to D_s l\nu$, $|V_{cb}|$ can be extracted.
- Precision tests of the standard model. There are currently a number of tensons between experiment and the standard model predictions of B decays, some closely related to $B_s \to D_s l\nu$ [28].

We will first discuss the general ideas in computing semileptonic decays.

2.1 Semileptonic Decays

Semileptonic decays are useful for studying the sector of the standard model (SM) which couples quarks to the weak force [31]:

$$\mathscr{L}_W = \frac{g}{\sqrt{2}} \left[\quad V_{ij} J^{\mu}_{ij} W^{+}_{\mu} + V^{\dagger}_{ij} J^{\mu\dagger}_{ij} W^{-}_{\mu} \quad \right] \tag{2.1}$$

$J^{ij}_{\mu} = \bar{u}_i \gamma_\mu \frac{1}{2}(1-\gamma_5) d_j$ are the weak currents, $\underline{u} = (u, c, t)$ and $\underline{d} = (d, s, b)$ are the quark fields, g is the weak coupling constant, $W^{\pm}$ are the charged weak bosons, and V is the (unitary) Cabibbo-Kobayashi-Maskawa (CKM) matrix. V_{ij} can be thought of as a matrix of couplings which dictate the probability of mixing between two quark flavours, for example the amplitude of a b decaying to a u (and emitting a W) is proportional to $V_{13} \equiv V_{ub}$.

A semileptonic decay is any $hardon \to hadron + leptons$ process. Fig. 2 shows the tree level (in weak coupling) contribution to a semileptonic decay. We denote a meson with valence quark content

q and q' as $M_{qq'}$. The $M_{q_1\bar{q}_3} \to M_{q_2\bar{q}_3} l\bar{\nu}_l$ decay (l is some charged lepton and $\bar{\nu}_l$ is it's neutrino) is proportional to $V_{q_1q_2}$ at tree level.

This is given by

$$\mathcal{M} = \left(\frac{ig}{\sqrt{2}}\right)^2 V_{q_1q_2} \langle M_{q_2\bar{q}_3}, l\bar{\nu}| J_\alpha^{q_1\bar{q}_2} D_W^{\alpha\beta}(p^2) J_\beta^{l\bar{\nu}} |M_{q_2\bar{q}_3}\rangle \tag{2.2}$$

$J^{l\bar{\nu}}$ is the analog of the quark currents $J^{q_1\bar{q}_2}$, since the W couples in the same way to leptons (with V replaced by a unit matrix). If the external momenta of the process p^2 are much smaller than the W mass, one can remove the W propagator from the tree level amplitude [6];

$$\left(\frac{ig}{\sqrt{2}}\right)^2 D_W^{\mu\nu}(p^2) = \left(\frac{ig}{\sqrt{2}}\right)^2 \left(\frac{-ig^{\mu\nu}}{p^2 - M_W^2}\right) = \underbrace{\frac{i}{M_W^2}\left(\frac{ig}{\sqrt{2}}\right)^2}_{\equiv -2\sqrt{2}G_F} g^{\mu\nu} + \mathcal{O}\left(\frac{p^2}{M_W^4}\right) \tag{2.3}$$

Then $\mathcal{M}$ can be factorised;

$$\begin{aligned}
\mathcal{M} &\simeq -2\sqrt{2} G_F V_{q_1q_2} \langle M_{q_1\bar{q}_3}, l\bar{\nu}| J_\mu^{q_1\bar{q}_2} J^{l\bar{\nu}\mu} |M_{q_2\bar{q}_3}\rangle \\
&= -2\sqrt{2} G_F V_{q_1q_2} \langle l\bar{\nu}| J^{l\bar{\nu}\mu} |0\rangle \langle M_{q_1\bar{q}_3}| J_\mu^{q_1\bar{q}_2} |M_{q_2\bar{q}_2}\rangle \\
&\equiv -2\sqrt{2} V_{q_1q_2} G_F L^\mu H_\mu.
\end{aligned} \tag{2.4}$$

L_μ can be computed in perturbation theory. The hadronic matrix element H_μ however, due to the non-perturbative nature of QCD at low energies, cannot be computed analytically. It is quantities such as H_μ that we wish to calculate in lattice QCD.

2.2 CKM Extraction

Speaking heuristically, if one can calculate the quantity $G_F L^\mu H_\mu$ from theory, and measures the amplitude of the process $\mathcal{M}$, they could then rearrange equation (2.4) to deduce $V_{q_1q_2}$. Speaking more practically, the relevant equation is, for example, for the $B_s \to D_s l\nu$ decay [28]:

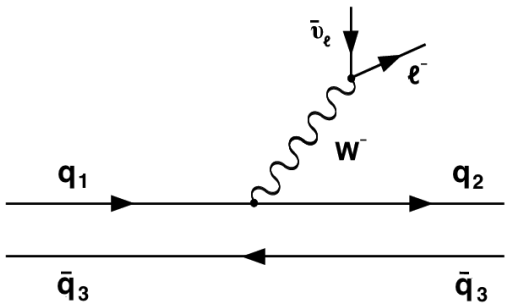

Figure 2.1: Semileptonic decay at tree level

$$\frac{d\Gamma}{dq^2} = \eta_{\text{EW}} \frac{G_F^2 |V_{cb}|^2}{24\pi^3 M_{B_s}^2} \left(1 - \frac{m_l^2}{q^2}\right)^2 |\underline{p}| \tag{2.5}$$

$$\times \left[\left(1 + \frac{m_l^2}{2q^2}\right) M_{B_s}^2 |\underline{p}| f_+^2(q^2) + \frac{3m_l^2}{8q^2}(M_{B_s}^2 - M_{D_s}^2)^2 f_0^2(q^2)\right] \tag{2.6}$$

where m_l is the mass of the charged lepton, η_{EW} is an electroweak correction factor, q^2 is the momentum transfer, $d\Gamma/dq^2$ is the branching fraction, and $f_{0,+}(q^2)$ are the form factors associated with the decay, to be defined in 2.4. Given experimental data for $d\Gamma/dq^2$ and theoretical data for $f_{0,+}(q^2)$, one can deduce $|V_{cb}|$.

What is the value of determining CKM elements? Firstly, uncertainty in CKM elements are the dominant error in many standard model predictions. V elements can also test the standard model. V is unitary by definition. However, if there were more than 3 generations of quark, requiring V to be 4x4 or larger, the 3x3 submatrix which couples the 3 known generations would not itself be unitary in general ([32] ch. 29). Therefore, if one can deduce the elements of the 3x3 V to high enough precision to show it is not unitary, this would be indirect evidence for new physics.

Unitarity imposes a constraint on each row of V. The current status of the unitarity of V can be read off from below (each quantity should be zero for unitarity to hold)[3]:

$$\begin{aligned} |V_{ud}|^2 + |V_{us}|^2 + |V_{ub}|^2 - 1 &= -0.020(9) \\ |V_{cd}|^2 + |V_{cs}|^2 + |V_{cb}|^2 - 1 &= 0.06(3) \end{aligned} \tag{2.7}$$

both of the above relations display a $\sim 2\sigma$ deviation from unitarity. More precise values of these parameters are needed to discover if CKM is infact non-unitary.

Now we consider $|V_{cb}|$ specifically. $|V_{cb}|$ is the dominant uncertainty in many standard model predictions of rare decays, such as $B_s \to \mu^+\mu^-$,$K \to \pi\nu\bar{\nu}$, and the CP violation parameter ϵ_K [28]. The FLAG working group quotes their average $|V_{cb}|$ to currently be [3]:

$$|V_{cb}| = 0.04085(95) \tag{2.8}$$

However, the water is slightly muddied by the presence of tensions between different determinations of $|V_{cb}|$, namely between that deduced from studying $B \to D^* l\nu$, and an inclusive analysis of $B \to X_c l\nu$ where X_c is any meson containing a c quark [28]:

$$|V_{cb}|_{\text{inclusive}} = 0.04221(78) \ , \ |V_{cb}|^{B\to D^* l\nu} = 0.03904(49)_{\text{expt.}}(53)_{\text{QCD}}(19)_{\text{QED}} \tag{2.9}$$

A recent review of $|V_{cb}|$ determinations from the lattice is given in [33]. This issue would benefit from a $|V_{cb}|$ determination from annother channel, like $B_s \to D_s l\nu$, to flesh out the landscape of $|V_{cb}|$ values and pinpoint the source of the tension, if there is any.

2.3 Lepton Flavour Violation

There are a number of tensions currently between theory and experiment in B decays, who's true nature can be eludicated by our $B_s \to D_s l\nu$ study, and eventual $B \to D l\nu$ study.

Define the ratios

$$R(X) = \frac{\mathcal{B}(B \to X\tau\nu_\tau)}{\mathcal{B}(B \to X l\nu_l)} \tag{2.10}$$

where $l = e$ or μ. These can be computed either purely from a lattice calculation, or purely from experimental data, independantly of the associated CKM element. It therefore can be used for comparison between experiment and the standard model. $R(D^*)$ contains a 2.1σ discrepancy [1]:

$$R(D^*)|_{\text{SM}} = 0.252(3) \ , \ R(D^*)|_{\text{LHCb}} = 0.336(27)_{sys}(30)_{stat} \tag{2.11}$$

The issue is similarly present for $R(D)$ [26]:

$$R(D)|_{\text{SM}} = 0.299(7) \ , \ R(D)|_{\text{exp}} = 0.391(28)_{sys}(41)_{stat} \tag{2.12}$$

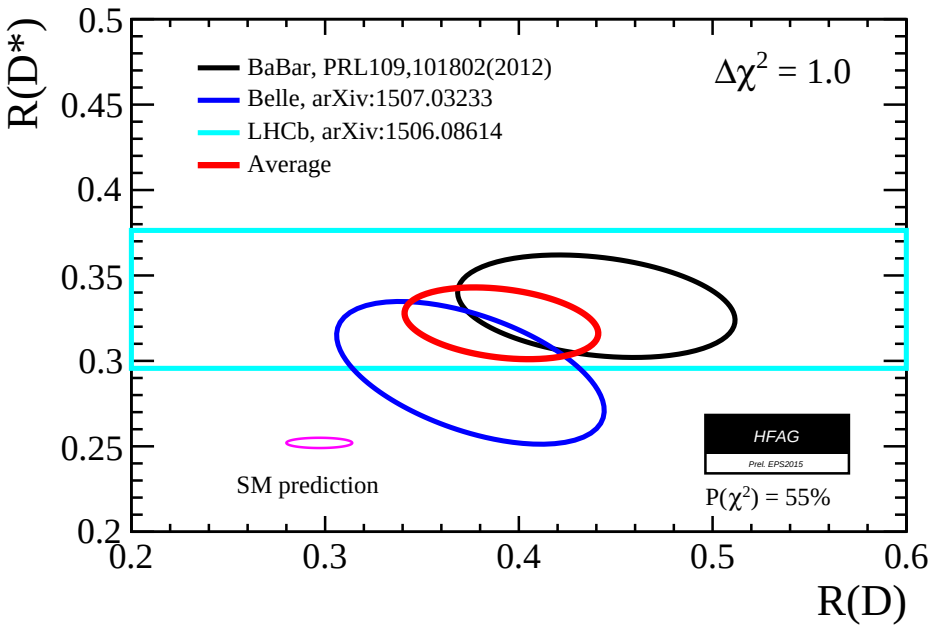

Figure 2: $R(D)/R(D^*)$ determinations from standard model and experiment [18]

where in this case $R(D)|_{\text{expt}}$ is a world average of experimental values. Our eventual study of $B \to Dl\nu$ will produce a new standard model determination of $R(D)$, helping shed light on the issue.

A natural question to ask is: is there an analagous tension in related channels, like $B_s \to D_s l\nu$? There is not yet any experimental data for this decay, but the current best standard model determination is given by $\mathcal{B}(B_s \to D_s \tau\nu_\tau)/\mathcal{B}(B_s \to D_s l\nu_l)|_{\text{SM}} = 0.314(6)$.

Besides these, there are also tensions in the quantitites [2]:

$$R'(K^{(*)}) = \frac{\mathcal{B}(B \to K^{(*)}\mu^+\mu^-)}{\mathcal{B}(B \to K^{(*)}e^+e^-)} \tag{2.13}$$

All of the above anomalies are suggestive of lepton flavour violating effects. Various BSM models have been suggested; hot topics include Leptoquarks, Z' models, and partial compositeness [2].

2.4 Kinematics and Form Factors

We now turn to some more technical details about semileptonic decays.

We will refer to the 4-momenta and mass of the initial and final states as $p_{1,2}, M_{1,2}$ respectively, and define the 4-momentum taken away by the $W^\pm$ boson as $q \equiv p_2 - p_1$. We work in the rest frame of the initial meson, in which

$$q^2 = M_1^2 + M_2^2 - 2M_1E_2. \tag{2.14}$$

There is a physically allowed range of values for an on-shell q^2. The minimum is when all of the 3-momentum of the initial state is taken by the final meson, $q^2_{\text{min}} = 0$. q^2 is maximised when all of the

3-momentum is taken by the boson, $\underline{p}_2^2 = 0 \to E_2 = \sqrt{M_2^2 + \underline{p}_2^2} = M_2 \to q^2_{\max} = M_1^2 - M_2^2 - 2M_1M_2 = (M_2 - M_1)^2$. So

$$0 \leq q^2 \leq (M_2 - M_1)^2 \tag{2.15}$$

This also creates an allowed range for the final meson 3-momentum:

$$0 < \underline{p}_2^2 < \left(\frac{M_1^2 + M_2^2}{2M_1}\right)^2 - M_2^2 \tag{2.16}$$

Hadronic matrix elements like H_μ from sec. 2.1 can be parameterised in terms of form factors. The current operator between between the states is a conserved current, so then the matrix element must be proportional to only conserved quantities, namely, elements of the stress-energy tensor. Lorentz invariance requires indices on either side of such a relation match, so a matrix element with a single Lorentz index can only be proportional to 4-momenta.

Let us consider the case where the two states in the matrix element are pseudoscalar mesons (as in $B_s \to D_s l\nu$), with an insertion of a left-handed current (i.e. the coupling to the $W^\pm$ in (2.1)), $J_\mu^{ij} = \bar{u}_i \gamma_\mu \frac{1}{2}(1 - \gamma_5) d_j$. This can be written as $J_\mu^{ij} = V_\mu^{ij} - A_\mu^{ij}$, these are the vector and axial vector currents.

The axial vector evaluated between two pseudoscalar states must vanish because the combination is not parity invariant thus does not contribute in pure QCD, leaving just the vector current. The most common parameterisation is:

$$\langle P_2(p_2)|V^\mu|P_1(p_1)\rangle = f_+(q^2)\left[p_1^\mu + p_2^\mu - \frac{M_1^2 - M_2^2}{q^2}q^\mu\right] + f_0(q^2)\frac{M_1^2 - M_2^2}{q^2}q^\mu \tag{2.17}$$

Where $|P_i(p_i)\rangle$ is a pseudoscalar meson state with momentum p_i.

3 Lattice QCD

3.1 Discretization

At low energies (∼200MeV and below), QCD becomes non-perturbative, in other words, the coupling α_s becomes $\mathcal{O}(1)$, and an expansion in α_s (as in perturbation theory) will not be dominated by the leading orders [32]. We require an alternative.

The expectation value of an observable $\mathcal{O}$ in a Yang-Mills theory can be expressed as a Euclidean path integral [22];

$$\langle \mathcal{O} \rangle = \int \mathcal{D}A\mathcal{D}\psi\mathcal{D}\bar{\psi}\mathcal{O}e^{-S[A,\psi,\bar{\psi}]}, \tag{3.1}$$

where A is a gauge field, $\psi(\bar{\psi})$ is an (anti)fermion field, S is the Euclideanised classical action, and $\mathcal{D}$ denotes integration over all configurations of a field. "Euclideanised" refers to a Wick rotation $t \to it$ in S. In the perturbative approach, we would expand exp(−interacting part of S) resulting in a power series in the gauge coupling populated by Feynman diagrams.

The other option is to instead carry out the integral directly. This can only be done numerically. Since it's not numerically feasible to carry out an infinite number of integrals, one must approximate spacetime as a discrete 4 dimensional lattice with spacing "a" between lattice sites, finite spacial volume L_x^3 and finite temporal extent L_t. The functional integral becomes

$$\int \mathcal{D}A\mathcal{D}\psi\mathcal{D}\bar{\psi} = \prod_n \int dU(x_n)d\psi(x_n)d\bar{\psi}(x_n), \tag{3.2}$$

where n is a 4-vector with integer components labelling the sites, and $x_n^\mu = an^\mu$. This has a second benefit which is to naturally regularize the theory with a momentum cutoff $\Lambda \sim 1/a$ [22]. The gauge field has been replaced with the gauge "link":

$$U_\mu(x) \equiv \exp\left(igaA_\mu\left(x + \frac{a\hat{\mu}}{2}\right)\right) \in SU(N_c), \tag{3.3}$$

g is the gauge coupling, $\hat{\mu}$ is a unit vector in the μ direction. Parameterizing the gauge fields this way is motivated by the geometrical interpretation of Yang-Mills theories on discrete spacetime and the requirement of exact gauge symmetry [27].

3.2 Meson Correlation Functions

A typical quantity that is computed on the lattice is a meson correlation function, i.e. when $\mathcal{O} = \Phi(x)\Phi^\dagger(y)$ and Φ is a meson creation operator. This is a good working example for showing the steps in a lattice calculation.

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

$$C(x,y) = \int \mathcal{D}\psi\mathcal{D}\bar{\psi}\mathcal{D}U \left(\bar{b}(x)\gamma_5 d(x)\bar{d}(y)\gamma_5 b(y) \right) e^{-S_G[U] - \sum_{w,z}\bar{\psi}(w)M(w,z)[U]\psi(z)} \tag{3.4}$$

where we have now broken the action up into a gauge part $S_G[U]$, and a fermion part. $M(x,y)[U]$ is the Dirac operator, and can be seen as a matrix in lattice site, flavor, color and spin. ψ is a vector of quark flavours.

The integral over fermions can be preformed analytically, since the fermion fields are Grassman valued. In our example, the result is [30],

$$C(x,y) = \int \mathcal{D}U \ \mathrm{Tr}\left[M_b^{-1}(y,x)[U]\gamma_5 M_d^{-1}(x,y)[U]\gamma_5 \right] e^{-S_G[U]}\det(M[U]) \tag{3.5}$$

The quantinties $M_f^{-1}(x,y)[U]$ are propagators of a quark of flavour f on a fixed gauge background U. The integration over gauge fields is generally carried out by an importance sampling method. A finite *ensemble* of gauge configurations $\{U_i\}$ is generated by a Monte Carlo Markov chain (MCMC), where the probability of a gauge configuration U_j being added to the ensemble is proportional to

$$e^{-S_G[U_j]}\det(M[U]) \tag{3.6}$$

See [11] ch. 7 for examples of such algorithms. In the case of our work, we use ensembles generated by the MILC collaboration [5].

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

$$C(x,y) \simeq \frac{1}{N}\sum_i \text{ Tr } [\ M_b^{-1}(y,x)[U_i]\gamma_5 M_d^{-1}(x,y)[U_i]\gamma_5\ ] \tag{3.7}$$

where N is the size of the ensemble. The calculation of the correlation function then is split into 3 steps:

1. Generate an ensemble of Gauge configurations $\{U_i\}$ by MCMC.
2. Compute quark propagators $M_f^{-1}(x,y)[U]$ on each Gauge configuration. This requires inverting the matrix M each time, this is typically done by conjugate gradient method.
3. Construct trace as in (3.7), and average over the ensemble.

We now turn to the issue of choosing lattice actions.

3.3 Gauge Action

One must choose a discrete version of S which becomes the original Yang-Mills action in the limit $a \to 0$. There are many choices of S which satisfy this. The simplest action with a continuum limit coinciding with an $SU(N_c)$ gauge theory is

$$S_G = \beta \sum_{x,\mu,\nu} \left[N_c - \text{Tr}\left(U_\mu(x)U_\nu(x+a\hat{\mu})U_\mu^\dagger(x+a\hat{\nu})U_\nu^\dagger(x)\right)\right]. \tag{3.8}$$

The trace is over colour, and $\beta = 2(L_x/a)/g^2$. The product of gauge links $U_\mu(x)U_\nu(x+a\hat{\mu})U_\mu^\dagger(x+a\hat{\nu})U_\nu^\dagger(x)$ is known as the elementary plaquette, and is the lattice analogue of the field strenght tensor $F_{\mu\nu}$. In practice this action is appended with irrelevant operators which act to cancel discretisation effects in the results of calculations, these are known as Symmanzik improvements ([11] ch.10). For example, in the MILC ensembles, the Gauge action used is a generalization of the above, the Luscher-Weisz action [15].

3.4 Fermion Actions

The interacting Dirac action is most naively discretised with

$$S_F = \sum_{x,\mu} \bar{\psi}(x)\gamma_\mu \nabla_\mu \psi(x) + m\sum_x \bar{\psi}(x)\psi(x), \tag{3.9}$$

where ∇_μ is the gauge covariant finite difference operator,

$$\nabla_\mu \psi(x) = \frac{1}{2a}(U_\mu(x)\psi(x+a\hat{\mu}) - U_\mu^\dagger(x-a\hat{\mu})\psi(x-a\hat{\mu})). \tag{3.10}$$

In appendix B we describe the doubling problem. This is the observation that the propagator for a fermion obeying (B.1), $M^{-1}(k)$ has the property

$$M^{-1}(k+\frac{\pi}{a}\zeta) = \gamma_{5\mu} M^{-1}(k)\gamma_{5\mu} \tag{3.11}$$

For 16 4-vectors $\zeta_\mu \in \mathbb{Z}_2$. This leads to 16 poles in the fermion specturm, therefore 16 distinct excitations (called *tastes*). We require a way of removing the 15 unphysical excitations.

3.4.1 Staggered Quarks

There are a number of solutions to this problem. The most straightforward is to modify the action to push the mass of the unwanted tastes above the momentum cutoff, preventing it from effecting the dynamics (" *Wilson fermions*")([11] ch.6.2). However, actions of this type explicity break Chiral symmetry. Among other issues, this causes additive renormalization of the fermion mass, immensely complicating the renormalization procedure.

Annother approach, known as *staggered quarks*([11] ch.6.3), partially resolves the doubling issue while retaining chiral symmetry. This is the method we use in our study. Other notable approaches besides Wilson and staggered quarks include *domain wall* [21] and *overlap* [29] fermions.

The general idea of staggered fermions is the following. Redefine the fields according to

$$\psi(x) = \prod_{\mu} (\gamma_\mu)^{x_\mu/a} \chi(x) \equiv \Omega(x)\chi(x) \tag{3.12}$$

In terms of the new spinor variables $\chi(x)$, the naive action (B.1) becomes

$$S_F = \bar{\chi}(x)[\alpha_\mu(x)\nabla_\mu + m]\chi(x) \tag{3.13}$$

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

$$M_\chi^{-1}(x,y)[U] = s(x,y)[U]\ 1_{\text{spin}} \tag{3.14}$$

One need only to include a single component of χ in a simulation (i.e. fix $\chi = (\chi_1, 0, 0, 0)$). Then they can compute $M_\chi^{-1}(x,y)[U]$ to obtain $s(x,y)$. Then, using the inverse of (3.12), $s(x,y)$ can be transformed to a propagator of the original spinors:

$$M_\psi^{-1}(x,y)[U] = s(x,y)[U]\Omega(x)\Omega^\dagger(y) \tag{3.15}$$

This is clearly computationally beneficial. But also, by shaving off the other spinor components, one reduces the number propagating degrees of freedom by a factor of 4. This cuts the number of tastes from 16 down to 4 (this is shown in appendix B).

The remaining multipilcity is tacked in 3 steps:

1. Ensure only one taste is created and destroyed in the propagator.
2. Minimize the interaction between tastes by modifying the action.
3. Remove contributions of extra tastes in the sea by taking $\det M \to \sqrt[4]{\det M}$ in (3.6).

Step 3 can be justified by the following - in the $a \to 0$ limit, $\det M$ tends to $(\det M^{(0)})^4$, where $M^{(0)}$ is the Dirac oprator for a single taste. Then, taking the 4th root (in principle) reduces the determinant to that of a sea containing 1 taste.

3.4.2 Criticisms of Rooting

3.4.3 Highly Improved Staggered Quarks

Step 2 above is the guiding principle for the action we use in our study, the Highly Improved Staggered Quark action (HISQ).

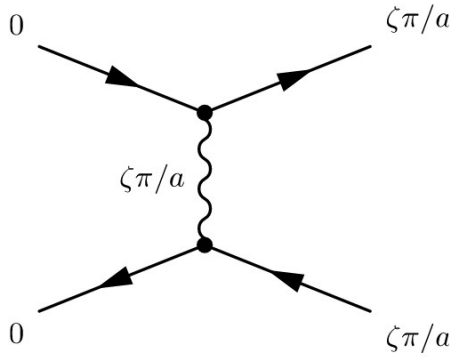

Figure 3: Taste mixing at tree level.

Interaction between different tastes ("taste mixing") is dominated by the process in fig. 3. In HISQ, this is suppressed by modifying the gauge fields in such a way as to minimize the coupling between a gluon with momentum π/a and the fermions, in other words, minimize the vertices in fig. 3. To this end, one can change the action so that fermions only couple to *smeared* gauge links, in which high frequency exitations have been removed.

Define the first and second covariant derivative operators:

$$\begin{aligned}\delta_\rho U_\mu(x) \equiv \frac{1}{a}\big(&U_\rho(x)U_\mu(x+a\hat{\rho})U_\rho^\dagger(x+a\hat{\mu}) \\ &- U_\rho^\dagger(x-a\hat{\rho})U_\mu(x-a\hat{\rho})U_\rho(x-a\hat{\rho}+a\hat{\mu})\big) \qquad (3.16)\\ \delta_\rho^{(2)} U_\mu(x) \equiv \frac{1}{a^2}\big(&U_\rho(x+a\hat{\rho})U_\rho^\dagger(x+a\hat{\mu}) \\ &- 2U_\mu(x) \\ &+ U_\rho^\dagger(x-a\hat{\rho})U_\mu(x-a\hat{\rho})U_\rho(x-a\hat{\rho}+a\hat{\mu})\big) \qquad (3.17)\end{aligned}$$

With this we can define the smearing operator;

$$\mathcal{F}_\mu = \prod_{\rho\neq\mu}\left(1+\frac{a^2\delta_\rho^{(2)}}{4}\right) \qquad (3.18)$$

HISQ uses two different smeared gauge fields defined by;

$$X_\mu(x) \equiv \mathcal{U}\mathcal{F}_\mu U_\mu(x) \qquad (3.19)$$

$$W_\mu(x) \equiv \left(\mathcal{F}_\mu - \sum_{\rho\neq\mu}\frac{a^2(\delta_\rho)^2}{2}\right)\mathcal{U}\mathcal{F}_\mu U_\mu(x) \qquad (3.20)$$

where $\mathcal{U}$ is a re-unitarization operator. The HISQ action can then be written as:

$$S_{\text{HISQ}} = \sum_x \bar{\psi}(x)\left(\gamma\cdot\left(\nabla_\mu(W) - \frac{a^2}{6}(1+\epsilon_{\text{Naik}})\nabla_\mu^3(X)\right)+m\right)\psi(x) \qquad (3.21)$$

Where $\nabla_\mu(Z)$ is the covariant derivative (B.2) with gauge links repaced with Z. This action in fact not only removes tree level interactions like fig. 3, but also all taste mixing interactions at 1-loop. The ∇_μ^3 term is introduced to remove "lattice artifacts", i.e. it reduces the size of $\mathcal{O}(a)$ terms in the $a \to 0$ limit of the action. In the same spirit, ϵ_{Naik} is fixed according to the constraint

$$\lim_{\underline{p}\to 0} \frac{E^2(\underline{p}) - m^2}{\underline{p}^2} = 1 \tag{3.22}$$

where $E(\underline{p})$ obeys the dispersion relation from HISQ. The motivation for the specific smearing of the gauge fields (3.20), and more details on HISQ in general, are given in [14].

3.4.4 NRQCD

The HISQ action is suitable for fermions in the mass range of up, down, strange and charm masses. Bottom quarks are an issue, since their mass is typically higher than the cutoff $1/a$ for computationally viable lattices, so it's dynamics cannot naively be resolved on the lattice.

The root of the problem is in the relativistic nature of the b quark. Consider the expansion in momentum $\underline{p}^2$ of the continuum relativistic dispersion relation:

$$\omega = \sqrt{\underline{p}^2 + m^2} \simeq m + \frac{\underline{p}^2}{2m} - \frac{\underline{p}^4}{4m^3} + ... \tag{3.23}$$

the first term (rest mass) is the source of the issue, when $m > 1/a$ the first term pushes the frequency of excitations ω over $1/a$ and it's wavelength becomes hidden between the lattice points.

The solution we employ is to persue instead a non-relativistic formulation, called NRQCD [24]. The leading order Lagrangian in the continuum is ([11] ch. 6.6)

$$\mathscr{L}^0_{NRQCD} = \psi_+^\dagger \left[i\partial_0 + \frac{\underline{\nabla}^2}{2m_b} \right] \psi_+. \tag{3.24}$$

ψ_+ is the first two components of a Dirac spinor, the second two components (the antiparticle) are not present since the dispersion relation from this Lagrangian has no antiparticle solutions. m_b is the bare mass for the b quark. The NRQCD action reproduces (3.23) with the first term chopped out.

Correction terms are chosen to be all gauge-invariant terms, grouped according to powers of the quark velocity v. For an example of deducing such powers: kinetic enegry $= m_b v^2 = \int d^3x \psi_+^\dagger \frac{\underline{\nabla}^2}{2m_b} \psi_+ \to |\underline{\nabla}|/m_b \sim v$.

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

The form of the action allows propagators $M^{-1}[U]$ to be computed using a simple recursion relation

$$M_b^{-1}(\underline{x}, t+1)[U] = e^{-aH}[U] M_b^{-1}(\underline{x}, t)[U] \tag{3.25}$$

which is numerically very fast. In the interest of numerical stability, the time evolution operator is re-cast as

$$e^{-aH} = \left(1 - \frac{a\delta H}{2}\right) \left(1 - \frac{aH_0}{2n}\right)^n U_0^\dagger(\underline{x}, t) \left(1 - \frac{aH_0}{2n}\right)^n \left(1 - \frac{a\delta H}{2}\right) \tag{3.26}$$

where n is an arbitrary integer (chosen in our study to be $n = 4$), and the Hamiltonian has been broken up into a leading part H_0 and correction δH. We use the $\mathcal{O}(\alpha_s v^4)$ corrected NRQCD Hamiltionian:

$$aH_0 = -\frac{\nabla^{(2)}}{2am_b}, \tag{3.27}$$

$$\begin{aligned} a\delta H = &- c_1 \frac{(\nabla^{(2)})^2}{8(am_b)^3} + c_2 \frac{i}{8(am_b)^2}(\underline{\nabla} \cdot \tilde{\underline{E}} - \tilde{\underline{E}} \cdot \underline{\nabla}) \\ &- c_3 \frac{1}{8(am_b)^2} \sigma \cdot (\underline{\nabla} \times \tilde{\underline{E}} - \tilde{\underline{E}} \times \underline{\nabla}) \\ &- c_4 \frac{1}{2am_b} \sigma \cdot \tilde{\underline{B}} + c_5 \frac{\nabla^{(4)}}{24am_b} \\ &- c_6 \frac{(\nabla^{(2)})^2}{16n(am_b)^2} \end{aligned} \tag{3.28}$$

where $\nabla^{(2)}$ is the second lattice derivative, $\tilde{\underline{E}}$ and $\tilde{\underline{B}}$ are the chromoelectric and chromomagnetic fields, the expression for these in terms of gauge links are given in [17]. The coefficients $\{c_i\}$ are fixed by matching lattice NRQCD to continuum QCD up to 1-loop (thus is the quoted $\mathcal{O}(\alpha_s)$ correction).

Once the propagator for the 2-component non-relativistic b quark has been found, it must be transformed back into a 4-component spinor. This is done by the inverse Fouldy-Wouthuysen transformation [13]:

$$\psi(x) = e^{-\frac{\gamma \cdot \nabla}{2m_b}} \begin{pmatrix} \psi_+(x) \\ 0 \end{pmatrix} \tag{3.29}$$

3.5 Correlation Functions from HISQ propagators and random walls

3.5.1 Preliminaries

The full set of spin-mixing matrices can be labelled according to

$$\gamma_n = \prod_\mu (\gamma_\mu)^{n_\mu} \quad n_\mu = \mathbb{Z}_2 \tag{3.30}$$

There are 16 such matrices representing corners of the hypercube. As $\gamma_\mu^2 = 1$, one can also use a general site vector x_μ to label the matrix, then $\gamma_x = \gamma_n$ where $n_\mu = x_\mu \mod 2$. One can show that for any x; $\gamma_x^\dagger \gamma_x = 1$.

Naive quarks $\psi(x)$ can be transformed into staggered quarks $\chi(x)$ via $\psi(x) = \gamma_x \chi(x)$. Then, Naive quark propagators (inverse Dirac operators) become

$$G_\psi(x, y) = \gamma_x \gamma_y^\dagger G_\psi(x, y) \tag{3.31}$$

By conjugating both sides and using γ_5-hermiticity $G_\psi^\dagger(y, x) = \gamma_5 G_\psi(y, x) \gamma_5$ it can be shown that

$$G_\psi(x, y) = \phi_5(x)\phi_5(y) G_\psi^\dagger(y, x) \tag{3.32}$$

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

3.5.2 2pt correlation functions

We will break down the correlation function to see what quantities must be computed by the simulation. Consider the generic 2-point correlator:

$$C(x,y) = \langle \Phi_X^\dagger(x)\Phi_Y(y)\rangle_{\psi,U} \quad , \quad \Phi_X(x) = \frac{1}{4}\bar{\psi}_a(x)\gamma_X\psi_b(x) \tag{3.33}$$

$$= \frac{1}{16}\langle tr_{c,s}\gamma_X G_{a,\psi}(x,y)\gamma_Y G_{b,\psi}(y,x)\rangle_U \tag{3.34}$$

$$= \frac{1}{16} tr_s\left(\gamma_x^\dagger\gamma_X\gamma_x\gamma_y^\dagger\gamma_Y\gamma_y\right)\langle tr_c\left(G_{a,\chi}(x,y)G_{b,\chi}(y,x)\right)\rangle_U \tag{3.35}$$

tr_s is a trace over spin and tr_c is over colour. To deal with the spin trace, define the family of phases $\{\phi_X(x)\}$ according to

$$\gamma_x^\dagger\gamma_X\gamma_x = \phi_X(x)\gamma_X \tag{3.36}$$

for example, if $X = 5$, then $\gamma_x^\dagger\gamma_5\gamma_x = (-1)^{\sum_\mu x_\mu}\gamma_x^\dagger\gamma_x\gamma_5 = \phi_5(x)\gamma_5$. The map from X to ϕ_X is structure preserving, i.e. if $\gamma_X = \gamma_A\gamma_B$, then $\phi_X(x) = \phi_A(x)\phi_B(x)$. The spin trace becomes $\phi_X(x)\phi_Y(y)tr_s\left(\gamma_X\gamma_Y\right)$. The will vanish unless $Y = X$, as one would expect physically for the correlation function. We end up with

$$C(x,y) = \frac{1}{4}\phi_X(x)\phi_X(y)\langle tr_c G_{a,\chi}(x,y)G_{b,\chi}(y,x)\rangle_U \tag{3.37}$$

It is useful in the simulation to replace $G_{b,\chi}(y,x)$ with it's conjugate via (3.32), resulting in

$$C(x,y) = \frac{1}{4}\phi_{5X}(x)\phi_{5X}(y)\langle tr_c G_{a,\chi}(x,y)G_{b,\chi}^\dagger(y,x)\rangle_U \tag{3.38}$$

where $\phi_{5X}(x) = \phi_5(x)\phi_X(x)$. To obtain the correlation function of a meson in an eigenstate with momentum $\underline{p}$, the above must be replaced with

$$C_{\underline{p}}(t_0,t) = \frac{1}{L^3}\sum_{\underline{x},\underline{y}} e^{i\underline{p}\cdot(\underline{x}-\underline{y})}C(\underline{x},t_0;\underline{y},t) \tag{3.39}$$

$$= \frac{1}{4L^3}\sum_{\underline{x},\underline{y}} e^{i\underline{p}\cdot(\underline{x}-\underline{y})}\phi_{5X}(x)\phi_{5X}(y)\langle tr_c G_{a,\chi}(x,y)G_{b,\chi}^\dagger(y,x)\rangle_U, \tag{3.40}$$

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

$$P_{a,\underline{p},X}^{t_0}(y) \equiv \frac{1}{\sqrt{L^3}}\sum_{\underline{x}} e^{i\underline{p}\cdot(\underline{x}-\underline{y})}\phi_{5X}(\underline{x},t_0)\xi(\underline{x})G_{a,\chi}(\underline{x},t_0;y), \tag{3.41}$$

where $\xi(\underline{x})$ is a random field of colour vectors, varying with U. This has the property

$$\langle f(\underline{x},\underline{x}')\xi^*(\underline{x}')\xi(\underline{x})\rangle_U = \delta_{\underline{x},\underline{y}}\langle f(\underline{x},\underline{y})\rangle_U. \tag{3.42}$$

Using this property it can be shown that the correlator can be build instead according to

$$C(\underline{x},t_0;\underline{y},t) \simeq \frac{1}{4}\sum_{\underline{y}}\phi_{5X}(y)\langle tr_c P_{a,\underline{p},X}^{t_0}(\underline{y},t)P_{b,\underline{0},1}^{t_0,\dagger}(\underline{y},t)\rangle_U \tag{3.43}$$

Now all the simulation has to do is compute $P^{t_0}_{a/b}(y)$ for general y, so 2·Vol operations, a reduction by a factor of Vol.

In the MILC code, "sources" are first created (the fields $\phi_{5X}(\underline{x}, t_0)\xi(\underline{x})$), then the objects $P^{t_0}(y)$ (referred to as "propagators") are generated from them. Any extra factors dependant on y (this is useful for "smeared" propagators, see ?) can be multiplied in. The resulting object $f(y) \cdot P^{t_0}(y)$ is referred to as a "quark". Finally, two of these quarks can be "tied together" according to (3.43), to produce correlation functions. The sources are chosen to be on some single timeslice t_0, resulting in a value for $C(t_0, t)$ at each t.

3.5.3 3-point correlators

The above discussion can be generalized to 3−(or N−)point correlation functions using *extended sources*. Consider a 3-pt correlator encoding the form-factors of a semileptonic decay from meson X to meson Z, via a current J. We start by evaluating

$$C(x,y,z) = \langle \Phi^\dagger_X(x) J(y) \Phi_Z(z) \rangle_{\psi,U} \quad , \quad \begin{aligned} \Phi_X(x) &= \frac{1}{4}\bar{\psi}_b(x)\gamma_X\psi_s(x) \\ J(y) &= \bar{\psi}_b(y)\gamma_J\psi_a(y) \\ \Phi_Z(z) &= \frac{1}{4}\bar{\psi}_a(z)\gamma_Z\psi_s(z) \end{aligned} \tag{3.44}$$

in the same way as before:

$$C(x,y,z) = \frac{1}{16} tr_s \left(\gamma^\dagger_x\gamma_X\gamma_x\gamma^\dagger_y\gamma_J\gamma_y\gamma^\dagger_z\gamma_Z\gamma_z\right) \langle tr_c G_{b,\chi}(x,y) G_{a,\chi}(y,z) G_{s,\chi}(z,x)\rangle_U \tag{3.45}$$

$$= \frac{1}{4}\phi_{5X}(x)\phi_J(y)\phi_{5Z}(z)\langle tr_c G_{b,\chi}(x,y) G_{a,\chi}(y,z) G^\dagger_{s,\chi}(z,x)\rangle_U \tag{3.46}$$

We have assumed that $tr_s\gamma_X\gamma_J\gamma_Z = 4$, requiring that each gamma matrix in this combination has a partner and therefore cancels. In any other situation the trace would vanish. For example, if the current is a temporal vector $J = 0$, and the two mesons represent pseudoscalars, one of the meson operators must have a γ_0, i.e. one could choose $\gamma_X = \gamma_0\gamma_5, \gamma_Z = \gamma_5$. why is it ok to have a non-goldstone for X?

Putting X into an eigenstate of zero momentum, and Y into an eigenstate of momentum $\underline{p}$, we get

$$C_{\underline{p}}(t_0,t,T) = \frac{1}{4L^3}\sum_{\underline{x},\underline{y},\underline{z}} e^{i\underline{p}\cdot(\underline{y}-\underline{z})}\phi_{5X}(x)\phi_J(y)\phi_{5Z}(z)\langle tr_c G_{b,\chi}(\underline{x},t_0;\underline{y},t) G_{a,\chi}(\underline{y},t,\underline{z},T) G^\dagger_{s,\chi}(\underline{z},T;\underline{x},t_0)\rangle_U \tag{3.47}$$

This can be built by first creating propagators for the b and s quarks: $P^{t_0}_{b,\underline{0},X}(y)$,$P^{t_0}_{s,\underline{0},1}(z)$. Then, build the a propagator using an extended source, i.e.:

$$P^T_{a,\underline{p},ext}(y) = \sum_{\underline{z}} P^{t_0}_{s,\underline{0},1}(\underline{z},T)\phi_{5Z}(\underline{z},T) G_{a,\chi}(\underline{z},T;y) \tag{3.48}$$

Then, by plugging $P^{t_0}_{b,\underline{0},X}(y)$ and $P^T_{a,\underline{p},ext}(y)$ into the MILC tie-together defined by (3.43),one ends up evaluating (3.47).

3.6 Fitting Correlation Functions

Once a correlation function like the in 3.2 has been computed, we can extract physics from it, namely the mass and decay constant of the meson we are studying. In practice the meson creation operators defined above are fourier transformed

$$\Phi(\underline{k},t) = \sum_{\underline{x}} e^{-i\underline{k}\cdot\underline{x}}\Phi(\underline{x},t) \tag{3.49}$$

which serves to change (3.7) into

$$C_{\underline{k}}(t) = \frac{1}{N}\sum_{i}\sum_{\underline{x},\underline{y}} e^{-i\underline{k}\cdot(\underline{x}-\underline{y})}\ \mathrm{Tr}\left[\ M_b^{-1}(\underline{y},t;\underline{x},0)[U_i]\gamma_5 M_d^{-1}(\underline{x},0;\underline{y},t)[U_i]\gamma_5\ \right] \tag{3.50}$$

(3.50) is computed for many t values with a lattice calculation following the principles detailed above. One performs a least-squares fit of this to a theoretically motivated function of t. To derive such a function, first construct a complete set of momentum $\underline{k}$ states with quantum numbers matching the meson:

$$1 = \sum_{n=0}\frac{1}{2E_n^r}|\lambda_n\rangle\langle\lambda_n|. \tag{3.51}$$

Where $E_n^r = \sqrt{M_n^2 + \underline{k}^2}$ are the relativistic energies of each state. Inserting this into the correlation function, and moving from the Heisenberg to Schroedinger picture;

$$\begin{aligned} C_{\underline{k}}(t) &= \sum_{n=0}\frac{1}{2E_n^r}\langle 0|e^{Ht}\Phi(\underline{k},0)e^{-Ht}|\lambda_n\rangle\langle\lambda_n|\Phi^\dagger(\underline{k},0)|0\rangle \\ &= \sum_{n=0}\left(\frac{\langle 0|\Phi(\underline{k},0)|\lambda_n\rangle}{\sqrt{2E_n^r}}\right)\left(\frac{\langle\lambda_n|\Phi^\dagger(\underline{k},0)|0\rangle}{\sqrt{2E_n^r}}\right)e^{-E_n^l t} \\ &\equiv \sum_{n=0}|a_n|^2 e^{-E_n^l t}. \end{aligned} \tag{3.52}$$

The fit results in a determination of the parameters a_n, and E_n^l. Since the lowest energies dominate the function at late times, one can afford to truncate the sum over n to some tractable range, in our case $n \in [1,6]$. We interpret $|\lambda_0\rangle$ to be the ground state of the meson we are studying. The exponential decays mean the fit function is dominated by the ground state at large t, and subsequent excited states become less important as E_n^l increases. Hence by only including $C_{\underline{k}}(t)$ at suitably large t values, we can affort to truncate the sum in n. In our fits we chose $n = 6$.

We maintain a distinction between E^l and E^r, since for example in simulations involving NRQCD quarks these differ. If this is not an issue, as is the case with HISQ, one can compute the correlation function at zero momentum $C_{\underline{0}}(t)$, then fit it to find the parameter E_0^l, which will equal the meson's mass M. Noting the definition of a meson decay constant f: $\langle 0|J_0|\mathrm{Meson}(\underline{k}=0)\rangle = Mf$, where J_0 is a temporal current with the same quantum numbers as the meson, we can see that the fit parameters a_n at zero momentum are related to the meson's decay constant via

$$f = \sqrt{\frac{2}{M}}a_0 \tag{3.53}$$

Hence the fit can also be used to extract decay constants.

The above discussion can be straightforwardly generalized to 3-point correlation functions, from which we are able to extract quantities like the hadronic transition amplitudes $H_\mu = \langle M_{q_1\bar{q}_3}|J_\mu^{q_1\bar{q}_2}|M_{q_2\bar{q}_2}\rangle$ from sec. 2.1. Specifically the quantity we require in order to deduce the $B_s \to D_s l\nu$ form factors is $\langle D_s|V_\mu|B_s\rangle$, where $V_\mu = \bar{c}\gamma_\mu b$.

The generalization of the above for 3pt functions is summarized here:

$$C_3(t,T) = \int \mathcal{D}\psi\mathcal{D}\bar{\psi}\mathcal{D}U \left(\Phi_{D_s}(\underline{0},0)V_\mu(-\underline{p},t)\Phi^\dagger_{B_s}(\underline{p},T) \right) e^{-S[\psi,\bar{\psi},U]} \tag{3.54}$$

$$\simeq \frac{1}{N}\sum_i \sum_{\underline{x},\underline{y},\underline{z}} e^{-i\underline{p}\cdot(\underline{y}-\underline{z})} \text{ Tr}\left[M_b^{-1}(\underline{x},0;\underline{y},t)[U_i]\gamma_\mu M_c^{-1}(\underline{y},t;\underline{z},T)[U_i]\gamma_5\gamma_5 M_s^{-1\dagger}(\underline{z},T;\underline{x},0)[U_i] \right] \tag{3.55}$$

$$= \sum_{n,m}\left(\frac{\langle 0|\Phi_{D_s}|\lambda_n\rangle}{\sqrt{2E_n^r}}\right)\left(\frac{\langle\lambda_n|V_\mu|\lambda_m\rangle}{2\sqrt{E_n^r E_m^r}}\right)\left(\frac{\langle\lambda_m|\Phi^\dagger_{B_s}|0\rangle}{\sqrt{2E_n^r}}\right) e^{-E_m^l(T-t)}e^{-E_n^l t} \tag{3.56}$$

$$\equiv \sum_{n,m} a_{D_s,n}V_{nm}a^*_{B_s,m}e^{-E_m^l(T-t)}e^{-E_n^l t}.$$

$C(t,T)$ is computed at different values of t and T, then a least-squares fit is performed to the fit function (3.57). a_n will vanish for states $|\lambda_n\rangle$ which have different quantum numbers to Φ_{B_s}, and similarly for b_m and Φ_{D_s}. Non-zero a_n's will match the analagous parameters extracted from fitting a 2pt function $\langle\Phi^\dagger_{B_s}\Phi_{B_s}\rangle$, similarly for b_n's and Φ_{D_s}. This carries on to the energies, $\{E_n^l\}$ is the spectrum for the D_s meson, and $\{E_m^l\}$ is the spectrum for the B_s. Therefore, we compute and fit the appropriate 2pt functions to deduce the parameters $\{a_n\}$,$\{b_m\}$,$\{E_n^l\}$. Then, fitting $C_3(t,T)$ results in an accurate determination of the remaining free parameters, V_{nm}. This set contains the quantity we need, recoginising that:

$$V_{00} = \frac{\langle D_s|V_\mu|B_s\rangle}{2\sqrt{E^{B_s}E^{D_s}}} \tag{3.57}$$

3.7 Renormalization of Currents

Once one has computed a matrix element from a lattice calculation, it needs to be translated into a continuum regularization scheme. Suppose we have some bare operator $\mathcal{O}_0$, we expect this to be related to the renormalized operator in $\overline{MS}$ at scale μ, $\mathcal{O}^{\overline{MS}}(\mu)$, via

$$\mathcal{O}^{\overline{MS}}(\mu) = Z^{\overline{MS}}(\mu)\mathcal{O}_0. \tag{3.58}$$

Similarly, in a lattice regularization,

$$\mathcal{O}^{\text{lat}}(1/a) = Z^{\text{lat}}(1/a)\mathcal{O}_0. \tag{3.59}$$

Hence we expect a multiplicative factor between the lattice matrix elements, and the continuum $\overline{MS}$ ones:

$$\langle\mathcal{O}\rangle^{\overline{MS}} = \mathcal{Z}(\mu,1/a)\langle\mathcal{O}\rangle^{\text{lat}} \tag{3.60}$$

where $\mathcal{Z}(\mu,1/a) = Z^{\overline{MS}}(\mu)/Z^{\text{lat}}(1/a)$. These "matching factors" $\mathcal{Z}$ can be deduced by equating observables calculated in both lattice QCD and continuum (appropriately regularized) QCD, producing equations which can be solved for $\mathcal{Z}$. The lattice side of the calculation can be done either through lattice

Set	a/fm	am_l	am_s	am_c	L_x/a	L_t/a	N_{cfg}
1	0.1474(15)	0.013	0.065	0.838	16	48	1020
2	0.1219(9)	0.0102	0.0509	0.635	24	64	1053
3	0.0884(6)	0.0074	0.037	0.440	32	96	1008

Table 1: Parameters for gauge ensembles [5]. a is the lattice spacing, deduced from a study of the Υ-Υ' splitting [12]. L_x is the spacial extent and L_t the temporal extent of the lattice. The masses of up, down, stange and charm quarks included in the sea are given as the dimensionless quantity am. Up and down quarks are taken to have the same mass, am_l. N_{cfg} is the number of configurations in the ensemble. We calculate propagators from 16 time sources on each ensemble to increase statistics.

perturbation theory ("perturbative matching"), or through a simulation ("non-perturbative matching").

It is a well-known result that conserved (or partially conserved) currents do not receive any renormalization in any scheme, i.e. $Z^{\text{any}} = 1$ ("absolutely normalized").

In principle this is of great help, since the currents we are calculating, namely V_μ, are partially conserved, so we are not required to include any matching factors. However in practice, this is complicated by the fact that the conserved current in the lattice theory is often computationally difficult or impossible to compute. For example, in NRQCD, the partially conserved current corresponding to $SU(N)_V$ is an infinite sum in $1/m_b$ where m_b is the bottom mass. The corresponding current in HISQ is also the sum of a large number of operators. This can be interpreted as a mixing in the renormalization:

$$\langle \mathcal{O}_i \rangle^{\overline{MS}} = \mathcal{Z}_{i,j} \langle \mathcal{O}_j \rangle^{\text{lat}} \tag{3.61}$$

In practice, lattice calculations often use only the dominant operators that contribute to the conserved current. Since these will be "close" to the conserved current, one can expect the matching factor to only be a small deviation from unity, and the more sub-dominant operators you add, the overall matching factor should tend towards one.

4 Calculation Details

4.1 Lattice Setup

We follow a largely similar approach to the recent $B \to Dl\nu$ and $B_s \to D_s l\nu$ calculations in [28] and [26].

Table 1 gives details of the ensembles used in this study. These are taken from the MILC HISQ ensembles [5]. They are generated by MCMC with the distribution given in (3.6), where M is the Dirac operator for the HISQ action. They take into accound up, down, strange and charm quarks, assuming the contribution from higher mass quarks in the sea to be negligable.

Following the procedure discussed in section 3.2, first we must compute 2-point correlators for the B_s and D_s. In both cases we use *smeared* creation operators:

$$\Phi^\alpha(\underline{k}, t) = \sum_{\underline{x},\underline{x'}} e^{-i\underline{k}\cdot\underline{x}} \; \bar{\psi}_1(\underline{x'}, t) \phi^\alpha(\underline{x'} - \underline{x}) \gamma_5 \psi_2(\underline{x}) \tag{4.1}$$

Set	am_s	am_c	am_b	u_0	ϵ_{Naik}	$c_{1,6}$	c_5	c_4	$\{T\}$
1	0.0705	0.826	3.297	0.8195	-0.3449	1.36	1.21	1.22	8, 11, 14
2	0.0541	0.645	2.66	0.8340	-0.2348	1.31	1.16	1.20	9, 12, 15
3	0.0376	0.450	1.91	0.8525	-0.1256	1.21	1.12	1.16	14, 19, 24

Table 2: Parameters used in our calculation. am_s and am_c are the bare masses of the strange, charm valence quarks, tuned in [8], am_b is the bare mass of the valence bottom quark, tuned in [12] u_0 is the 'tadpole improvement parameter' as used in [12]. ϵ_{Naik} is the Naik parameter in the HISQ action [14]. $\{c_i\}$ are the coefficients for the kinetic and chromomagnetic terms in the NRQCD action [19]. $\{T\}$ is the set of temporal seperations between source (B_s creation operator) and sink (D_s anihilation operator).

where $\bar{\psi}_1$,ψ_2 are $\bar{b}$,s for the B_s and $\bar{c}$,s for the D_s. The smearing functions ϕ^α approximate the wave-function of one quark in a potential set up by the other, giving a psuedo-realistic representation of the meson. This makes the operator Φ^α couple stronger to the ground state of the meson, maximizing the $a_0 \propto \langle 0|\Phi^\alpha|\lambda_0\rangle$ parameter in the fit (3.52) relative to the excited states $a_{n>0}$. This increases the dominance of the first term in (3.52) resulting in better convergence of the n sum, therefore a better fit and better statistics of the results.

For the B_s, we use a combination of $\phi^0(\underline{z}) = \delta(\underline{z})$ ("local"), and $\phi^{1,2}(\underline{z}) \propto \exp(-|\underline{z}|/r_0)$ ("smeared"), with $r_0 \simeq 2.5\text{fm}, 5\text{fm}$. For the D_s we use one local and one smeared with $r_0 \simeq 2.5\text{fm}$. Correlation functions between each smearing combination is calculated, i.e. for the D_s, we compute 4 correlators $\langle \Phi_D^{0\dagger}\Phi_D^0\rangle$, $\langle \Phi_D^{1\dagger}\Phi_D^0\rangle$, $\langle \Phi_D^{0\dagger}\Phi_D^1\rangle$, $\langle \Phi_D^{1\dagger}\Phi_D^1\rangle$, and 9 for the B_s.

A good way to see the benefit of the smearing is by plotting the effective mass $m_{\text{eff}}(t) = \log(C(t)/C(t+1))$ against t as in fig. 4. The flatness of this as a function in t shows how well it is approximated by the fit function with only the ground state (i.e. not contaminated by excited states), therefore how easily the fit will determine the ground-state energy. As can be seen, The smeared effective mass "flattens out" quicker than the local, as the smeared couples mostly to the ground state.

The 3pt correlation function calculated is $\langle \Phi_{D_s}^{\alpha\dagger} V_\mu \Phi_{B_s}^\beta\rangle$, including all smearings of the B_s and D_s used in the 2-point functions. We must be careful choosing the appropriate operator for V_μ. Recall that NRQCD quarks contain an inverse Fouldy-Wouthuysen (FW) transformation (3.29). This must be encorparated into the V_μ operator, since it couples to the b which obeys the NRQCD action in our simulation. Hence the V_μ operator must be of the form

$$V_\mu = \bar{c}\gamma_\mu \left(1 - \frac{1}{2m_b}\underline{\gamma}\cdot\underline{\nabla} + \mathcal{O}\left(\frac{1}{m_b^2}\right)\right) b \tag{4.2}$$

where we can afford to expand the exponential in the FW transformation since $1/m_b$ is small. However, this is only a tree level result. According to the discussion in sec. 3.7, these terms require renormalization constants, and also new terms in the above expansion may appear. Therefore the full expression for the

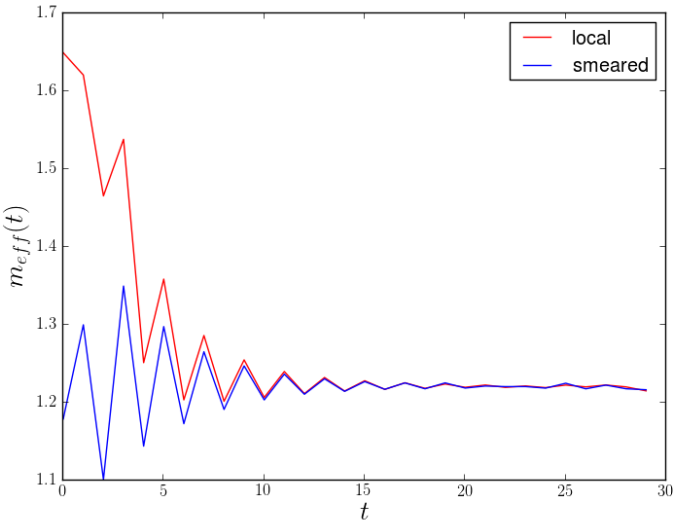

Figure 4: Effective mass of D_s correlation functions, one with local D_s operators $\langle \Phi^0 \Phi^0 \rangle$ and one with smeared $\langle \Phi^1 \Phi^1 \rangle$.

vector current that we use is (accurate to $\mathcal{O}(\alpha_s, 1/m_b)$):

$$\begin{aligned} V_\mu &= (1 + z_0 \alpha_s) \left[V_\mu^{(0)} + (1 + z_1 \alpha_s) V_\mu^{(1)} + z_2 \alpha_s V_\mu^{(2)} \right] \\ V_\mu^{(0)} &= \bar{c} \gamma_\mu b \\ V_\mu^{(1)} &= -\frac{1}{2m_b} \bar{c} \gamma_\mu \underline{\gamma} \cdot \underline{\nabla} b \\ V_\mu^{(2)} &= -\frac{1}{2m_b} \bar{c} \gamma_\mu \underline{\gamma} \cdot \overleftarrow{\underline{\nabla}} b \end{aligned} \tag{4.3}$$

The coefficients $\{z_i\}$ are set by a matching procedure between the HISQ/NRQCD currents and continuum QCD in [25]. We calculate correlation functions $\langle \Phi_{D_s} V_\mu^{(i)} \Phi_{B_s} \rangle$ for each current and combine them after the fitting procedure.

4.2 Correlator Fits

2-point correlation functions are then fitted as in sec. 3.6. The fit function we use is modified a little from (3.52), we use:

$$\begin{aligned} C^{\alpha\beta}(t) = &\sum_n a_n^{\alpha *} a_n^{\beta} (e^{-E_n^l t} - s e^{-E_n^l (T-t)}) \\ &+ \sum_n a_n^{'\alpha *} a_n^{'\beta} (-1)^t (e^{-E_n^{'l} t} - s e^{-E_n^{'l} (T-t)}) \end{aligned} \tag{4.4}$$

Firstly, the parameters $\{a_n\}$ must vary between source and sink to account for the different operators. Secondly, the periodicity of the lattice in the time direction means an extra exponential term is required,

but not in the case of the B_s since NRQCD quarks do not experience the periodicity of the lattice. Hence s is set to 0 for the B_s correlator and 1 for the D_s. T is the time extent of the lattice. The second term is to account for the so-called "oscillating state", which is in fact the $\zeta = (1,0,0,0)$ doubler fermion appearing due to the doubling in the HISQ action (see sec. B.9). No other doublers contribute, since $\Phi_{\underline{k}}$ has a 3-momentum fixed at $\underline{k}$, which we always take to be small relative to π/a, hence does not couple to the states at $k \sim (0, \pi/a, 0, 0)$, $k \sim (0, 0, \pi/a, 0)$ etc. However, $\Phi_{\underline{k}}$ can couple to arbitrarily high energy states, so the $k \sim (\pi/a, 0, 0, 0)$ doubler contributes. The second term in (4.4) is justified by performing the doubling operation $\mathcal{B}_0$ defined in (3.52), the quark fields in $\Phi_{\underline{k}}$ which obey the HISQ action. See appendix G of [14] for details.

We use the *CorrFitter* package [16] for performing Baysian least-squares fitting to the correlation functions. The package employs the trust region method of least-squares fitting. The fits require priors for each of the fit parameters. The "amplitude" parameters $a^{\alpha}_{n,B_s/D_s}$ are given priors of 0.1(1.0), thus inserting only the assumption that they are of $\mathcal{O}(1)$. The ground state energies are given priors motivated by the meson masses, and excited state energies are given loose, evenly spaced priors with 600MeV between each level.

2-point correlation functions for B_s and D_s are fit to (4.4), resulting in $a^{\alpha}_{n,B_s/D_s}$,$E^{l}_{n,B_s/D_s}$. Since the HISQ action is fully relativitstic, E^{l}_{0,D_s} at $\underline{k} = 0$ can be interpreted as the D_s meson mass. The same cannot be done for the B_s. The decay constants for B_s and D_s can be deduced from a^{0}_{0,B_s} and a^{0}_{0,D_s}, since $\Phi^{0}_{B_s,D_s}$ are also temporal axial currents. This is a good avenue for consistency checks, we compared $M_{D,s}$,f_{D_s} and f_{B_s} to those computed in [9],[26] amongst others, and found them to be consistent (modulo small shifts we can reasonably expect due to differing choices of bare quark masses).

We now discuss fitting the 3-point correlation functions. The same considerations as those that went into (4.4) lead us to our 3-point fit function:

$$\begin{aligned} C_3^{\alpha\beta}(t,T) = \sum_{k,m} \big(& a^{\alpha}_{k,D_s} V^{nn}_{km} a^{*\beta}_{m,B_s} e^{-E^{l}_{m} t} e^{-E^{l}_{k}(T-t)} \\ & + a^{\alpha}_{k,D_s} V^{no}_{km} a^{'*\beta}_{m,B_s} e^{-E^{\prime l}_{m} t} e^{-E^{l}_{k}(T-t)} \\ & + a^{'\alpha}_{k,D_s} V^{on}_{km} a^{*\beta}_{m,B_s} e^{-E^{l}_{m} t} e^{-E^{\prime l}_{k}(T-t)} \\ & + a^{'\alpha}_{k,D_s} V^{oo}_{km} a^{'*\beta}_{m,B_s} e^{-E^{\prime l}_{m} t} e^{-E^{\prime l}_{k}(T-t)} \big) \end{aligned} \tag{4.5}$$

The 2-point and 3-point correlators are fit simultaniously, according to fit functions (4.4) and (4.5). The parameters involved in the 2pt fits are mostly fixed by the data in the 2pt correlation functions, so the fit can use most of the data in the 3pt correlation functions to determine the transition amplitudes V^{ab}_{km}. This is carried out for each B_s and D_s smearing, each direction μ and each current correction i of the vector current $V^{(i)}_{\mu}$.

In this large 2pt/3pt fit, there is a huge χ^2 manifold with many local minema, and it is crucial to impose strong priors in order to ensure the fit finds the true minemum. Priors for ground state 2-point amplitudes and energies $a^{\alpha}_{0,B_s/D_s}$,$E^{l}_{0,B_s/D_s}$ are taken to be the results from individual fits of the 2-point functions, with the errors expanded by a factor of 2. The excited state amplitudes and energies are given the same priors as in the 2-point fits. The transition amplitueds V^{ab}_{km} are given the prior 0.1(1.0), assuming it to be $\mathcal{O}(1)$.

Finally, we end up with the sought-after parameters V_{00}^{nn} representing $V_\mu^{(i)}$, via the relation

$$\langle D_s|V_\mu^{(i)}|B_s\rangle = 2\sqrt{M^{B_s}E^{D_s}}V_{00}^{nn}|_{V=V_\mu^{(i)}\text{in simulation}} \tag{4.6}$$

We have asserted the ground state of B_s to be it's mass, but not the D_s, as we give D_s spacial momenta in the calculation (to be expanded on in the next section). Then, the full vector currents $\langle D_s|V_\mu^{(i)}|B_s\rangle$ can be build from a linear combination of these according to 4.3.

4.3 Deduction of Form Factors

We wish to use these currents to deduce the $B_s \to D_s l\nu$ form factors discussed in section 2.4 over the full q^2 range. To this end, the above procedure to compute $\langle D_s|V_\mu^{(i)}|B_s\rangle$ is repeated while giving the D_s meson a range of different spacial momenta. We chose spacial momentum $\underline{p} = |\underline{p}|(1,1,1)$, in this each direction is equivilant, and it allows us to average V_μ over the 3 spacial directions, reducing work in the fit and increasing the statistics of the averaged current $V_k \equiv \sum_{i=1}^{3} V_i/3$.

As discussed in appendix C, the most accurate results will come from $|\underline{p}| = 0$, then the correlators become noise-dominated as one increases $|\underline{p}|$ towards $|\underline{p}|_{\max}$. Hence, the approach is to start at the $|\underline{p}| = 0$ end and move up in momentum, monitoring statistical errors as you go. So far we have performed runs for $|a\underline{p}| = 0, 0.30, 0.45$ on the fine ensemble, and $|a\underline{p}| = 0, 0.25, 0.50$ on the coarse ensemble. Statistical errors have not yet become uncontrollable at these momenta, but from experience of previous calculations, they are expected to become problematic at $|a\underline{p}| \sim 0.7$.

With values for $\langle D_s|V_\mu|B_s\rangle$ at varying $\underline{p}$ therefore varying q^2, we can perform a fit of this data in order to extract $f_{0,+}(q^2)$ via (2.17). To do the fit, we need some anzats for the functional form of $f_{0,+}(q^2)$. We use the BCL parameterization [7]. This involves first reparameterizing q^2 to

$$z(q^2) = \frac{\sqrt{t_+ - q^2} - \sqrt{t_+ - t_0}}{\sqrt{t_+ - q^2} + \sqrt{t_+ - t_0}} \tag{4.7}$$

where we take $t_0 = t_+(1 - \sqrt{1 - t_-/t_+})$, and $t_\pm = (M_{B_s} \pm M_{D_s})^2$, as in [20]. $z(q^2)$ has a very small magnitude throughout the entire q^2 range, in our case $|z|_{\max} \sim 0.032$. We can then accurately model $f_{+,0}$ as a series expansion in z:

$$f_{0,+}(q^2) = \frac{1}{P_{0,+}(q^2)} \sum_n a_n^{0,+} z(q^2)^n \tag{4.8}$$

we truncate this at z^2, adding further terms have no effect on the fit. The so-called Blaschke factors $P(q^2)$ are defined by

$$P_{0,+}(q^2) = \left(1 - \frac{q^2}{M_{0,+}^2}\right) \tag{4.9}$$

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

4.3.1 Using only Scalar and Temporal Vector Currents

The continuum currents V_0 and S are given by

$$V_0 = (1 + \alpha_s(\eta_0^{V_0} - \tau_{01}^{V_0})) \times \left(J_{V_0}^{(0)} + J_{V_0}^{(1)}\right) \tag{4.10}$$

$$m_b S = (1 + \alpha_s(\eta_0^{V_0} + \tau_{01}^{V_0} + C_M)) \times \frac{am_b}{a} \times \left(J_{V_0}^{(0)} - J_{V_0}^{(1)}\right) \tag{4.11}$$

The scalar current $F_0(q^2)$ can be extracted from $m_b S$ via

$$m_b \left(1 - \frac{m_c}{m_b}\right) \langle D|S|B\rangle = (M_B^2 - M_D^2) F_0(q^2) \tag{4.12}$$

Once $F_0(q^2)$ is constrained, all the information in V_0 can be used to constrain $F_+(q^2)$:

$$\begin{aligned} \langle D|V_0|B\rangle = \left(p_B^0 + p_D^0 - \frac{M_B^2 - M_D^2}{q^2} q^0\right) F_+(q^2) \\ + \frac{M_B^2 - M_D^2}{q^2} q^0 F_0(q^2) \end{aligned} \tag{4.13}$$

Let's write these relations as

$$m_b S \equiv R_{S0} F_0 \tag{4.14}$$

$$V_0 \equiv R_{0+} F_+ + R_{00} F_0 \tag{4.15}$$

To deduce F_0 we use (4.14), then to deduce F_+ via:

$$F_+ = \frac{V_0 - R_{00} F_0}{R_{0+}} \tag{4.16}$$

Close to q^2_{max}, there is something of a "fine tuning problem" in this equation. Consider the size of each of these terms:

$$\mathcal{O}(1) = \frac{\mathcal{O}(10) - \mathcal{O}(10)}{\mathcal{O}(0.1)} \tag{4.17}$$

V_0 and $R_{00}F_0$ must cancel to a high precision in order to reproduce an F_+ of $\mathcal{O}(1)$. The correlation between V_0 and S is smaller than I expected, $\sim$0.5, so this does not improve much the precision of $V_0 - R_{00}F_0$. Even if it did, any error in matching factors will cause a large difference between V_0 and $R_{00}F_0$.

At the moment, I am getting

$$F_+ \sim 10 \tag{4.18}$$

at the lowest non-zero momentum point on very coarse, while we expect $F_+ \sim 1$.

A possible direction to go in is to define something like

$$\begin{aligned} V_\alpha &\equiv V_0 - \alpha_s V_k \\ &= (R_{00} - \alpha_s R_{k0}) F_0 + (R_{0+} - \alpha_s R_{k+}) F_+ \\ &\equiv R_{\alpha 0} F_0 + R_{\alpha +} F_+ \end{aligned} \tag{4.19}$$

V_α does not contain the V_k loop corrections we were worried about, but adds some information from it's leading order. Then F_+ can be deduced instead from

$$F_+ = \frac{V_\alpha - R_{\alpha 0} F_0}{R_{\alpha +}} \tag{4.20}$$

which suffers from less of a tuning problem, via this method I get

$$F_+ \sim 4 \tag{4.21}$$

Flipping the sign of $J^{(1)}_{V_0}$ seems to ease this problem further, are we confident that the sign we are using is correct? With the sign flipped, we get to

$$F_+ \sim 2 \tag{4.22}$$

Still not at a point where this is useful, but there could be other improvements to be made...

4.4 Continuum & Chiral Extrapolation

The form factors $f_{0,+}(q^2)$ we deduce from the above will contain systematic errors due to 1) discretization, and 2) unphysical quark masses and mistunings.

2) requires a little explaination. In lattice simulations it is computationally expedient (and sometimes neccesary) to take the up/down quark masses m_l to be much larger than their physical values. This is because the condition number of the Dirac operator M_l is proportional to am_l, a small condition number makes it difficult or impossible to numerically invert M_l to obtain the propagator M_l^{-1} as part of the process in section 3.2. Fortunately, since the $B_s \to D_s l\nu$ decay involves no up/down valence quarks, we mostly need not worry about this problem. There is also the issue of mistuning: the quark masses we use are tuned by a calculation of some process, see caption in table 2. The uncertainty in these determinations must be accounted for somehow.

The above issues are in general dealt with by computing the above form factors at a number of lattice spacings and quark masses, and results extrapolated to $a \to 0$ and $m \to m_{\text{physical}}$. The $a \to 0$ extrapolation is performed by involving data from all ensembles in the fit to $f_{0,+}$, and modifying (4.8):

$$a_n^{0,+} \to a_n^{0,+} \times (1 + b_n^{0,+}(am_c)^2) \tag{4.23}$$

where $b_n^{0,+}$ are new fit parameters, and m_c is the charm mass. $am_c \to 0$ in the continuum limit, and, since the charm mass is the largest mass parameter involved in our calculation, it serves as a good order parameter for discretization effects. The extrapolation in masses has not yet been implemented in this work.

5 Preliminary Results

Figure 5 shows form factors for $B \to D$ and $B_s \to D_s$ respectively, extraced from the continuunm vector current constructed as in eq. (4.3), along with our preliminary kinematic extrapolation.

The subleading lattice currents (those of $\mathcal{O}(\alpha_s, \Lambda_{\text{QCD}}/m_b)$ that we neglect in our continuum current) in the temporal direction are all smaller than $\sim 5\%$ of the leading current $J^{(0)}_{0,(s)}$. Some subleading

Set	$\lvert a\vec{p}_D\rvert$	$\langle J_0^{(0)}\rangle/\sqrt{M_B}$	$\langle J_0^{(1)}\rangle/\sqrt{M_B}$	$\langle J_k^{(0)}\rangle/\sqrt{M_B}$	$\langle J_k^{(1)}\rangle/\sqrt{M_B}$
2	0.00	2.188(16)	-0.0056(92)	0.0	0.0
	0.36	2.164(16)	-0.0310(95)	0.1965(41)	-0.0013(33)
	0.51	2.140(27)	-0.028(16)	0.2746(72)	-0.0077(52)
	1.03	2.12(10)	-0.041(26)	0.499(45)	0.005(28)
3	0.00	1.8978(60)	-0.0258(26)	0.0	0.0
	0.30	1.8266(78)	-0.0303(38)	0.1889(38)	-0.0045(16)
	0.45	1.757(14)	-0.0279(60)	0.2694(69)	-0.0068(34)

Table 3: NRQCD-HISQ currents calculated in our simulation, for the $B \to D$ case. $\langle\mathcal{O}\rangle$ denotes $\langle D(p)|\mathcal{O}|B(0)\rangle$. Errors are statistical. The spatial currents are averaged over spatial directions to produce $J_k^{(i)}$.

Set	$\lvert a\vec{p}_{D_s}\rvert$	$\langle J_{0,s}^{(0)}\rangle/\sqrt{M_{B_s}}$	$\langle J_{0,s}^{(1)}\rangle/\sqrt{M_{B_s}}$	$\langle J_{k,s}^{(0)}\rangle/\sqrt{M_{B_s}}$	$\langle J_{k,s}^{(1)}\rangle/\sqrt{M_{B_s}}$
1	0.00	2.6016(90)	-0.0097(41)	0.0	0.0
	0.25	2.5888(81)	-0.0085(24)	0.1383(17)	-0.0031(12)
	0.50	2.5767(91)	-0.0117(64)	0.2631(35)	-0.0057(30)
	0.75	2.584(17)	-0.0164(30)	0.367(10)	-0.0089(91)
	1.00	2.695(46)	0.014(50)	0.444(37)	-0.038(31)
2	0.00	2.3713(61)	-0.0157(25)	0.0	0.0
	0.25	2.3437(75)	-0.0175(17)	0.1472(14)	-0.0041(12)
	0.50	2.2961(91)	-0.0213(23)	0.2778(27)	-0.0082(20)

Table 4: NRQCD-HISQ currents calculated in our simulation, for the $B_s \to D_s$ case. $\langle\mathcal{O}\rangle$ denotes $\langle D_s(p)|\mathcal{O}|B_s(0)\rangle$. Errors are statistical. The spatial currents are averaged over spatial directions to produce $J_{k,s}^{(i)}$.

spatial currents however are relatively large. Namely, these are $J_{\mu,(s)}^{(2)} = -\bar{c}\gamma_\mu\gamma\cdot\overleftarrow{\nabla}b/2am_b$ and $J_{\mu,(s)}^{(4)} = -\bar{c}\overleftarrow{\nabla}_\mu b/2am_b$. This may cause large matching errors due to their neglection in our continuum current.

Figure 5 shows the form factors extracted from these expectation values. As can be seen here, both in $B \to D$ and $B_s \to D_s$, there are large differences between ensembles entering below $q^2 \sim 8\text{GeV}^2$ ($(ap_{D_{(s)}})^2 \sim 0.5$). These could be at least partially explained by discretization effects. They could also be partially explained by the matching errors associated with negating the large subleading currents since some of these are expected to grow with $\left|a\vec{p}_{D_{(s)}}\right|$.

A useful check of how well the discretization errors can be controlled as we add larger $|a\vec{p}_{D_{(s)}}|$ is to compute the speed of light $c = (E_{D_s}(\vec{p}_{D_s})^2 - M_{D_s}^2)/\vec{p}_{D_s}^2$ using our data. We can show that c^2 tends to unity in the continuum limit ($am_c, |a\vec{p}_{D_s}| \to 0$), see fig. 6 (see also [**?**]).

References

[1] Roel Aaij et al. Measurement of the ratio of branching fractions $\mathcal{B}(\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau)/\mathcal{B}(\bar{B}^0 \to D^{*+}\mu^-\bar{\nu}_\mu)$. *Phys. Rev. Lett.*, 115(11):111803, 2015. [Erratum: Phys.

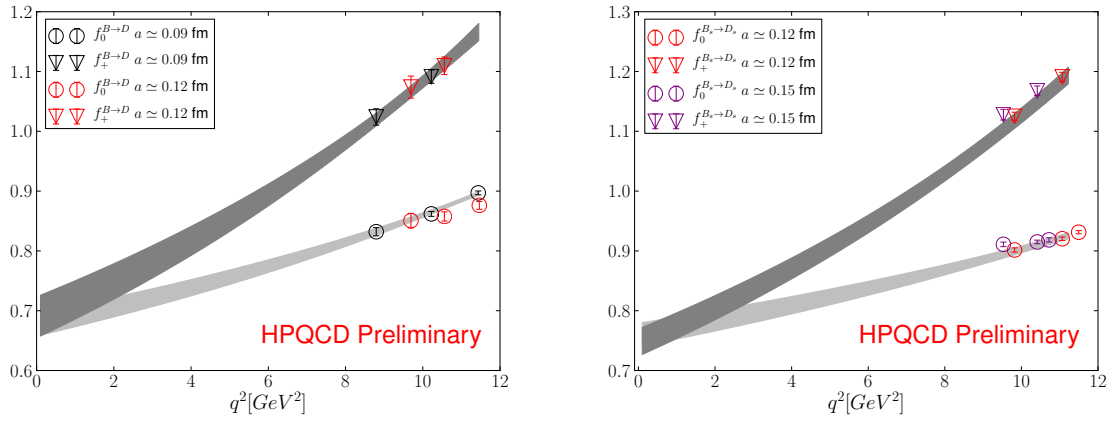

Figure 5: Form factors for the $B \to D$ (left) and $B_s \to D_s$ (right) cases. Errors on data points are statistical. The bands are produced from a fit to the data using (4.8) as the fit form.

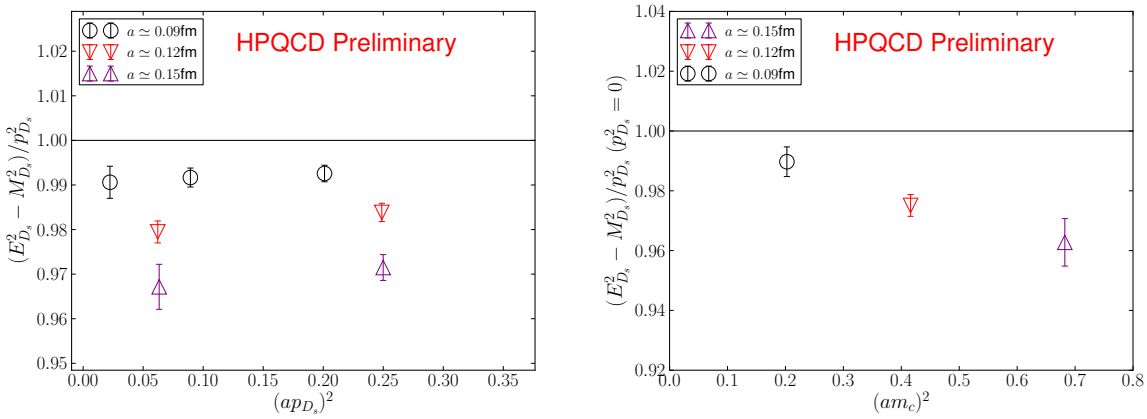

Figure 6: Left: The speed of light against D_s momentum. Discretization errors cause this to differ from unity. Right: On each ensemble, the speed of light at $|a\vec{p}_{D_s}| = 0$ is extrapolated from the $a\vec{p}_{D_s} \neq 0$ data (below $(a\vec{p}_{D_s})^2 = 0.5$), and plotted here. As can be seen, in the limit $am_c \to 0$, the speed of light tends towards unity, showing that discretization effects are controllable below $(a\vec{p}_{D_s})^2 = 0.5$.

Rev. Lett.115,no.15,159901(2015)].

[2] Wolfgang Altmannshofer, Peter Stangl, and David M. Straub. Interpreting Hints for Lepton Flavor Universality Violation. 2017.

[3] S. Aoki et al. Review of lattice results concerning low-energy particle physics. *Eur. Phys. J.*, C77(2):112, 2017.

[4] Gunnar S. Bali, Bernhard Lang, Bernhard U. Musch, and Andreas Schäfer. Novel quark smearing for hadrons with high momenta in lattice QCD. *Phys. Rev.*, D93(9):094515, 2016.

[5] A. Bazavov et al. Lattice QCD ensembles with four flavors of highly improved staggered quarks. *Phys. Rev.*, D87(5):054505, 2013.

[6] Bugra Borasoy. Introduction to Chiral Perturbation Theory. *Springer Proc. Phys.*, 118:1–26, 2008.

[7] Claude Bourrely, Laurent Lellouch, and Irinel Caprini. Model-independent description of $b \to \pi l \nu$ decays and a determination of $|V_{ub}|$. *Phys. Rev. D*, 79:013008, Jan 2009.

[8] Bipasha Chakraborty, C. T. H. Davies, B. Galloway, P. Knecht, J. Koponen, G. C. Donald, R. J. Dowdall, G. P. Lepage, and C. McNeile. High-precision quark masses and qcd coupling from $n_f = 4$ lattice qcd. *Phys. Rev. D*, 91:054508, Mar 2015.

[9] B. Colquhoun, C. T. H. Davies, R. J. Dowdall, J. Kettle, J. Koponen, G. P. Lepage, and A. T. Lytle. B-meson decay constants: a more complete picture from full lattice QCD. *Phys. Rev.*, D91(11):114509, 2015.

[10] Brian Colquhoun, Christine Davies, Jonna Koponen, Andrew Lytle, and Craig McNeile. B_c decays from highly improved staggered quarks and NRQCD. *PoS*, LATTICE2016:281, 2016.

[11] Thomas DeGrand and Carleton E. Detar. *Lattice methods for quantum chromodynamics.* 2006.

[12] R. J. Dowdall et al. The Upsilon spectrum and the determination of the lattice spacing from lattice QCD including charm quarks in the sea. *Phys. Rev.*, D85:054509, 2012.

[13] Leslie L. Foldy and Siegfried A. Wouthuysen. On the dirac theory of spin 1/2 particles and its non-relativistic limit. *Phys. Rev.*, 78:29–36, Apr 1950.

[14] E. Follana, Q. Mason, C. Davies, K. Hornbostel, G. P. Lepage, J. Shigemitsu, H. Trottier, and K. Wong. Highly improved staggered quarks on the lattice, with applications to charm physics. *Phys. Rev.*, D75:054502, 2007.

[15] Christof Gattringer, Meinulf Gockeler, Paul E. L. Rakow, Stefan Schaefer, and Andreas Schaefer. A Comprehensive picture of topological excitations in finite temperature lattice QCD. *Nucl. Phys.*, B618:205–240, 2001.

[16] G.P.Lepage. Corrfitter, 2012.

[17] A. Gray, I. Allison, C. T. H. Davies, Emel Dalgic, G. P. Lepage, J. Shigemitsu, and M. Wingate. The Upsilon spectrum and m(b) from full lattice QCD. *Phys. Rev.*, D72:094507, 2005.

[18] Heavy Flaour Averaging Group. Rd plane, 2015.

[19] T. C. Hammant, A. G. Hart, G. M. von Hippel, R. R. Horgan, and C. J. Monahan. Radiative improvement of the lattice nonrelativistic QCD action using the background field method with applications to quarkonium spectroscopy. *Phys. Rev.*, D88(1):014505, 2013. [Erratum: Phys. Rev.D92,no.11,119904(2015)].

[20] Richard J. Hill. The Modern description of semileptonic meson form factors. *eConf*, C060409:027, 2006.

[21] Karl Jansen. Domain wall fermions and chiral gauge theories. *Phys. Rept.*, 273:1–54, 1996.

[22] G. P. Lepage. Lattice QCD for novices. In *Strong interactions at low and intermediate energies. Proceedings, 13th Annual Hampton University Graduate Studies, HUGS'98, Newport News, USA, May 26-June 12, 1998*, pages 49–90, 1998.

[23] G. P. Lepage, B. Clark, C. T. H. Davies, K. Hornbostel, P. B. Mackenzie, C. Morningstar, and H. Trottier. Constrained curve fitting. *Nucl. Phys. Proc. Suppl.*, 106:12–20, 2002. [,12(2001)].

[24] G. Peter Lepage, Lorenzo Magnea, Charles Nakhleh, Ulrika Magnea, and Kent Hornbostel. Improved nonrelativistic QCD for heavy quark physics. *Phys. Rev.*, D46:4052–4067, 1992.

[25] Christopher Monahan, Junko Shigemitsu, and Ron Horgan. Matching lattice and continuum axial-vector and vector currents with nonrelativistic QCD and highly improved staggered quarks. *Phys. Rev.*, D87(3):034017, 2013.

[26] Christopher J Monahan, Heechang Na, Chris M Bouchard, G Peter Lepage, and Junko Shigemitsu. $B_s \to D_s \ell \nu$ Form Factors and the Fragmentation Fraction Ratio f_s/f_d. 2017.

[27] G. Munster and M. Walzl. Lattice gauge theory: A Short primer. In *Phenomenology of gauge interactions. Proceedings, Summer School, Zuoz, Switzerland, August 13-19, 2000*, pages 127–160, 2000.

[28] Heechang Na, Chris M. Bouchard, G. Peter Lepage, Chris Monahan, and Junko Shigemitsu. $B \to Dl\nu$ form factors at nonzero recoil and extraction of $|V_{cb}|$. *Phys. Rev.*, D92(5):054510, 2015. [Erratum: Phys. Rev.D93,no.11,119906(2016)].

[29] R. Narayanan. Tata lectures on overlap fermions. 2011.

[30] Michael E. Peskin and Daniel V. Schroeder. *An Introduction to quantum field theory*. 1995.

[31] Jeffrey D. Richman and Patricia R. Burchat. Leptonic and semileptonic decays of charm and bottom hadrons. *Rev. Mod. Phys.*, 67:893–976, 1995.

[32] Matthew D. Schwartz. *Quantum Field Theory and the Standard Model*. Cambridge University Press, 2014.

[33] Matthew Wingate. $|V_{cb}|$ using lattice QCD. In *9th International Workshop on the CKM Unitarity Triangle (CKM2016) Mumbai, India, November 28-December 3, 2016*, 2017.

A Chiral Currents and Ward Identities

The currents V_μ and A_μ are conserved in the Chiral limit (quark masses $\to 0$). This can be shown via Noether's theorem: consider a theory containing a number of fields $\{\varphi_i\}$, which transform under a group G by

$$\varphi_i \to \varphi_i(x) - i\epsilon_a(x) F_{i,a}[\{\varphi\}] \tag{A.1}$$

It can be shown that there exists currents

$$J^\mu_a = -i \frac{\partial \mathscr{L}}{\partial \partial_\mu \varphi_i} F_{i,a} \tag{A.2}$$

which obey

$$J_a^\mu = \frac{\partial \delta \mathscr{L}}{\partial \partial_\mu \epsilon_a} \tag{A.3}$$

$$\partial_\mu J_a^\mu = \frac{\partial \delta \mathscr{L}}{\delta \epsilon_a} \tag{A.4}$$

where $\delta\mathscr{L}$ is the result of an infinitesimal G operation, i.e. $G : \mathscr{L} \to \mathscr{L} + \delta\mathscr{L}$. From (A.4), we see that if the theory is symmetric under the generator parameterised by ϵ_a, then J_a^μ is a conserved current.

In the chiral limit, QCD with N flavors has the accidental global symmetry ("chiral symmetry") $U(N)_L \times U(N)_R$:

$$q_{L/R} \to \exp(-i\theta_a^{L/R}\lambda_a) q_{L/R} \tag{A.5}$$

where q is a vector in flavor space, and λ_a are the generators of $U(N)$ in the fundemental representation and acts on flavor. Applying (A.3) and (A.4) to QCD we get the conserved currents:

$$J_{L/R}^{\mu,a} = \bar{q}_{L/R}\gamma^\mu \lambda_a q_{L/R} \tag{A.6}$$

$$\partial_\mu J_{L/R}^{\mu,a} = 0 \tag{A.7}$$

$J_{L/R}$ are often expressed instead in terms of *vector* and *axial* currents

$$V^{\mu,a} = J_L^{\mu,a} + J_R^{\mu,a} = \bar{q}\gamma^\mu \lambda_a q \tag{A.8}$$

$$A^{\mu,a} = J_L^{\mu,a} - J_R^{\mu,a} = \bar{q}\gamma^\mu \gamma_5 \lambda_a q \tag{A.9}$$

which are conserved currents corresponding to the vector and axial symmetries $U(N)_V$ and $U(N)_A$, consisting of simultanious L/R transformations above with constraints $\theta_a^L = \pm\theta_a^R$. Vector and axial currents between any two individal flavours, i.e.

$$V_{ij}^\mu = \bar{q}_i \gamma^\mu q_j \tag{A.10}$$

$$A_{ij}^\mu = \bar{q}_i \gamma^\mu \gamma_5 q_j \tag{A.11}$$

are also conserved, since they can be constructed from linear combinations of $V^{\mu,a}$ and $A^{\mu,a}$.

$U(N)$ can be broken up into $SU(N) \times U(1)$, where $U(1)$ is a singlet transformation, i.e., of the form of (A.5) with $\lambda = 1$. When QCD is quantized, it develops an anomaly which breaks the singlet axial symmetry. This reduces the symmetry group to $SU(N)_V \times SU(N)_A \times U(1)_V$, and prevents the corresponding axial singlet current $A^{\mu,0}$ from being conserved.

In the case of non-zero quark mass, the chiral symmetry is broken. This leads to (A.4) for the vector and axial currents evaluating instead as

$$\partial_\mu V^{\mu,a} = i\bar{q}[M, \lambda_a]q \tag{A.12}$$

$$\partial_\mu A^{\mu,a} = i\bar{q}\{M, \lambda_a\}q \tag{A.13}$$

where M is the quark mass matrix in flavor space. These are the parially conserved axial and vector current identities (PCAC and PCVC). By taking a linear combination of these equations, one can derive PCAC and PCVC for individual flavors:

$$\partial_\mu V_{ij}^\mu = i(m_i - m_j)\bar{q}_i q_j \equiv i(m_i - m_j)S_{ab} \tag{A.14}$$

$$\partial_\mu A_{ij}^\mu = i(m_i - m_j)\bar{q}_i \gamma_t q_j \equiv i(m_i - m_j)P_{ij} \tag{A.15}$$

We refer to S as the scalar current and P as the pseudoscalar. These identities translates straightforwardly to expectation values:

$$(p_1 - p_2)_\mu \langle \psi(p_2)|V^\mu_{ij}|\phi(p_1)\rangle = (m_i - m_j)\langle \psi(p_2)|S_{ij}|\phi(p_1)\rangle \tag{A.16}$$

$$(p_1 - p_2)_\mu \langle \psi(p_2)|A^\mu_{ij}|\phi(p_1)\rangle = (m_i - m_j)\langle \psi(p_2)|P_{ij}|\phi(p_1)\rangle \tag{A.17}$$

where $|\psi(p)\rangle$ and $|\phi(p)\rangle$ are arbitrary states containing momentum p.

These are powerful tools for connecting expectation values of different currents. For example, by combining (A.16) and (2.17), we see that:

$$\langle P_2(p_2)|S_{ij}|P_1(p_1)\rangle = \frac{M_1^2 - M_2^2}{m_i - m_j} f_0(q^2) \tag{A.18}$$

Hence we have an extra route to accessing f_0. In a lattice simulation which computes expectation values of V^μ, one could constrain f_0 by also computing the scalar current on the lattice, then use the rest of the information in V^μ to constrain f_+.

B The Doubling Problem

The interacting Dirac action is most naively discretised with

$$S_F = \sum_{x,\mu} \bar{\psi}(x)\gamma_\mu \nabla_\mu \psi(x) + m \sum_x \bar{\psi}(x)\psi(x), \tag{B.1}$$

where ∇_μ is the gauge covariant finite difference operator,

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

An issue arises with fermions on the lattice, known as the doubling problem. S_F is invariant under a so-called "doubling symmetry", which is generated by

$$\psi(x) \to \mathcal{B}_\mu \psi(x) \equiv (-1)^{x_\mu/a}\gamma_{5\mu}\psi(x) \tag{B.3}$$

$$\bar{\psi}(x) \to \bar{\psi(x)}\mathcal{B}^\dagger_\mu \equiv (-1)^{x_\mu/a}\bar{\psi}(x)\gamma^\dagger_{5\mu} \tag{B.4}$$

where $\gamma_{5\mu} = i\gamma_\mu\gamma_5$. The product space of these form a group of 16 elements $\{\mathcal{B}_\zeta\}$, labeled by vectors ζ with $\zeta_\mu \in \mathbb{Z}_2$ (e.g.the element $\mathcal{B}_0\mathcal{B}_1$ is labeled by $\zeta = (1, 1, 0, 0)$).

The physical signifiance of this symmetry can be seen when we study it's effect on the action. First, notice that

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

$$= \gamma_{5\mu} \sum_k \tilde{\psi}(k - \frac{\pi}{a}\hat{\mu}) e^{ik\cdot x} \tag{B.6}$$

where k is a set of discrete 4-momenta. The action in momentum space can be written as

$$S = \sum_k \bar{\tilde{\psi}}(k) M(k) \tilde{\psi}(k) \tag{B.7}$$

after the operation of $\mathcal{B}_\mu$ it becomes

$$S \to \sum_k \bar{\tilde{\psi}}(x)(k)\gamma_{5\mu}M(k+\frac{\pi}{a}\hat{\mu})\gamma_{5\mu}\tilde{\psi}(k) \tag{B.8}$$

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

$$M^{-1}(k+\frac{\pi}{a}\hat{\mu}) = \gamma_{5\mu}M^{-1}(k)\gamma_{5\mu} \tag{B.9}$$

M^{-1} is the propagator for the fermion field, so B.9 shows that the spectrum of the fermion is periodic, with a period of π/a. We expect a pole in $M^{-1}(k)$ where $k \sim m$, where m is the pole mass of the fermion, but there will now be a second pole at $m + \pi/a$. This will be around the natural cutoff imposed by the lattice $1/a$, and any higher poles like $m+2\pi/a$ is far above the cutoff so will not contribute.

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

$$M^{-1}(k+\frac{\pi}{a}\zeta) = \gamma_{5\mu}M^{-1}(k)\gamma_{5\mu} \tag{B.10}$$

leading to 16 poles in the fermion specturm, therefore 16 distinct excitations (called *tastes*).

We now show that there are only 4 tastes in the staggered quark formalism. One can isolate a single taste by a block-scaling procedure:

$$\psi^{(\zeta)}(x_B) = \sum_{\delta x_\mu \in \mathbb{Z}_2} \mathcal{B}_\zeta \psi(x_B + \delta x) \tag{B.11}$$

For example, for $\zeta = 0$, it would only contain the original non-doubler taste, since all other poles at $|k| \sim \psi/a$ have been integrated out. For $\zeta \neq 0$, the $\mathcal{B}_\zeta$ operator pushes the ζ doubler to where the $\zeta = 0$ taste originally was in k space, then the blocking procedure integrates out the rest. Now writing this isolated taste in terms of χ we arrive at:

$$\psi^{(\zeta)}(x_B) = \sum_{\delta x_\mu \in \mathbb{Z}_2} \Omega(\delta x)\mathcal{B}_\zeta(0)\chi(x+\delta x) \tag{B.12}$$

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

C Signal/Noise Ratio

One of the main obstacles in our calculation is the *signal degredation* of correlation functions computed on the lattice.

A random variable x has mean and standard deviation

$$\hat{x} = \langle x \rangle \tag{C.1}$$

$$\sigma^2 = \frac{1}{N}(\langle x^2 \rangle - \langle x \rangle^2), \tag{C.2}$$

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

$$\frac{\hat{x}^2}{\sigma^2} = N\left(\frac{\langle x^2\rangle}{\langle x\rangle^2} - 1\right)^{-1}. \tag{C.3}$$

Consider 2 point correlators where $x = \Phi^\dagger(t)\Phi(0)$, where Φ is a meson operator with zero spacial momentum.

$\langle x^2\rangle$ and $\langle x\rangle$ can be written as

$$\langle x\rangle = \sum_k \frac{1}{2E_n}\langle 0|\Phi^\dagger(t)|\lambda_n\rangle\langle\lambda_n|\Phi(0)|0\rangle e^{-E_n t} \simeq_{t\to\infty} e^{-E_0 t} \tag{C.4}$$

$$\langle x^2\rangle = \sum_n \frac{1}{2E_n}\langle 0|\Phi^{\dagger 2}(t)|\lambda_n\rangle\langle\lambda_n|\Phi^2(0)|0\rangle e^{-E_n t} \simeq_{t\to\infty} e^{-E_0' t} \tag{C.5}$$

where we have assumed the ratio of matrix elements and energies are $\mathcal{O}(1)$. The two ground state energies E_0 and E_0' need not be the same, since the lowest states for which $\langle\lambda_n|\Phi(0)|0\rangle \neq 0$ and $\langle\lambda_n|\Phi^2(0)|0\rangle \neq 0$ may differ.

The operator Φ^2 will contain two quark and two antiquark operators, connected by some matrices in spin space. Φ^2 can create a combination of all possible 2 meson states where the mesons are made of the available quark species, and quantum numbers. For example, If Φ is a pion, Φ^2 is a 2 pion state, and $E_0' = 2m_\pi$. If Φ is a D_s meson, then $E_0' = m_{\tilde{\pi}} + m_{\eta_c}$ ($\tilde{\pi}$ is a pseudoscalar $s\bar{s}$ state).

Define $\mu_0 = E_0'/2$. Then

$$\frac{\hat{x}^2}{\sigma^2} \simeq N\left(e^{-2(\mu_0 - m_\Phi)t} - 1\right)^{-1} \tag{C.6}$$

In the case of pions, $\mu_0 = m_\Phi$, the ratio becomes simply $\sim N$. For mesons heavier than the pion, $\mu_0 < m_\Phi$, so at large times $e^{-2(\mu_0 - m_\Phi)t} >> 1$, and upon taylor expanding the inverse of this phase we arrive at

$$\frac{\hat{x}}{\sigma} \simeq \sqrt{N} e^{-(m_\varphi - \mu_0)t} \tag{C.7}$$

From this we see there are 3 variables which effect the quality of the signal:

1. The size of the sample N.
2. At large t, the correlators undergo *signal degredation*, i..e, become dominated by noise.
3. The degree of signal degradation is decided by $m_\varphi - \mu_0$. Heavier mesons will tend to experience more signal degredation.

Relevant to our calculation is how giving mesons non-zero spacial momenta $\underline{p}$ can exaserbate this problem. In this case, m_Φ in (C.7) is replaced with $\sqrt{m_\Phi^2 + \underline{p}^2}$. As $\underline{p}$ increases, the signal/noise ratio will degrade more and more and statistics will suffer.

In the $B_s \to D_s l\nu$ calculation, to deduce form factors over the whole range of q^2 values, we need to simulate the process with the D_s having a range of momenta $0 < |\underline{p}| < 2.32$GeV, as discussed in sec. 2.4. Correlation functions at the higher end of this range may be too noisy for any meanigful results to be extracted. We are investigating ways of taming this problem, see sec. **??**.